Journal of Vacuum Science & Technology B

JVST B

Second Series
Volume 16, Number 4
July/August - 1998

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

19981231 117

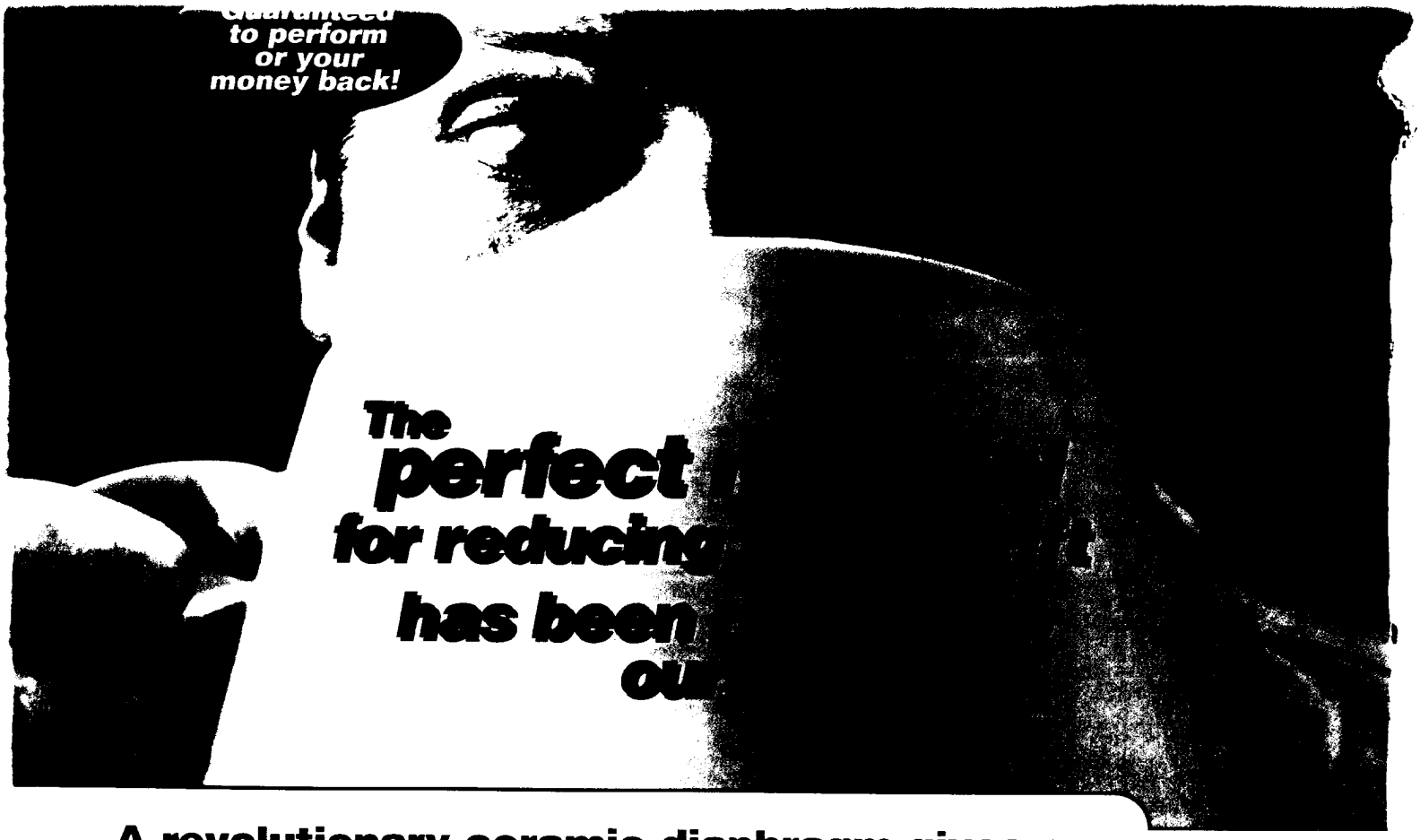
Papers from the 25th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Papers from the Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing

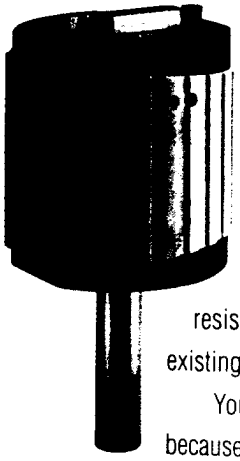
DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

An official journal of the American Vacuum Society
Published for the Society by the American Institute of Physics

Reproduced From
Best Available Copy



A revolutionary ceramic diaphragm gives our new CDG superior performance—at lower cost.



Old-fashioned metal-diaphragm CDGs are about to get a wake-up call. Our new CDG uses a remarkable aluminum oxide ceramic diaphragm to produce measurements with superior accuracy and repeatability.

Even small temperature changes affect metal diaphragms, creating zero drift that can dramatically impact a CDG's accuracy. Alumina ceramic cuts temperature-induced zero drift by almost 90%.

Ceramic is also more rigid than metal, so it's not deformed by repeated pressure cycling or overpressure. This further decreases zero drift and shortens the time needed for readings to stabilize during each process cycle, as well as eliminating the need for isolation valves. In addition, ceramic has nearly perfect corrosion resistance—so our CDG maintains its superior repeatability through every process. It is also fully compatible with existing installations.

You'd expect a CDG with such dramatically improved performance to cost more, but it actually costs *less*. And because it needs so little maintenance, it keeps on costing less.

If a CDG that delivers better repeatability, lower zero drift and lower cost of ownership sounds like your cup of tea, call us today.

LEYBOLD LEYBOLD INFICON

The Instrumental Difference™

United States
Phone: (315) 434-1100
Fax: (315) 437-3803

Germany
Phone: (+49) 221 347-0
Fax: (+49) 221 347-1250

France
Phone: (+33) 01 69 076400
Fax: (+33) 01 69 075738

United Kingdom
Phone: (+44) 0181 9717000
Fax: (+44) 0181 9717005

Netherlands
Phone: (+31) 03480 77411
Fax: (+31) 03480 20489

Japan
Phone: (+81) 045 471 3311
Fax: (+81) 045 471 3324

Korea
Phone: (+82) 02 270 133
Fax: (+82) 02 271 156

Taiwan
Phone: (+886) 03 542567
Fax: (+886) 03 542123

Visit our website for other sales offices worldwide: http://www.leyboldinficon.com

E-mail: reachus@inficon.com

REGISTERED ISO 9001 MAY 1993

OFFICE OF NAVAL RESEARCH

FINAL REPORT

for

Grant No. N00014-98-1-0137

PR Number 98Pr02114-00

Twenty-fifth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Under the direction of
C. R. Schulte
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, Arizona 85252-0036

Reproduction in whole or in part is permitted for any purpose of the United States Government.

*This document has been approved for public release and sale; its distribution is unlimited.

Twenty-fifth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Under the direction of
C. R. Schulte
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, Arizona 85252-0036

Conference

The Twenty-fifth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in Salt Lake City, Utah, in January of 1998, and was attended by over 100 workers in the field of semiconductor interfaces, including researchers from all over the world.

The conference featured extensive discussion sessions and a workshop format, and was judged a success by its attendees — largely because it facilitated informal discussion among the attendees, and especially involving young attendees (*e.g.*, graduate students) interacting with senior scientists.

Publications

The Proceedings of the conference are enclosed, and feature the papers presented which have been published in the Journal of Vacuum Science and Technology.

Expenditure of funds

The following represents the expenditure of Department of Defense funds for this conference:

EXPENDITURES OF DEPARTMENT OF DEFENSE FUNDS: ONR

A. Travel
Domestic travel by Committee: $4,967.18
Domestic travel and conference fees for speakers/discussion leaders: 800.00

TOTAL TRAVEL	3,049.23
B. Secretarial Services	4,997.00
C. Publications	0.00
D. Graphics, Printing, Postage, Communications	203.77
E. Audio-visual equipment and related facilities	0.00
F. All other conference facilities and services	0.00
G. Accounting, overhead, and legal services (a flat fee)	750.00
H. **GRAND TOTAL**	$9,000.00

FINANCIAL STATUS REPORT

(Long Form)

(Follow instructions on the back)

1. Federal Agency and Organizational Element to Which Report is Submitted	2. Federal Grant or Other Identifying Number Assigned By Federal Agency	OMB Approval No.	Page	of
Office of Naval Research Regional Office San Diego	N00014-98-1-0137	0348-0039	1	1 pages

3. Recipient Organization (Name and complete address, including ZIP code)

Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, AZ 85252-0036

4. Employer Identification Number	5. Recipient Account Number or Identifying Number	6. Final Report	7. Basis
86-0760869	PCSI25	☐ Yes ☐ No X	☐ Cash ☐ Accrual X

8. Funding/Grant Period (See Instructions) From: (Month, Day, Year)	To: (Month, Day, Year)	9. Period Covered by this Report From: (Month, Day, Year)	To: (Month, Day, Year)
Nov 1 1997-	Oct. 31, 1998	Nov. 1, 1997	Oct. 31, 1998

10. Transactions:	I Previously Reported	II This Period	III Cumulative
a. Total outlays	$8100.00	$900.00	$9000.00
b. Refunds, rebates, etc.	0	0	0
c. Program income used in accordance with the deduction alternative			
d. Net outlays *(Line a, less the sum of lines b and c)*	$8100.00	$900.00	$9000.00
Recipient's share of net outlays, consisting of:			
e. Third party (in-kind) contributions	0	0	0
f. Other Federal awards authorized to be used to match this award	0	0	0
g. Program income used in accordance with the matching or cost sharing alternative	0	0	0
h. All other recipient outlays not shown on lines e, f or g	0	0	0
i. Total recipient share of net outlays *(Sum of lines e, f, g and h)*	0	0	0
j. Federal share of net outlays (line d less line i)			$9000.00
k. Total unliquidated obligations			0
l. Recipient's share of unliquidated obligations			0
m. Federal share of unliquidated obligations			0
n. Total federal share (sum of lines j and m)			$9000.00
o. Total federal funds authorized for this funding period			$9000.00
p. Unobligated balance of federal funds *(Line o minus line n)*			0
Program income, consisting of:			
q. Disbursed program income shown on lines c and/or g above			0
r. Disbursed program income using the addition alternative			0
s. Undisbursed program income			0
t. Total program income realized *(Sum of lines q, r and s)*			0

11. Indirect Expense	a. Type of Rate *(Place "X" in appropriate box)* ☐ Provisional ☒ Predetermined ☐ Final ☐ Fixed	b. Rate	c. Base	d. Total Amount	e. Federal Share
		$750.00 flat	$9000.00	$750.00	$750.00

12. *Remarks: Attach any explanations deemed necessary or information required by Federal sponsoring agency in compliance with governing legislation.*

13. Certification: **I certify to the best of my knowledge and belief that this report is correct and complete and that all outlays and unliquidated obligations are for the purposes set forth in the award documents.**

Typed or Printed Name and Title	Telephone (Area code, number and extension)
C. R. Schulte, Principal Investigator	602-423-8540
Signature of Authorized Certifying Official	**Date Report Submitted**
C. R. Schulte	December 19, 1998

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	19 December 1998	Final 01 Nov 97 through 31 Oct 98

4. TITLE AND SUBTITLE
Final Report:
Twenty-Fifth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

5. FUNDING NUMBERS
Grant No: N00014-98-1-0137

6. AUTHOR(S)
C. R. Schulte

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, AZ 85252-0036

8. PERFORMING ORGANIZATION REPORT NUMBER
F1-1998

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
ONR 251: Elizabeth L. Ford
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660 CFDA No. 12.300

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Office of Naval Research position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Twenty-Fifth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in Salt Lake City, Utah, and the enclosed report contains abstracts of papers presented.

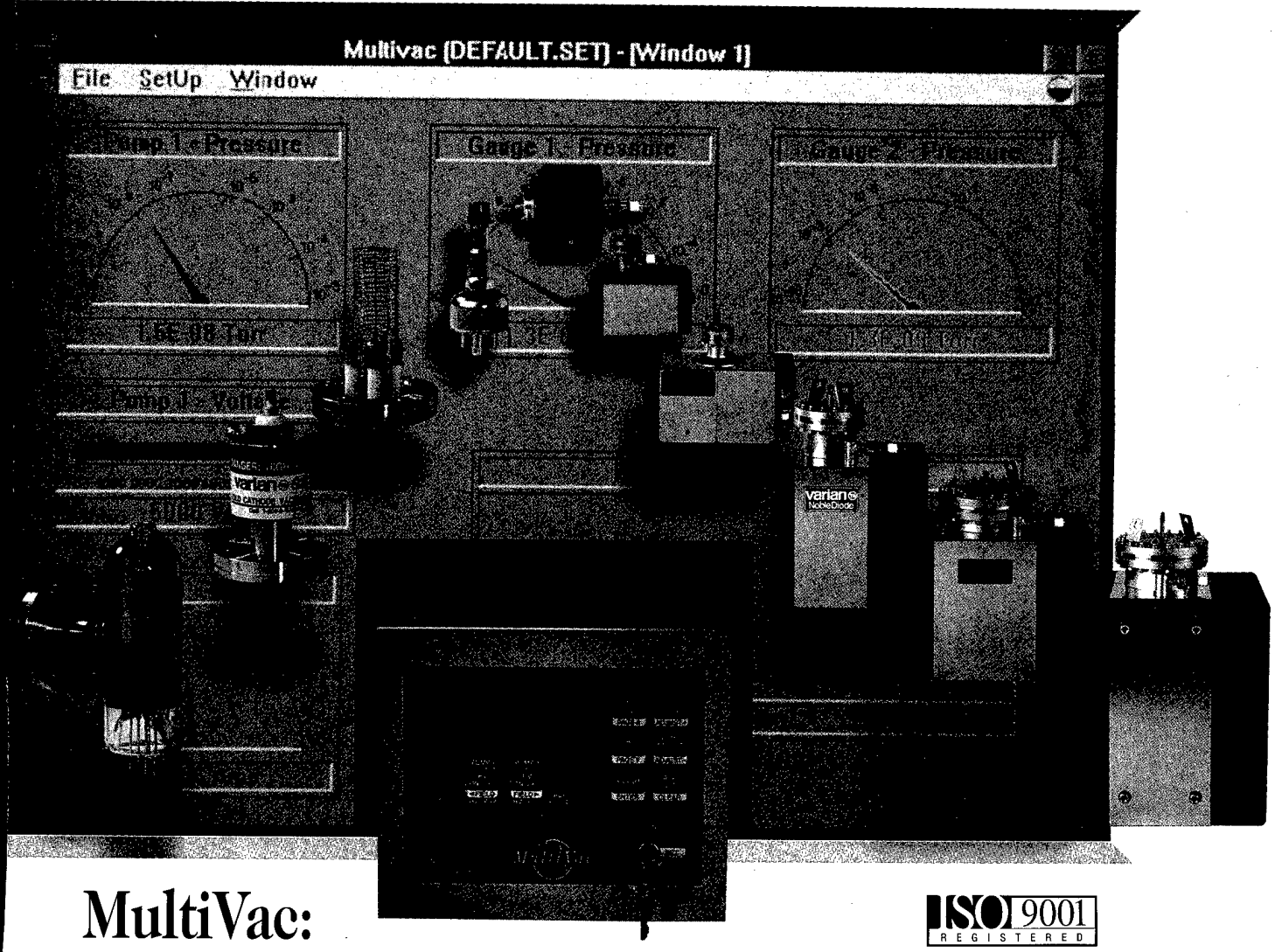
14. SUBJECT TERMS
Semiconductor Interfaces

15. NUMBER OF PAGES
301

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Take control over your ion pumping needs



MultiVac:
a new window to the high-vacuum world

ISO 9001 REGISTERED

Varian's MultiVac Controller operates up to two different ion pumps and two gauges, simultaneously and independently. So whether you need to operate a single pump, or have an application that demands multiple pumps, gauges, and computer control, MultiVac is the only controller you'll need.

Unlike conventional controllers that operate at fixed voltages, MultiVac automatically optimizes voltage according to operating pressure, giving you the best possible pumping performance at any pressure. This also minimizes leakage current, delivering more reliable pressure readings down to the UHV region (10^{-10} mbar range).

MultiVac is small, lightweight, and cost-effective. And with optional Windows® software, it's very easy to use.

To find out more about the *only* controller you need, call Varian today.

In the U.S.	1.800.8VARIAN
In Canada	1.800.663.2727
In the U.K.	1932.898.000
Rest of Europe	39.011.997.9111
In Japan	81.3.5232.1253

http://www.varian.com/vacuum

varian 50 YEARS

VAC-8430/468C

Journal of Vacuum Science & Technology B

JVST B

ISSN: 0734-211X
CODEN: JVTBD9

Microelectronics and Nanometer Structures
Processing, Measurement, and Phenomena

Editor: G. E. McGuire,
MCNC, Research Triangle Park, NC 27709. Telephone: (919) 248-1910

Supervisor Editorial Office: Rebecca York,
Editorial Assistant: Mary P. Hipsher,
Journal of Vacuum Science and Technology, Caller Box 13994, 10 Park Plaza, Ste. 4A,
Research Triangle Park, NC 27709, Telephone: (919) 361-2787 and (919) 361-2342;
FAX: (919) 361-1378; E-mail: jvst@jvst.org

Associate Editors:

Eric Kay, Review Articles
C. R. Abernathy (98) University of Florida
S. Felch (2000) Varian Associates
J. Givens (99) Micron Technology, Inc.
J. P. McVittie (2000) Stanford University
G. Owen (2000) Hewlett Packard Labs.
J. L. Shaw (98) Naval Research Lab.

JVST Publication Committee:
Dorota Temple, Chair, MCNC
John E. Crowell, UC San Diego
Joseph D. Geller, Geller Microanalytical Lab.
J. William Rogers, Jr., Univ. of Washington
Pete Sheldon, NREL

JVST Editorial Board:
S. A. Chambers (99) Pacific NW National Lab.
R. J. Colton (98) Naval Research Lab.
A. Czanderna (2000) NREL
M. Kushner (2000) Univ. of Illinois
T. Mayer Sandia Natl. Labs
B. S. Meyerson (98) IBM Watson Research Ctr.
D. Monroe (2000) Lucent Bell Labs
D. Ruzic (99) Univ. of Illinois

JVST Editorial Staff at AIP: Editorial Supervisor: Deborah McHone; Journal Coordinator: Margaret Reilly; Chief Production Editor: Mary Ellen Mormile

The Journal of Vacuum Science & Technology B is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) by the American Vacuum Society (AVS) through the American Institute of Physics (AIP). It is an official publication of the AVS and is received by all members of the Society. It is devoted to reports of original research and Review articles. The *JVST B* has been established to provide a vehicle for the publication of research dealing with microelectronics and nanometer structures. The emphasis will be on processing, measurement, and phenomena, and will include vacuum processing, plasma processing, materials and structural characterization, microlithography, and the physics and chemistry of submicron and nanometer structures and devices. This journal will publish the proceedings of conferences and symposia that are sponsored by the AVS and its divisions.

Submit Manuscripts to the Editorial Office of the *Journal of Vacuum Science & Technology*, 10 Park Plaza, Caller Box 13994, Research Triangle Park, North Carolina 27709-3994; e-mail jvst@jvst.org. Manuscripts of papers presented at AVS-sponsored conferences and symposia and being submitted to *JVST B* should be sent to the Guest Editor appointed for that particular conference. Before preparing a manuscript, authors should read "Information for Contributors," printed in the first issue of each volume of the journal. Submission of a manuscript is a representation that the manuscript has not been published previously nor currently submitted for publication elsewhere. Upon receipt of a manuscript the Editor will send the author a Transfer of Copyright Agreement form. This must be completed by the author and returned *only* to the Editorial Office prior to publication of an accepted paper in the *Journal of Vacuum Science & Technology B*. This written transfer of copyright, which previously was assumed to be implicit in the act of submitting a manuscript, is necessary under the 1978 copyright law in order for the AVS and AIP to continue disseminating research results as widely as possible. Further information may be obtained from AIP.

Publication Charge: To support the cost of wide dissemination of research results through the publication of journal pages and production of a database of articles, the author's institution is requested to pay a *page charge* of $95 per page (with a one-page minimum). The charge (if honored) entitles the author to 100 free reprints. For Errata the minimum page charge is $10, with no free reprints.

Physics Auxiliary Publication Service (PAPS) and/or Electronic PAPS (E-PAPS): For a nominal fee, authors may submit material that is part of and supplemental to a paper, but is too long to be included in the journal. PAPS deposits may be on paper or in electronic media, and can include text, data in graphic or numeric form, computer programs, etc. Retrieval instructions are footnoted in the related published paper. Direct requests to the Editor.

Advertising Rates will be supplied on request from AIP's Advertising Division, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Telephone: (516) 576-2440. Fax: (516) 576-2481. E-mail: advtsg@aip.org. All insertion orders and advertising material should be sent to that division.

Copying: Single copies of individual articles may be made for private use or research. Authorization is given (as indicated by the Item Fee Code for this publication) to copy articles beyond the use permitted by Sections 107 and 108 of the U.S. Copyright Law, provided the copying fee of $15 per copy per article is paid to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, USA. Persons desiring to photocopy materials for classroom use should contact the CCC Academic Permissions Service. The Item Fee Code for this publication is 0734-211X/98 $15.00.

Authorization does not extend to systematic or multiple reproduction, to copying for promotional purposes, to electronic storage or distribution, or to republication in any form. In all such cases, specific written permission from AIP must be obtained.

Permission for Other Use: Permission is granted to quote from the journal with the customary acknowledgment of the source. To reprint a figure, table, or other excerpt requires the consent of one of the authors and notification to AIP.

Requests for Permission: Address requests to AIP Office of Rights and Permissions, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Telephone: (516) 576-2268; Fax: (516) 576-2327; Internet: rights@aip.org.

Copyright © 1998 American Vacuum Society. All rights reserved.

American Vacuum Society

Officers

Jerry M. Woodall, *President*
Purdue University

Stephen Rossnagel, *President-Elect*
IBM T. J. Watson Research Center

Gary E. McGuire, *Immediate Past-President*
Microelectronics Center of North Carolina

Joseph E. Greene, *Secretary*
University of Illinois

N. Rey Whetten, *Treasurer*
American Vacuum Society

Directors

Dawn A. Bonnell
University of Pennsylvania

Yip-Wah Chung
Northwestern University

Elizabeth Dobisz
Naval Research Laboratory

Richard A. Hoffman
Northwestern University

Rudolf Ludeke
IBM T. J. Watson Research Center

Peter Sheldon
National Renewable Energy Laboratory

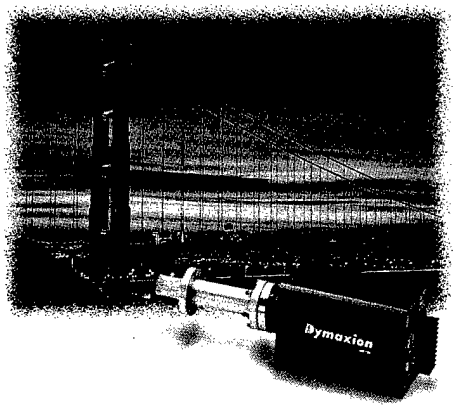
JVST

G. Lucovsky, *Editor-in-Chief*
North Carolina State University

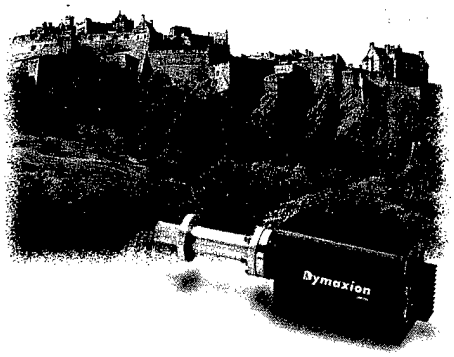
AVS Membership Information may be obtained from

Angela Mulligan
AVS Membership and Scholarship Coordinator
120 Wall Street
32nd Floor
New York, NY 10005
(212) 248-0200

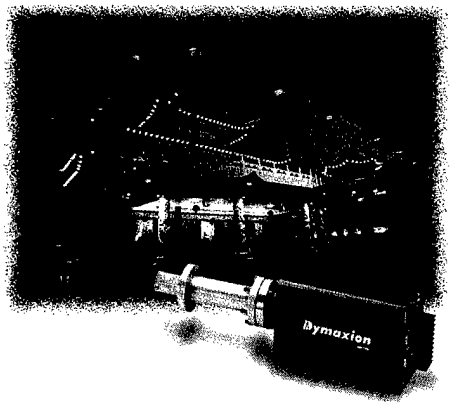
NOW THERE'S AN RGA THAT CAN PROVIDE REPEATABLE DATA FROM INSTRUMENT TO INSTRUMENT, PLANT TO PLANT,



SAN FRANCISCO, CALIFORNIA



EDINBURGH, SCOTLAND

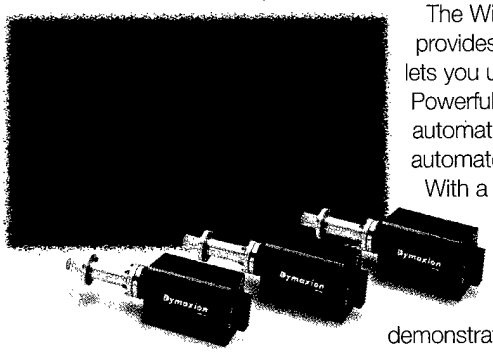


KUALA LUMPUR, MALAYSIA

ANYWHERE AROUND THE WORLD.

The brand new Dymaxion sensor and powerful System 2000 software makes the DYCOR® RGA an on-line process tool you can count on.

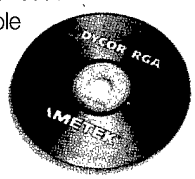
The Dymaxion smart sensor features a zero-drift amplifier and on-board calibration routines that give you data integrity from sensor to sensor. And, with its zero-footprint, it mounts directly on the tool to conserve clean room space.



The Windows® NT-based System 2000 software provides a pre-emptive, multi-tasking environment that lets you use that data for real time process control. Powerful scripting capabilities, and DDE and OLE automation mean you can, for the first time, truly automate equipment and process monitoring.

With a DYCOR RGA, you'll get dependable data from every sensor you install. With the resources of AMETEK, Inc., you'll get dependable support anywhere you install it.

Call or fax for a CD-ROM demonstration of the Dymaxion RGA and System 2000 software capabilities.



AMETEK®
PROCESS INSTRUMENTS
www.ametek.com

AMETEK Process Instruments
150 Freeport Road
Pittsburgh, PA 15238. U.S.A.
Tel: 412-828-9040
Fax: 412-826-0399

AMETEK Singapore Pvt. Ltd.
10 Ang Mo Kio Street 65, #05-12 TECHPOINT
Singapore 569059
Tel: (65) 484-2388
Fax: (65) 481-6588

AMETEK Precision Instruments Europe GmbH
Rudolf-Diesel-Strasse 16, D-40670
Meerbusch, Germany
Tel: (49) 21 59 9136 0
Fax: (49) 21 59 9136-39

CHA Industries

HIGH VACUUM DEPOSITION SYSTEMS

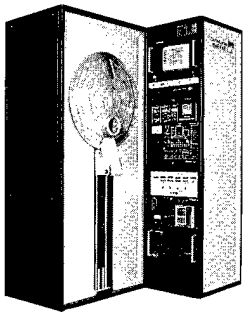
Evaporation and sputtering systems for optical and semiconductor manufacturing and precision industrial coating applications...

CHA field-proven deposition systems are the ultimate in precision and quality. They feature multi-substrate processing, fast loading and unloading, high throughput, and accommodate a variety of substrate sizes, shapes, and materials.

CHA systems utilize moving substrates for consistently high film uniformity. They incorporate multiple sources, vertical sealing planes on all valves, simplified high vacuum construction, and fast pump-down. Options include substrate heat, substrate preclean, ion beam etch, RF etch, and DC or RF substrate bias.

Serving worldwide optical and semiconductor manufacturing and precision coating industries for over 40 years.

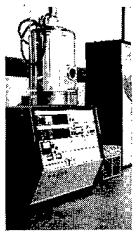
COMBINATION EVAPORATION AND SPUTTERING SYSTEMS



MARK 40/50

Dual process — sputtering and/or evaporation. Features include rotating fixture, 360-degree substrate heating, and unique shielding. Source Isolation option. Through-the-wall clean-room installation.

EVAPORATION SYSTEMS

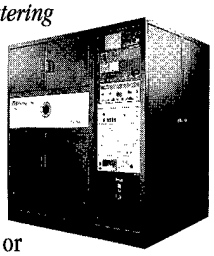


SE-SERIES

Bell-jar type system accepts all commercially-available sources and can be equipped with a variety of rotating, single or multi-planet fixturing.

SPUTTERING SYSTEMS

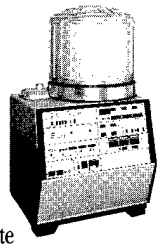
Superior sputtering since 1968



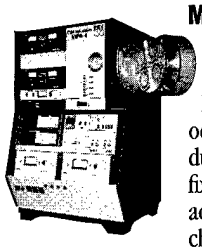
SSC-600/1000

Sputter up or sputter down, up to four DC or RF cathodes. Substrate and source access outside of chamber. Through-the-wall clean-room installation.

MAG-DRUM 1000™



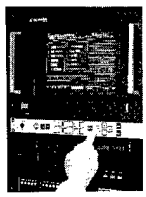
Side sputtering, up to three DC or RF cathodes. Featuring rotating, multi-planet fixturing, 360-degree substrate heating, and unique shielding.



MPS-4 FAMILY

Sputter sideways, up to four DC or RF cathodes. Features dual-axis planetary fixturing, easy access outside of chamber, and uncommon flexibility.

CONTROL SYSTEMS



Optional Windows™-based, computerized control system for fully automatic cycling of system and process. CHA Autotech control or PLC also available for one-button automatic operation.

POWER SUPPLIES

SR-Series E-Beam Power Supplies feature superior emission current and unique gun-programming capabilities. 3KW through 15KW models.

ACCESSORIES/HARDWARE

CHA offers a full range of sources, fixturing, pumping systems, cryo-coil assemblies, bell jars, feed-throughs, couplings, and controls.

FOR DETAILS, CALL US AT...

510-683-8554

CHA Industries

4201 Business Center Drive
Fremont, CA 94538
Fax: 510-683-3848
www.chaindustries.com

Windows is a trademark of Microsoft Corp.

Journal of Vacuum Science & Technology B

JVST B

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Second Series
Volume 16, Number 4
Jul/Aug 1998

Regular Articles

Strain-induced birefringence in $Si_{1-x}Ge_x$ optical waveguides
M. Robillard, P. E. Jessop, D. M. Bruce, S. Janz, R. L. Williams, S. Mailhot, H. Lafontaine, S. J. Kovacic, and J. J. Ojha **1773**

Relaxation of strained, epitaxial $Si_{1-x}Sn_x$
M. F. Fyhn, J. Chevallier, and A. Nylandsted Larsen **1777**

Layer by layer growth mode stabilization and step-edge smoothing of cation-stabilized $In_{1-x}Ga_xAs$ strained layers grown on InP
Y. Robach, A. Soléré, M. Gendry, and L. Porte **1786**

(Continued)

Journal of Vacuum Science & Technology B (ISSN: 0734-211X) is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) by the American Vacuum Society through the American Institute of Physics, 500 Sunnyside Blvd., Woodbury, NY 11797-2999. Membership in the American Vacuum Society includes $17.50 from membership dues to be applied towards a subscription to *Journal of Vacuum Science & Technology B.*

Known office of publication is Thorofare, NJ 08086. Periodicals postage paid at Thorofare, NJ 08086, and at additional mailing offices. POSTMASTER: Send address changes to *Journal of Vacuum Science & Technology B,* SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.

Subscription Prices (1998)

	U.S.A. and Poss.	Can., Mex., Central & S. America & Caribbean	Europe, Asia, Africa & Oceania*
JVST B†	$727	$755	$775
JVST B‡	$835	$862	$885
JVST A and B†	$835	$890	$930
JVST A and B#	$835	$862	$862
JVST A and B‡	$945	$1000	$1040

†Paper only.
‡Paper and CD-ROM.
#CD-ROM only.
*Nonmember subscriptions include air-freight service.

Back-number Prices: 1998 single copy: $135. Prior to 1998 single copies: $125.

Subscriptions, renewals, and address changes should be addressed to *Subscription Fulfillment Division, SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.* Allow at least six weeks advance notice. For address changes please send both old and new addresses, and, if possible, include a label from the plastic mailing wrapper of a recent issue. For your convenience a **change-of-address form is included in every issue of *Physics Today;* please use it.** AIP will honor a request for a missing journal issue only within six months of that issue's actual date of publication. Requests received for missing issues beyond six months of the actual publication date will not be honored. The issue may, however, be purchased at the single-copy price. (Subscription Fulfillment offices are located at SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.)

Single-copy orders (current and back issues) should be addressed to American Institute of Physics, Circulation and Fulfillment Division, 500 Sunnyside Blvd., Woodbury, NY 11797-2999; Telephone 800-344-6902 (or 516-576-2270 outside the U.S.A.), Fax at 516-349-9704, or E-mail at subs@aip.org.

Reprints: Reprints can be ordered with or without covers only in multiples of 50 (with a minimum of 100 in each category) from AIP, Circulation and Fulfillment/ Reprints, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6909 (U.S. and Canada) or 516-576-2234.

Document Delivery: Copies of journal articles can be ordered for online delivery from the new Articles in Physics online document delivery service (URL: http://www.aip.org/articles.html).

Copying fees: The code that appears on the first page of articles in this journal gives the fee for each copy of the article made beyond the free copying permitted by AIP. (See statement under "Copyright" elsewhere in this journal.) If no code appears, no fee applies. The fee for pre-1978 articles is $0.25 per copy. With the exception of copying for advertising and promotional purposes, the express permission of AIP is not required provided the fee is paid through the *Copyright Clearance Center, Inc. (CCC), 222 Rosewood Dr., Danvers, MA 01923.* Contact the CCC for information on how to report copying and remit payment.

Microform: *Journal of Vacuum Society & Technology B* is available on microfiche issued at the same frequency as the printed journal and annually on microfilm. A microform catalog is available from AIP, Circulation and Fulfillment/ Single, Copy Sales, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6908 (U.S. and Canada) or 516-576-2277.

Online Access: The *Journal of Vacuum Science and Technology B* is available online to AVS members at no additional charge; for details, please see http://www.aip.org/ojs/entry.html. Abstracts of journal articles are available from AIP's SPIN Database, via AIP's PINET Plus Service (http://www.aip.org/pinet).

Reduced 980 nm laser facet absorption by band gap shifted extended cavities
P. G. Piva, R. D. Goldberg, I. V. Mitchell, S. Fafard, M. Dion, M. Buchanan, S. Charbonneau, G. Hillier, and C. Miner 1790

Localized interface states and the optical spectra of AlSb/InAs heterostructures
M. J. Shaw, G. Gopir, P. R. Briddon, and M. Jaros 1794

Chemical beam epitaxial growth of $GaAs_{1-x}P_x$ on GaAs (100) substrates
D. Wildt, B. J. García, J. L. Castaño, J. Piqueras, and C. J. Pastor 1804

Effect of a deep-level trap on hole transport in $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ metal–semiconductor–metal photodetectors
Kun-Jing Lee, F. G. Johnson, W. B. Johnson, Junghwan Kim, and Chi H. Lee 1808

Effect of the forward biasing the source-substrate junction in *n*-metal–oxide–semiconductor transistors for possible low power complementary metal–oxide–semiconductor integrated circuits' applications
F. J. De la Hidalga-W., M. J. Deen, E. A. Gutierrez-D., and F. Balestra 1812

Reduced damage reactive ion etching process for fabrication of InGaAsP/InGaAs multiple quantum well ridge waveguide lasers
B. C. Qiu, B. S. Ooi, A. C. Bryce, S. E. Hicks, C. D. W. Wilkinson, R. M. De La Rue, and J. H. Marsh 1818

X-ray photoelectron spectroscopy damage characterization of reactively ion etched InP in CH_4–H_2 plasmas
Y. Feurprier, Ch. Cardinaud, and G. Turban 1823

Germanium etching in high density plasmas for 0.18 μm complementary metal–oxide–semiconductor gate patterning applications
C. Monget, A. Schiltz, O. Joubert, L. Vallier, M. Guillermet, and B. Tormen 1833

Estimation of the activation energy for Ar/Cl_2 plasma etching of InP via holes using electron cyclotron resonance
E. W. Sabin 1841

Abrupt reduction in poly-Si etch rate in HBr/O_2 plasma
Shinji Kuroda and Hiroaki Iwakuro 1846

Structural, optical, and electrical properties of nanocrystalline silicon films deposited by hydrogen plasma sputtering
B. Garrido, A. Pérez-Rodríguez, J. R. Morante, A. Achiq, F. Gourbilleau, R. Madelon, and R. Rizk 1851

Deep, vertical etching of flame hydrolysis deposited hi-silica glass films for optoelectronic and bioelectronic applications
A. J. McLaughlin, J. R. Bonar, M. G. Jubber, P. V. S. Marques, S. E. Hicks, C. D. W. Wilkinson, and J. S. Aitchison 1860

Sidewall slope control of chemically assisted ion-beam etched structures in InP-based materials
J. Daleiden, K. Czotscher, C. Hoffmann, R. Kiefer, S. Klussmann, S. Müller, A. Nutsch, W. Pletschen, S. Weisser, G. Tränkle, J. Braunstein, and G. Weimann 1864

Evaluation of trifluoroiodomethane as SiO_2 etchant for global warming reduction
F. Fracassi and R. d'Agostino 1867

Electrical characterization and annealing behavior of defect introduced in Si during sputter etching in an Ar plasma
P. N. K. Deenapanray, F. D. Auret, and G. Myburg 1873

Low frequency noise in heavily doped polysilicon thin film resistors
M. J. Deen, S. Rumyantsev, and J. Orchard-Webb 1881

Damage-free cleaning of Si(001) using glancing-angle ion bombardment
Jose Gregorio C. Labanda, Scott A. Barnett, and L. Hultman 1885

Inductively coupled plasma etching of $(Ba,Sr)TiO_3$ thin films
H. M. Lee, D. C. Kim, W. Jo, and K. Y. Kim 1891

Reactive ion etching of $Pb(Zr_xTi_{1-x})O_3$ thin films in an inductively coupled plasma
Chee Won Chung 1894

Structural study of $YSi_{1.7}$ layers formed by channeled ion beam synthesis
M. F. Wu, Shude Yao, A. Vantomme, S. Hogg, H. Pattyn, G. Langouche, Qinqing Yang, and Qiming Wang 1901

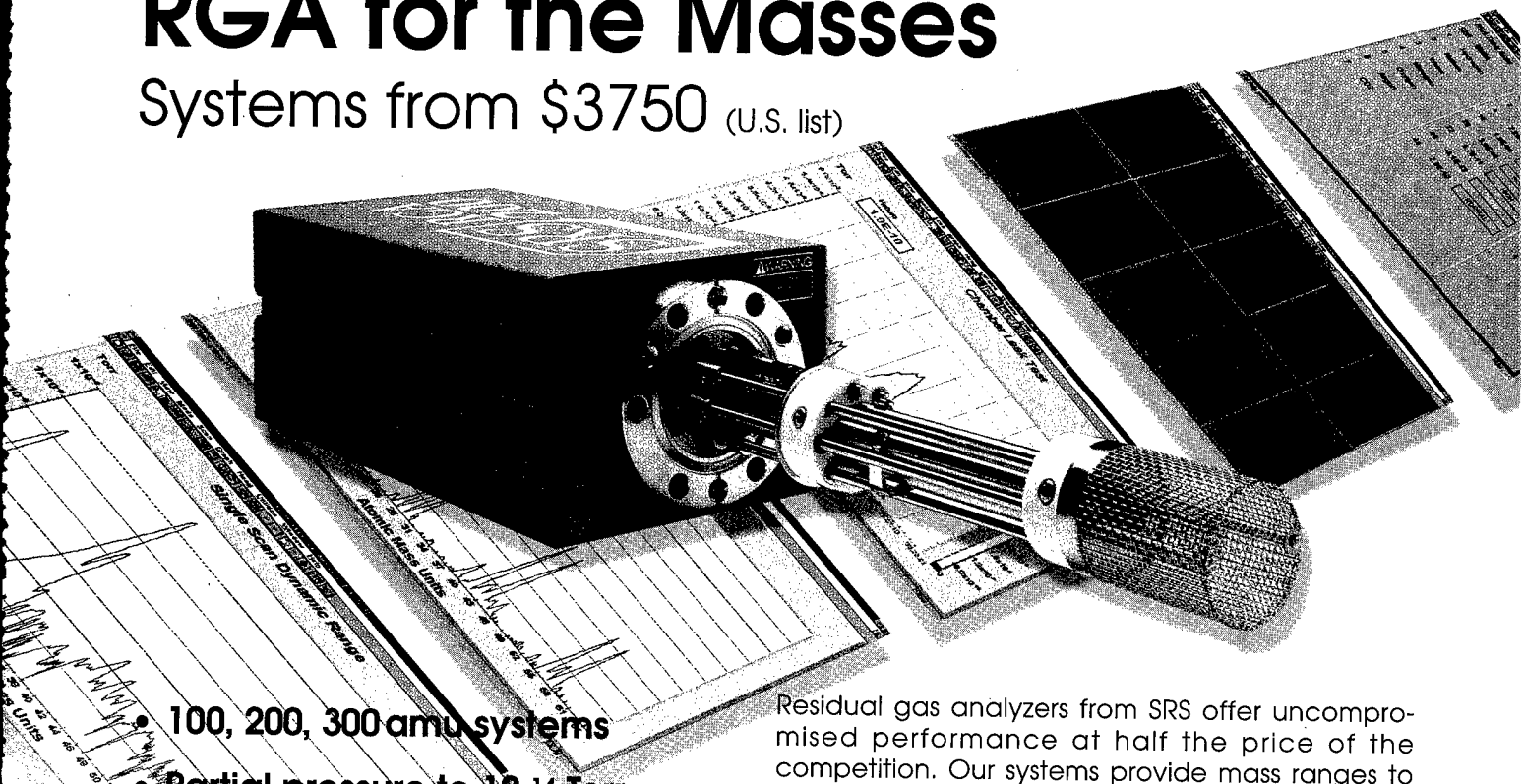
Implanted gallium-ion concentrations of focused-ion-beam prepared cross sections
T. Ishitani, H. Koike, T. Yaguchi, and T. Kamino 1907

Widely changing probability of surface damage creation induced by a single ion in the MeV ion energy range
Hisato Ogiso, Hiroshi Tokumoto, Shizuka Nakano, and Kazushi Yamanaka 1914

(Continued)

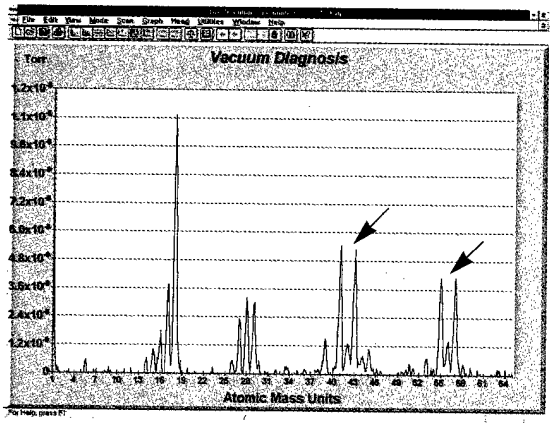
RGA for the Masses

Systems from $3750 (U.S. list)

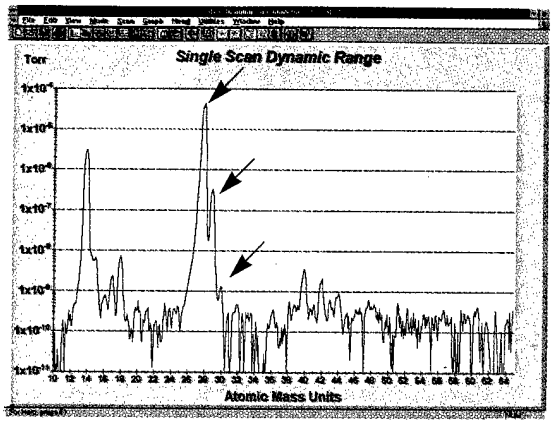


- **100, 200, 300 amu systems**
- **Partial pressure to 10^{-14} Torr**
- **6 orders of magnitude dynamic range in a single scan**
- **Real-time Windows® software**
- **Mass spectra, leak detection and pressure vs. time modes**
- **Multi-head operation**
- **Optional Electron Multiplier**

Residual gas analyzers from SRS offer uncompromised performance at half the price of the competition. Our systems provide mass ranges to 300 amu, detectable partial pressures to 10^{-14} Torr and better than 1 amu resolution. With our interactive Windows® software package you can fully control the RGA sensor, acquire data and analyze your vacuum system in real-time. The simple, rugged quadrupole mass spectrometer attaches directly to a 2 3/4" CF flange. A dual ThO_2Ir filament and a unique continuous dynode electron multiplier provide increased stability and longer life than other designs. Both are field replaceable in a matter of minutes – a first for RGAs! If your application involves residual gas analysis, leak detection or vacuum processing, contact us for a copy of our RGA systems brochure and demo software package.



The RGA is an invaluable vacuum diagnostic tool. Shown above is the mass spectrum of a vacuum chamber contaminated with oil.



With dynamic range that spans 6 orders of magnitude, three isotopes of Nitrogen ($^{14}N_2$, $^{14}N\,^{15}N$, $^{15}N_2$) are clearly detected in a single scan.

Stanford Research Systems
1290-D Reamwood Avenue, Sunnyvale, CA 94089
Telephone (408)744-9040 • FAX (408)744-9049
Email: info@srsys.com • WWW: http://www.srsys.com/srsys

Windows is a registered trademark of Microsoft Corporation

Evidence of depth and lateral diffusion of defects during focused ion beam implantation 1919
C. Vieu, J. Gierak, M. Schneider, G. Ben Assayag, and J. Y. Marzin

On the improvement in thermal quenching of luminescence in SiGe/Si structures grown by molecular beam epitaxy 1928
W. M. Chen, I. A. Buyanova, G. Pozina, B. Monemar, W. X. Ni, and G. V. Hansson

Changes in morphology using atomic hydrogen during $Si/Si_{1-x}Ge_x$ molecular beam epitaxy growth on Si (100) 1933
C. Silvestre, G. G. Jernigan, M. E. Twigg, and P. E. Thompson

Stabilizing the surface morphology of $Si_{1-x-y}Ge_xC_y/Si$ heterostructures grown by molecular beam epitaxy through the use of a silicon-carbide source 1937
E. T. Croke, J. J. Vajo, A. T. Hunter, C. C. Ahn, D. Chandrasekhar, T. Laursen, David J. Smith, and J. W. Mayer

Low temperature scanning tunneling microscope-induced luminescence of GaN 1943
S. Evoy, C. K. Harnett, H. G. Craighead, T. J. Eustis, W. A. Davis, M. J. Murphy, W. J. Schaff, and L. F. Eastman

High efficiency, dual collection mode near-field scanning optical microscope 1948
Stephan J. Stranick, Lee J. Richter, and Richard R. Cavanagh

Probing conducting particles buried in a $Ni_x(SiO_2)_{1-x}$ composite by conducting atomic force microscopy 1953
E. Z. Luo, I. H. Wilson, J. B. Xu, J. X. Ma, and X. Yan

Microroughness of polymer thin films studied by total-reflection x-ray fluorescence and atomic force microscopy 1958
W. L. Wu and W. E. Wallace

Plasticity study of deformed materials by *in situ* atomic force microscopy 1964
C. Coupeau, J. C. Girard, and J. Grilhé

Secondary ion mass spectrometry and atomic force spectroscopy studies of surface roughening, erosion rate change and depth resolution in Si during 1 keV 60° O_2^+ bombardment with oxygen flooding 1971
Z. X. Jiang and P. F. A. Alkemade

Novel approach to atomic force lithography 1983
S. Hu, S. Altmeyer, A. Hamidi, B. Spangenberg, and H. Kurz

Dry-etching development characteristics of $Se_{75}Ge_{25}$ resist for focused-ion-beam lithography 1987
Hyun-Young Lee and Hong-Bay Chung

Blurring effect analysis of an x-ray mask for synchrotron radiation lithography 1992
Il Yong Kim, Byung Man Kwak, Young Jin Jeon, and Sang Soo Choi

Photoresist erosion studied in an inductively coupled plasma reactor employing CHF_3 1998
M. F. Doemling, N. R. Rueger, G. S. Oehrlein, and J. M. Cook

Apparent tunnel barrier heights of PtIr–Au interfaces in relation to the Au surface composition 2006
Lionel Boyer, Sophie Noël, and Frédéric Houzé

Structural and electrical properties of chemical vapor deposition tungsten overgrowth on physical vapor deposited and metalorganic chemical vapor deposited TiN adhesion layers 2013
Y. C. Peng, L. J. Chen, W. Y. Hsieh, Y. R. Yang, and Y. F. Hsieh

Characterization of TiN barriers against Cu diffusion by capacitance–voltage measurement 2019
Sa-Kyun Rha, Seung-Yun Lee, Won-Jun Lee, Yong-Sup Hwang, Chong-Ook Park, Dong-Won Kim, Youn-Seoung Lee, and Chung-Nam Whang

Thermal stability of selective chemical vapor deposited tungsten contact and effects of *in situ* N_2 plasma treatment 2026
M. T. Wang, P. C. Wang, M. C. Chuang, L. J. Chen, and M. C. Chen

Effects of sintering on Au/Ti/GaAs Schottky barrier submicron metal-semiconductor field-effect transistors characteristics 2034
M. M. Ahmed

Advanced emitters for next generation rf amplifiers 2038
K. L. Jensen, J. E. Yater, E. G. Zaidman, M. A. Kodis, and A. Shih

Field emission from nitrogen doped tetrahedral amorphous carbon prepared by filtered cathodic vacuum arc technique 2049
L. K. Cheah, X. Shi, B. K. Tay, and Z. Sun

Field emission of nitrogen-doped diamond films 2052
I. T. Han, N. Lee, S. W. Lee, S. H. Kim, and D. Jeon

(Continued)

We're really big on Inorganics.

"We're glad that new composition worked the way you thought it would. Look for your first 120 kg right on schedule."

"Tom tried out the evaporation material we sent him last week. With a little tweaking it'll be just what they need. Here's what we've got to do..."

"Sectioned targets? Sure, we can do that. Just send us your specs..."

"Tell your research guys not to worry, Nancy. We keep that powder in stock all the time. It'll ship out today."

And bigger still on custom services.

Whether you need that stock item your boss wanted yesterday, a custom size that's got to be just right or bulk quantities for your next production run, just call CERAC. We'll work with you to provide what you need:

- **Virtually Any Size or Shape: Powders, Evaporation Materials, Sputtering Targets**
- **Quantities from Grams to Thousands of Kilograms**
- **The World's Widest Selection of Inorganics**
- **Nearly Any Composition**
- **Custom Manufacturing**

Call Today For Our Free Catalog

414-289-9800

Or visit our website at: **www.cerac.com**

CERAC™ incorporated

407 N. 13th St. • P.O. Box 1178 • Milwaukee, WI 53201 USA

Tel: 414/289-9800 • Fax: 414/289-9805 • E-mail: marketing@cerac.com • Home Page: www.cerac.com

YOUR SINGLE SOURCE FOR INORGANICS FROM R&D TO FULL SCALE UP.

Field emission from ZrC and ZrC films on Mo field emitters
W. A. Mackie, Tianbao Xie, M. R. Matthews, B. P. Routh, Jr., and P. R. Davis 2057

Experimental evaluation of the extended Schottky model for ZrO/W electron emission
M. J. Fransen, J. S. Faber, Th. L. van Rooy, P. C. Tiemeijer, and P. Kruit 2063

Surface work function studies in porous silicon
Tejashree M. Bhave and S. V. Bhoraskar 2073

Computer-controlled fabrication of ultra-sharp tungsten tips
Young-Guon Kim, Eun-Ha Choi, Seung-Oun Kang, and Guangsup Cho 2079

Electrophysical characteristics and low-energy cathodoluminescence of vacuum fluorescent display and field emission display screens
S. A. Bukesov, N. V. Nikishin, A. O. Dmitrienko, S. L. Shmakov, and J. M. Kim 2082

Resistivity effect of $ZnGa_2O_4$:Mn phosphor screen on cathodoluminescence characteristics of field emission display
Shin-Sung Kim, Sung Hee Cho, Jae Soo Yoo, Sung Ho Jo, and Jong Duk Lee 2086

Critical role of degassing for hot aluminum filling
J. Proost, E. Kondoh, G. Vereecke, M. Heyns, and K. Maex 2091

Rapid Communications

Analysis of scanning force microscope force-distance data beyond the Hookian approximation
Fredy R. Zypman and Steven J. Eppell 2099

Microtrenching resulting from specular reflection during chlorine etching of silicon
Robert J. Hoekstra, Mark J. Kushner, Valeriy Sukharev, and Phillipe Schoenborn 2102

Aligned pipe arrays formation by silicon anodic etching
M. C. dos Santos and O. Teschke 2105

Shop Notes

Side-by-side wafer bonding of InP for use with stepper-based lithography
R. W. Ryan, R. F. Kopf, A. Tate, J. Burm, and R. A. Hamm 2110

Simple technique for measuring grating periods made using e-beam lithography
K. S. Feder and D. M. Tennant 2113

Brief Reports and Comments

Electrical characteristics of Ta_2O_5 films on Si prepared by dc magnetron reactive sputtering and annealed rapidly in N_2O
G. Eftekhari 2115

Tungsten silicide and tungsten polycide anisotropic dry etch process for highly controlled dimensions and profiles
R. Bashir, A. E. Kabir, F. Hebert, and C. Bracken 2118

Perturbation solution for the interfacial oxidation of silicon
Vijay R. Mhetar and L. A. Archer 2121

Atomic force microscopy observation of the ferroelectric domain structure on the (010) cleavage surface of K_2ZnCl_4
D.-Y. Kim, J. W. Hong, S.-I. Kwun, and S.-Y. Jeong 2125

***In situ* ultraviolet illumination of porous silicon during scanning tunneling microscopy**
D. Schwall, F. A. Otter, and J. M. Galligan 2127

Papers from the 25th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Preface 2133

Silicon and Ultrathin Oxides

What can electron paramagnetic resonance tell us about the Si/SiO_2 system?
P. M. Lenahan and J. F. Conley, Jr. 2134

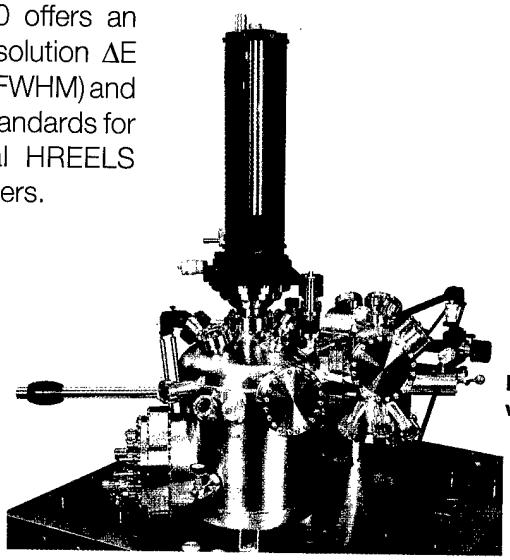
Interface formation and thermal stability of advanced metal gate and ultrathin gate dielectric layers
B. Claflin and G. Lucovsky 2154

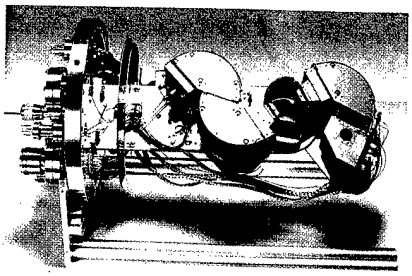
Formation of ultrathin oxynitride layers on Si(100) by low-temperature electron cyclotron resonance N_2O plasma oxynitridation process
Sahana Chakraborty, Toshiyuki Yoshida, Tamotsu Hashizume, Hideki Hasegawa, and Takamasa Sakai 2159

(Continued)

IB 500 High Resolution EELS

The IB 500 offers an ultimate resolution ΔE < 0.5 meV (FWHM) and sets new standards for commercial HREELS spectrometers.





MULIPROBE *HREELS XP*
with IB 500

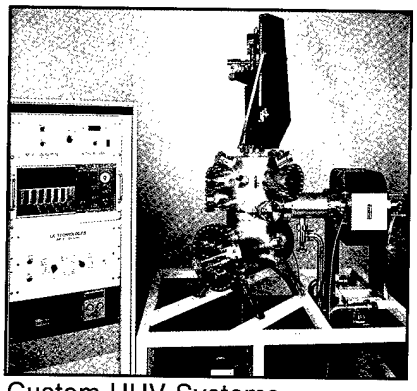
OMICRON
INSTRUMENTS FOR SURFACE SCIENCE

Headquarters:
OMICRON VAKUUMPHYSIK GmbH
Idsteiner Str. 78 • D-65232 Taunusstein • Germany
Tel. +49 (0) 6128 9 87 - 0 • Fax +49 (0) 6128 9 87 - 185

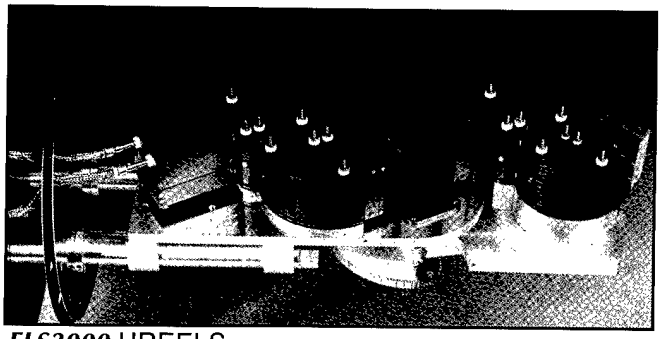
USA:
OMICRON Associates
1226 Stoltz Road • Bethel Park, PA 15102
Tel. (412) 831-2262 • Fax (412) 831-9828

http://www.omicron-instruments.com/

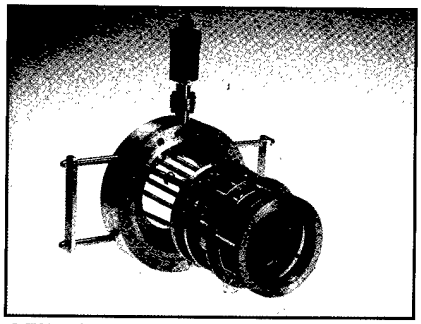
Surface Analysis Components and Systems Since 1985



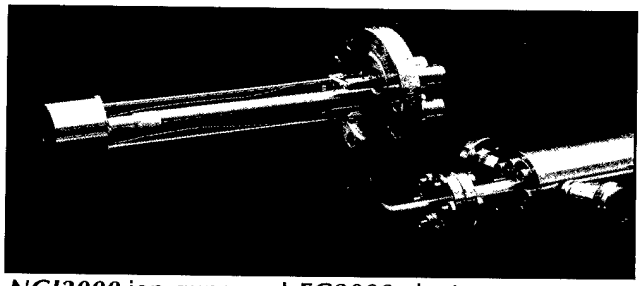
Custom UHV Systems



ELS3000 HREELS



LEED/Auger



NGI3000 ion guns and ***EG3000*** electron guns

LK TECHNOLOGIES

LK Technologies, Inc.
3910 Roll Avenue, Bloomington, IN 47403
Tel (812) 332-4449
Fax (812) 332-4493
http: //www.lktech.com
e-mail: lktech@lktech.com

Surface roughness of SiO_2 from a remote microwave plasma enhanced chemical vapor deposition process 2165
M. J. Rack, D. Vasileska, D. K. Ferry, and M. Sidorov

Investigation of postoxidation thermal treatments of Si/SiO_2 interface in relationship to the kinetics of amorphous Si suboxide decomposition 2171
B. J. Hinds, F. Wang, D. M. Wolfe, C. L. Hinkle, and G. Lucovsky

Cathodoluminescence measurements of suboxide band-tail and Si dangling bond states at ultrathin $Si–SiO_2$ interfaces 2177
A. P. Young, J. Schäfer, G. H. Jessen, R. Bandhu, L. J. Brillson, G. Lucovsky, and H. Niimi

Effects of interface roughness and conducting filaments in metal–oxide–semiconductor tunnel structures 2182
D. Z.-Y. Ting and T. C. McGill

Equilibrium shape of $CoSi_2$ hut clusters on Si(100) 2188
S. H. Brongersma, M. R. Castell, D. D. Perovic, and M. Zinke-Allmang

Heterointerface dipoles: Applications to (a) $Si–SiO_2$, (b) nitrided $Si–N–SiO_2$, and (c) $SiC–SiO_2$ interfaces 2191
G. Lucovsky, H. Yang, and H. Z. Massoud

Multilayer oxidation of AlAs by thermal and electron beam induced decomposition of H_2O in ultrahigh vacuum 2199
S. I. Yi, W. J. Mitchell, C.-H. Chung, E. L. Hu, and W. H. Weinberg

GaN and Related Nitrides

Exciton localization in InGaN quantum well devices 2204
Shigefusa Chichibu, Takayuki Sota, Kazumi Wada, and Shuji Nakamura

Dynamics of localized excitons in InGaN/GaN quantum wells 2215
Hongbin Yu, H. Htoon, Alex deLozanne, C. K. Shih, P. A. Grudowski, R. D. Dupuis, K. Zeng, R. Mair, J. Y. Lin, and H. X. Jiang

Investigation of the chemistry and electronic properties of metal/gallium nitride interfaces 2218
C. I. Wu and A. Kahn

Negative electron affinity of cesiated *p*-GaN(0001) surfaces 2224
M. Eyckeler, W. Mönch, T. U. Kampen, R. Dimitrov, O. Ambacher, and M. Stutzmann

Nucleation and growth of GaN layers on GaAs, Si, and SiC substrates 2229
Klaus H. Ploog, Oliver Brandt, Hui Yang, Bin Yang, and Achim Trampert

Reflection high-energy electron diffraction studies of wurtzite GaN grown by molecular beam epitaxy 2237
O. H. Hughes, D. Korakakis, T. S. Cheng, A. V. Blant, N. J. Jeffs, and C. T. Foxon

Reconstructions of GaN(0001) and $(000\bar{1})$ surfaces: Ga-rich metallic structures 2242
A. R. Smith, R. M. Feenstra, D. W. Greve, M. S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup

Soft x-ray emission studies of the bulk electronic structure of AlN, GaN, and $Al_{0.5}Ga_{0.5}N$ 2250
Kevin E. Smith, Laurent C. Duda, Cristian B. Stagarescu, James Downes, Dimitris Korakakis, Raj Singh, Theodore D. Moustakas, Jinghua Guo, and Joseph Nordgren

Formation of an Sb–N compound during nitridation of InSb (001) substrates using atomic nitrogen 2254
L. Haworth, J. Lu, P. Hill, D. I. Westwood, J. E. Macdonald, N. Hartmann, A. Schneider, and D. R. T. Zahn

Influence of growth conditions, inversion domains, and atomic hydrogen on growth of $(000\underline{1})$ GaN by molecular beam epitaxy 2261
T. H. Myers, L. S. Hirsch, L. T. Romano, and M. R. Richards-Babb

Magnetic Materials and BEEM

Epitaxial growth and properties of III–V magnetic semiconductor (GaMn)As and its heterostructures 2267
Masaaki Tanaka

Ferromagnetic resonance imaging of Co films using magnetic resonance force microscopy 2275
B. J. Suh, P. C. Hammel, Z. Zhang, M. M. Midzor, M. L. Roukes, and J. R. Childress

***In situ* controlled reactions and phase formation of thin films on GaAs** 2280
D. A. Caldwell, L. C. Chen, A. H. Bensaoula, J. K. Farrer, C. B. Carter, and C. J. Palmstrøm

Metal/GaN Schottky barriers characterized by ballistic-electron-emission microscopy and spectroscopy 2286
L. D. Bell, R. P. Smith, B. T. McDermott, E. R. Gertner, R. Pittman, R. L. Pierson, and G. J. Sullivan

Ballistic electron emission microscopy spectroscopy study of AlSb and InAs/AlSb superlattice barriers 2291
X.-C. Cheng and T. C. McGill

(Continued)

READ TODAY'S BEST NONLINEAR RESEARCH ON THE WEB FREE

Select articles from the journal Chaos are at your fingertips

Already recognized for its accessible presentation of the latest international research on nonlinear phenomena, **Chaos** reaches out to you in a whole new way. You can now read and download—at no cost—some of the finest recent articles from this leading interdisciplinary journal published by the American Institute of Physics. Simply visit the journal's Web site at:

www.aip.org/journals/chaos

There's even more **FREE** online access for you when you subscribe to the print edition, including:

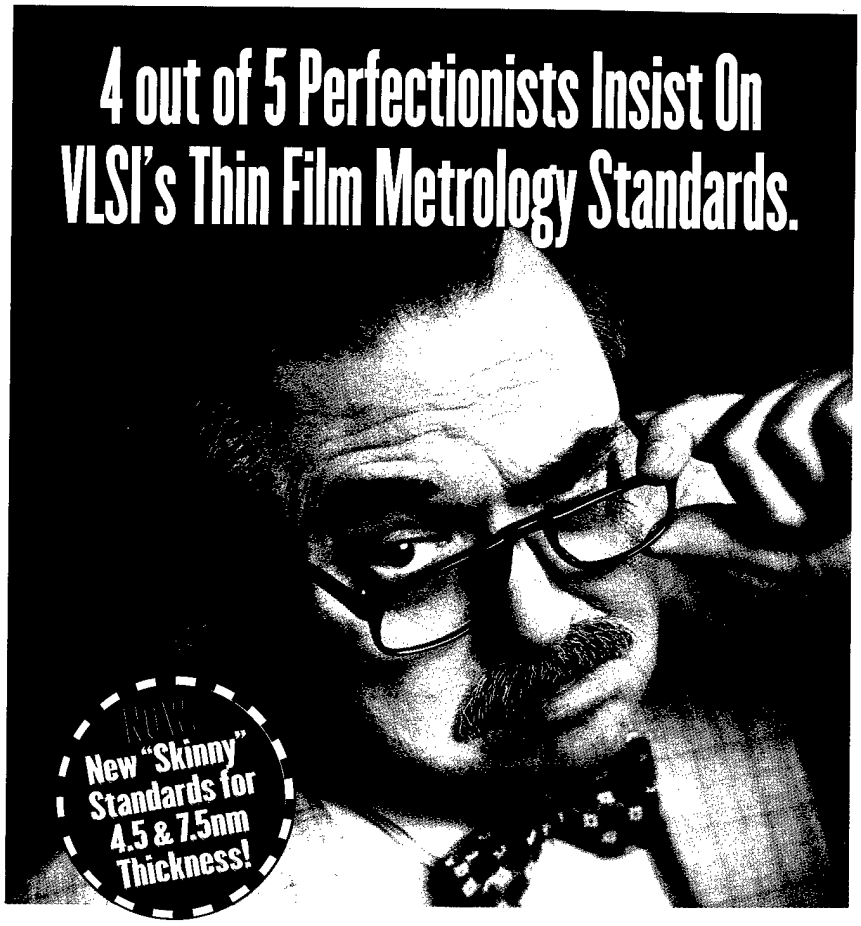
- Access to an online version of **Chaos** up to 3 weeks before print publication.
- Online access to every **Chaos** issue ever published! The complete **Chaos** archive includes 7 years of cutting-edge nonlinear research — and it's yours to tap any time you want via the Web.
- A subscription to PINET Plus, giving you access to AIP's SPIN abstracts database, plus searchable directories from leading science organizations, and tables of contents for AIP and Member Society publications.

For more information and subscription rates call toll free: (800) 344-6902, fax: (516) 349-9704, or e-mail to: mktg@aip.org.

AMERICAN INSTITUTE OF PHYSICS

500 Sunnyside Boulevard Woodbury, NY11797

4 out of 5 Perfectionists Insist On VLSI's Thin Film Metrology Standards.



If you're responsible for thin film thickness measurements, you want them to be right. And you definitely don't want to be embarrassed by a metrology tool that decides to drift at a critical time.

That's why perfectionists insist on VLSI's suite of thin-film metrology standards. For silicon dioxide and silicon nitride. The broadest selection in the industry.

And now, oxide standards are available for 4.5nm and 7.5nm! It's a VLSI exclusive.

So if you're a metrology perfectionist, flaunt it! Call now for your free "Good Enough ISN'T" button along with your free VLSI catalog...

VLSI Standards: (800) 228-8574. Or on the Internet: www.vlsistd.com

Good Enough VLSI

VLSI Standards Incorporated

The Measurement Standards for the Industry.

Current oscillations in thin metal–oxide–semiconductor structures observed by ballistic electron emission microscopy
H. J. Wen, R. Ludeke, and Andreas Schenk 2296

Investigation of ultrathin SiO_2 film thickness variations by ballistic electron emission microscopy
B. Kaczer, H.-J. Im, and J. P. Pelz 2302

Compliant Substrates, Schottky Barriers, and Interface States

Compliant substrate technology: Status and prospects
April S. Brown 2308

Interface states induced in GaAs by growth interruption during an *in situ* process
F. Wakaya, T. Matsubara, M. Nakayama, J. Yanagisawa, Y. Yuba, S. Takaoka, K. Murase, and K. Gamo 2313

Influence of sulfur interlayers on the Mg/GaAs(100) interface formation
St. Hohenecker, T. U. Kampen, D. R. T. Zahn, and W. Braun 2317

Direct imaging of the evolving Au/InSb(111) *B* interface
Tetsuya Mishima, Jun Nakamura, Kenji Tsukada, Masayasu Nishizawa, Toyoaki Eguchi, and Toshiaki Osaka 2324

Microwave modulated photoluminescence used to measure surface recombination velocities
C. E. Inglefield, M. C. DeLong, P. C. Taylor, and W. A. Harrison 2328

Optical Methods and II–VI Materials

Native extended defects in $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures
B. Müller, S. Heun, R. Lantier, S. Rubini, J. J. Paggel, L. Sorba, A. Bonanni, M. Lazzarino, B. Bonanni, A. Franciosi, E. Napolitani, F. Romanato, A. Drigo, J.-M. Bonard, J.-D. Ganière, L. Lazzarini, and G. Salviati 2334

***In situ* determination of in-plane strain anisotropy in ZnSe(001)/GaAs layers using reflectance difference spectroscopy**
K. Hingerl, T. Yasuda, T. Hanada, S. Miwa, K. Kimura, A. Ohtake, and T. Yao 2342

Interpretation of reflectance anisotropy spectroscopy spectra of ZnSe(001) grown on GaAs(001) in terms of bulk, interface, and surface contributions
A. M. Frisch, Ch. Schultz, T. Herrmann, V. Emiliani, D. Wolffframm, D. A. Evans, M. Korn, U. Rossow, N. Esser, and W. Richter 2350

Reflectance difference spectroscopy spectra of clean (3×2), (2×1), and *c*(2×2) 3*C*-SiC(001) surfaces: New evidence for surface state contributions to optical anisotropy spectra
U. Rossow, K. Lindner, M. Lübbe, D. E. Aspnes, and D. R. T. Zahn 2355

Processes of quantum dot formation in the InAs on GaAs(001) system: A reflectance anisotropy spectroscopy study
D. I. Westwood, Z. Sobiesierski, C. C. Matthai, E. Steimetz, T. Zettler, and W. Richter 2358

Photon-induced localization and final-state correlation effects in optically absorbing materials
D. E. Aspnes, L. Mantese, K. A. Bell, and U. Rossow 2367

Epitaxy and Heteroepitaxy

Reflection high-energy electron diffraction/scanning tunneling microscopy study of InAs growth on the three low index orientations of GaAs: Two-dimensional versus three-dimensional growth and strain relaxation
B. A. Joyce, T. S. Jones, and J. G. Belk 2373

Structure of InAs/AlSb/InAs resonant tunneling diode interfaces
B. Z. Nosho, W. H. Weinberg, J. J. Zinck, B. V. Shanabrook, B. R. Bennett, and L. J. Whitman 2381

Scanning tunneling microscopy and x-ray photoelectron spectroscopy studies of atomic level structure and Fermi level pinning on GaAs(110) surfaces grown by molecular beam epitaxy
Yasuhiko Ishikawa, Naohiro Tsurumi, Takashi Fukui, and Hideki Hasegawa 2387

Atomic-scale compositional structure of InAsP/InP and InNAsP/InP heterostructures grown by molecular-beam epitaxy
S. L. Zuo, W. G. Bi, C. W. Tu, and E. T. Yu 2395

Chemistry of $[(t\text{-butyl})GaS]_4$ on Si(100)-(2×1)
R. I. Pelzel, A. B. Hopcus, J. H. G. Owen, B. Z. Nosho, and W. H. Weinberg 2399

New model for reflection high-energy electron diffraction intensity oscillations
W. Braun, L. Däweritz, and K. H. Ploog 2404

Anisotropic microstructure development during the reaction of Mg with GaAs
S. W. Robey 2413

(Continued)

Take a look from the inside...

Process Gas Purity

Process Contaminants

Base Pressure Residuals

VACUUM PROCESSING

Leak Checking

Pump Down Profiles

Vacuum Diagnostics

Hiden's HPR-30 - See The Whole Picture

For hands free, automatic data acquisition in base pressure residual gas analysis, vacuum diagnostics and in-situ process gas analysis the HPR-30 offers real analytical capability in a compact RGA package. Our one-touch **"point and click"** Windows™ MASsoft PC software provides full multi-tasking capability, a range of networking options, *"intelligent trip"* process control and instant *"process views"* from our integral Scan Gallery.

For further details visit our Web site at: **http://www.hiden.co.uk**
or e-mail: **info@hiden.demon.co.uk**

HIDEN ANALYTICAL
Excellence in Analytical Technology

CORPORATE OFFICE AND SERVICE CENTER: HIDEN ANALYTICAL INC., 75 HANCOCK ROAD, SUITE D, PETERBOROUGH, NH 03458-1100. TEL: (603) 924 5008 FAX: (603) 924 5009
MANUFACTURED IN ENGLAND BY: HIDEN ANALYTICAL LTD., 420 EUROPA BOULEVARD, WARRINGTON WA5 5UN, ENGLAND. TEL: +44 (0) 1925 445225 FAX: +44 (0) 1925 416518

For Convenient Training at Your Facility ...

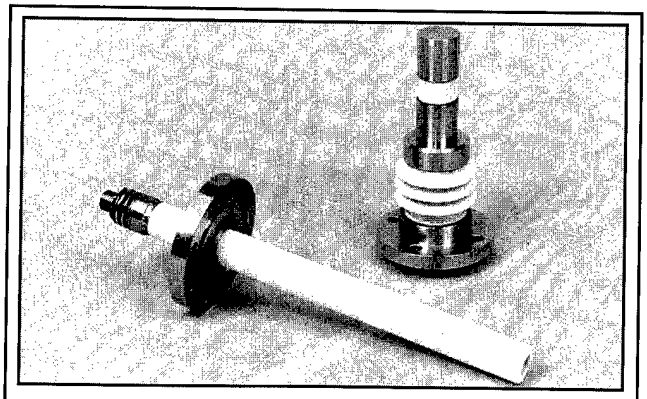
Try AVS Industrial Training Videos

We bring vacuum technology training directly to your workplace—on video—for convenient employee training, review, and self-paced learning. Training programs currently available are:

- *Properties of Vacuum System Materials*
- *Sputter Deposition and Ion Beam Processes*
- *Fundamentals of Capture Pumping*

Each program provides 6-8 hours of training and includes 20 illustrated workbooks.

Order your tapes today! Call the AVS at 212-248-0200.



Omley Industries

Vacuum Braze Technology Specialists

Metal to Metal & Ceramic to Metal Applications

For over 35 years Omley Industries has been and continues to be, a leader in custom and production vacuum brazing.
Whatever your needs in:
- Feedthrus: Ultra high vacuum, high voltage, high pressure, high current, instrumentation, corrosive, liquid and cryogenic.
- Windows: Beryllium, sapphire, diamond and stainless steel.
- Chambers: Flanged, evacuated, dewar and beryllium.

1-800-541-3355 • Fax: 1-800-717-3355
See our web site at, www.empnet.com/omley

Equilibrium critical thickness for strained-layer growth 2417
K. E. Khor and S. Das Sarma

Surface roughness, strain, and alloy segregation in lattice-matched heteroepitaxy 2421
Catherine Priester and Genevieve Grenet

Chemical bonding features for faultily stacked interfaces of GaAs{111} 2426
Jun Nakamura, Tetsuya Mishima, Moto-hisa Masui, Mineo Sawayanagi, Sung-Pyo Cho, Masayasu Nishizawa, Toyoaki Eguchi, and Toshiaki Osaka

Papers from the Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing 2438

Preface

General Interest

Material-wave nanotechnology: Nanofabrication using a de Broglie wave 2439
Shinji Matsui and Jun-Ichi Fujita

Economic and technical case for ion projection lithography 2444
I. L. Berry

Microelectromechanical tunneling sensor fabrication and post-processing characterization using focused ion beams 2449
F. P. Stratton, R. L. Kubena, H. H. McNulty, R. J. Joyce, and J. Vajo

Sources and Systems

Multipurpose nanobeam source with supertip emitter 2455
S. Kalbitzer and A. Knoblauch

Focused ion beam optical column design and consideration on minimum attainable beam size 2462
Kiyoshi Sakaguchi and Tetsu Sekine

Overlay accuracy tests for direct write implantation 2469
S. Mogren and I. L. Berry

Development of ion and electron dual focused beam apparatus for high spatial resolution three-dimensional microanalysis of solid materials 2473
Zhaohui Cheng, Tetsuo Sakamoto, Masanori Takahashi, Yasuyuki Kuramoto, Masanori Owari, and Yoshimasa Nihei

Key technologies of a focused ion beam system for single ion implantation 2479
T. Matsukawa, T. Shinada, T. Fukai, and I. Ohdomari

Development of wide range energy focused ion beam lithography system 2484
M. Kinokuni, H. Sawaragi, R. Mimura, R. Aihara, and A. Forchel

Current status of single ion implantation 2489
Takahiro Shinada, Yoshinori Kumura, Jun Okabe, Takashi Matsukawa, and Iwao Ohdomari

Beam Induced Chemical and Physical Effects

Chemically and geometrically enhanced focused ion beam micromachining 2494
P. E. Russell, T. J. Stark, D. P. Griffis, J. R. Phillips, and K. F. Jarausch

Focused ion beam technology applied to microstructure fabrication 2499
Michael J. Vasile, Raja Nassar, and Jushan Xie

Focused-ion-beam-assisted etching of diamond in XeF_2 2506
Jun Taniguchi, Naoto Ohno, Shuuichi Takeda, Iwao Miyamoto, and Masanori Komuro

Focused ion beam etching of resist/Ni multilayer films and applications to metal island structure formation 2511
Masayoshi Nakayama, Fujio Wakaya, Junichi Yanagisawa, and Kenji Gamo

Focused ion beam direct deposition and its applications 2515
Shinji Nagamachi, Masahiro Ueda, and Junzo Ishikawa

Microanalysis

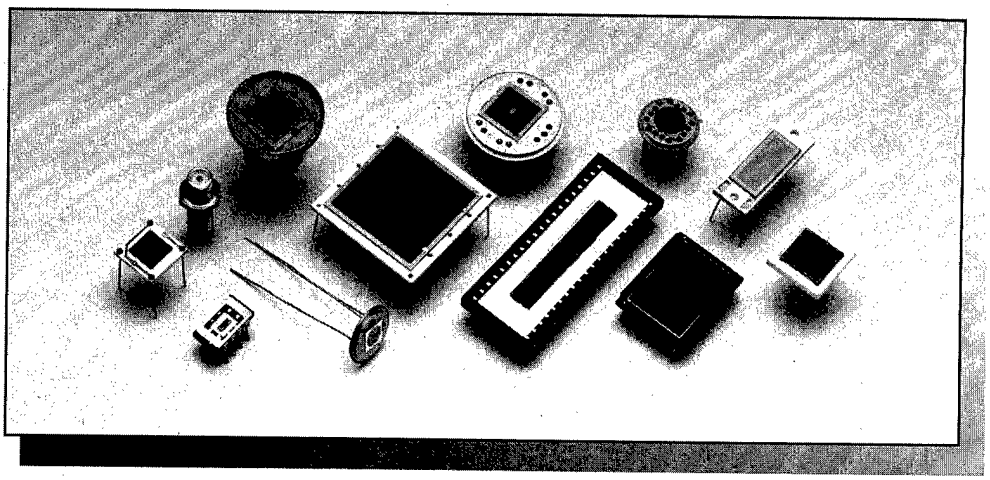
Transmission electron microscopy observation of thin foil specimens prepared by means of a focused ion beam 2522
H. Saka

Dynamic Monte Carlo simulation for depth profiling by ion-sputter etching: Application to the AlAs/GaAs multilayered system 2528
Hyung-Ik Lee and Ryuichi Shimizu

(Continued)

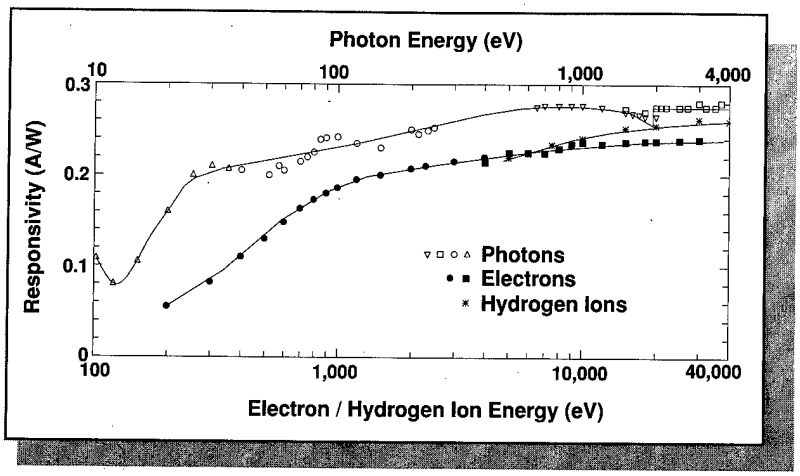
ABSOLUTE

XUV SILICON PHOTODIODES



Features

- 100% Internal Quantum Efficiency
- 8 Decades of Linearity
- No Change in QE with 100 Mrad (Si) 124 eV Photons
- May be Operated Without Bias
- Cryogenically and UHV Compatible
- With Integrated Bandwidth Limiting Filters



TYPICAL RESPONSIVITY OF AXUV PHOTODIODES TO PHOTONS, ELECTRONS AND HYDROGEN IONS

INTERNATIONAL RADIATION DETECTORS INC.
2545 West 237th Street, Unit I ◆ Torrance, California 90505-5229
(310) 534-3661 ◆ FAX (310) 534-3665 ◆ email: ird@kaiwan.com

IRD

DEVELOPED IN COLLABORATION WITH NIST, NIH, LLNL, LANL, NCAR

Proposals for exact-point transmission-electron microscopy using focused ion beam specimen-preparation technique
T. Ishitani, Y. Taniguchi, S. Isakozawa, H. Koike, T. Yaguchi, H. Matsumoto, and T. Kamino 2532

Nanofabrication, Devices and Implantation

GaAs microcrystal growth on semiconductor surfaces by low energy focused ion beam
Toyohiro Chikyow and Nobuyuki Koguchi 2538

AlGaAs/GaAs heterostructures grown on a focused-Be-ion-beam written backgate
Y. Hirayama and T. Saku 2543

Formation of GaAs/AlGaAs constricted-channel field-effect transistor structures by focused Ga implantation and transport of electrons via focused ion beam induced localized states
H. Kim, T. Noda, and H. Sakaki 2547

Hopping conduction and localized states in p-Si wires formed by focused ion beam implantations
Hirotaka Iwano, Shigeaki Zaima, and Yukio Yasuda 2551

End-point detection using focused ion beam-excited photoemissions in milling deep small holes in large scale integrated circuit structures
Hiroshi Yamaguchi, Keiya Saitoh, and Michinobu Mizumura 2555

Focused ion beam implantation for opto- and microelectronic devices
H. König, N. Mais, E. Höfling, J. P. Reithmaier, A. Forchel, H. Müssig, and H. Brugger 2562

Direct writing of active loads by focused ion beams
C. Wiemann, M. Versen, and A. D. Wieck 2567

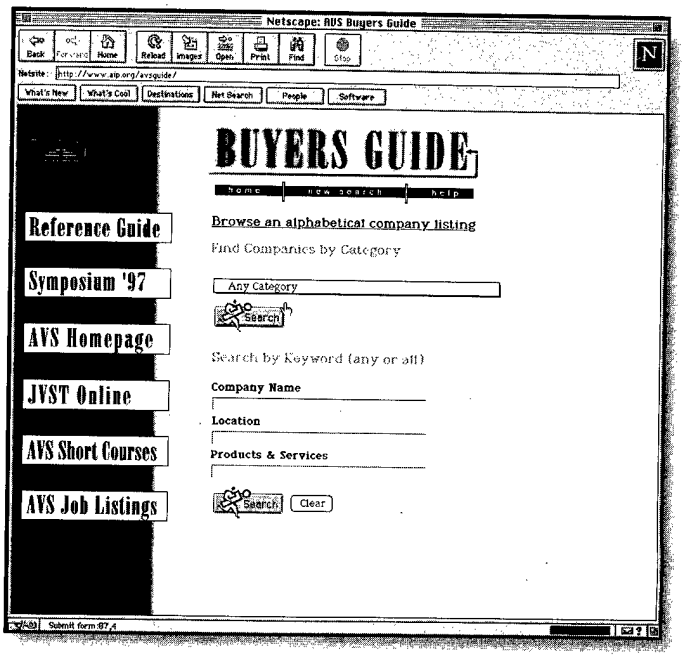
***In situ* scanning tunneling microscope studies of high-energy, focused ion implantation of Ga into GaAs: Direct observation of ion beam profiles**
G. A. C. Jones, P. D. Rose, and S. Brown 2570

Ion beam synthesis of cobalt disilicide using focused ion beam implantation
J. Teichert, L. Bischoff, and S. Hausmann 2574

AUTHOR INDEX 2578

CUMULATIVE AUTHOR INDEX 2579

Vacuum specialists and specifiers of vacuum systems, equipment, and services...
Bookmark_this_page!



When you pick out a novel you want a real page-turner. But when you need information, you want it fast—you can't waste time thumbing through cumbersome directories. That's why the American Vacuum Society has created an information-packed **Buyers Guide** and **Reference Guide** and put them online. Now whether you need critical reference data or helpful product information, you'll enjoy the speed and versatility of these new electronic editions.

www.aip.org/avsguide

The AVS Marketplace Buyers Guide and Reference Guide. Your one-stop online resource for all your vacuum needs.

With the **Buyers Guide** you get:

- ✦ Company and product information from more than 250 suppliers of vacuum equipment and services
- ✦ Product directory containing more than seventy categories
- ✦ Sophisticated search engine, which lets you search by company, product category, and location

With the **Reference Guide** you get:

- ✦ Full glossary of vacuum terminology
- ✦ Spectra library
- ✦ Working pressure ranges of vacuum pumps and gauges
- ✦ Vapor pressure data
- ✦ Charts of common and derived units of measurement and throughput leak rates

Log on to the **Buyers Guide** now and take advantage of these two valuable guides in a powerful new electronic format.

AVS

KEEP AN EYE ON AVS EVENTS ...

WWW.VACUUM.ORG

VISIT THE AVS WEB SITE TO:

- Learn more about AVS and its member services, including its local chapters.
- Obtain information on its short courses and symposia and even register online.
- Find employment opportunities.
- Subscribe to the AVS journals, including the *Journal of Vacuum Science and Technology* and *Surface Science Spectra*.
- Read the bimonthly *AVS Newsletter*.
- Order educational training materials.
- Learn about other international vacuum societies and other meetings of interest.

Strain-induced birefringence in $Si_{1-x}Ge_x$ optical waveguides

M. Robillard,[a] P. E. Jessop, and D. M. Bruce
Centre for Electrophotonic Materials and Devices, McMaster University, Hamilton, Canada L8S 4L7

S. Janz, R. L. Williams, S. Mailhot, and H. Lafontaine
Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Canada K1A 0R6

S. J. Kovacic[b] and J. J. Ojha
Nortel Technology, P.O. Box 3511, Station C, Ottawa, Canada K1Y 4H7

(Received 25 August 1997; accepted 17 April 1998)

For the design of $Si_{1-x}Ge_x$ optical waveguide devices, one of the most important material parameters is the refractive index difference, δn, between the alloy layer and the silicon substrate. We have measured δn for pseudomorphic waveguide layers with germanium fractions between 1% and 9% by fitting measured mode profiles to theoretical mode shapes for a wavelength of 1.3 μm. For transverse electric modes, the measured δn varied with composition as $\delta n = (0.34 \pm 0.05)x$. Transverse magnetic modes were more tightly confined to the waveguide layer and the index was determined to be $\delta n = (0.55 \pm 0.05)x$. The large difference between the two polarizations results from strain-induced birefringence. Bulk photoelastic theory, using constants appropriate for pure silicon, predicts strain contributions to the index of $-0.080x$ and $+0.095x$ for light polarized parallel and perpendicular, respectively, to the substrate plane, consistent with experimental observations. © *1998 American Vacuum Society.* [S0734-211X(98)02004-6]

I. INTRODUCTION

The growing interest in silicon photonic components is motivated primarily by the prospect of using well established silicon device processing technology to fabricate devices that perform optical as well as electrical functions. In this context, one of the most attractive approaches to optical waveguide fabrication is to use $Si_{1-x}Ge_x$ epitaxial layers that are kept below the critical thickness[1,2] for pseudomorphic growth so that dislocation formation is avoided. The presence of strain has a significant effect on the refractive index of the $Si_{1-x}Ge_x$ layer, and there are widely varying reports in the literature for the dependence of the refractive index on the germanium fraction, x, in waveguide structures.

While some experimental work has been done in this field,[3] resulting estimates have varied over a wide range. In addition, most of the work that has been done has addressed only transverse electric (TE) polarized light. Figure 1 shows some of the recently reported theoretical predictions and experimental determinations of refractive index variation in strained $Si_{1-x}Ge_x$. Since in many of the proposed applications of strained $Si_{1-x}Ge_x$ optoelectronics the light entering the device will contain both TE and transverse magnetic (TM) polarization, it seems reasonable that the propagation characteristics of both polarizations should be studied prior to the design of any complex device.

In $Si_{1-x}Ge_x$ layers that are sufficiently thin and of sufficiently low germanium content so as to be pseudomorphic, the alloy is strained to match the lattice constant of the silicon substrate. The resulting strain is compressive in the plane of the substrate, and tensile in the perpendicular direction. The strain-induced perturbations in the refractive index, therefore, have opposite signs for light that is polarized parallel or perpendicular to the surface, resulting in a high degree of strain-induced birefringence which is superimposed upon the natural birefringence of the optical waveguide structure.

In this article, we present the results of a study of the refractive index variation in strained $Si_{1-x}Ge_x$ as a function of germanium concentration for both TE and TM polarization. Our findings indicate a strong birefringence in the material, which we show to be consistent with a simple theoretical model. We do not believe that the birefringence of strained $Si_{1-x}Ge_x$ has been previously reported quantitatively in the literature, although it has been proposed as a possible cause of unexpected behavior in previously observed waveguiding results.[4–6]

II. THEORY

In this section, we develop a simple theoretical model that can provide an estimate of the degree of birefringence expected in strained $Si_{1-x}Ge_x$. Two assumptions are made in this model. First, we have used the strain-optic tensor for silicon[7] to calculate the effect of the two strains in the $Si_{1-x}Ge_x$ core layer of the waveguide. Due to the small germanium concentrations in the samples, we expect the strain-optic behavior to closely mimic that of bulk silicon. Second, it is assumed that the strain is confined uniformly to the $Si_{1-x}Ge_x$ core layer. That is, there is no strain at all in the silicon cladding, and the in-plane $Si_{1-x}Ge_x$ lattice constant is matched to silicon everywhere. While it is unrealistic to assume that the silicon cladding does not experience any deformation at all, even at the interface, we expect that the

[a]Present address: Newbridge Networks, Kanata, Ontario, Canada.
[b]Present address: SiGe Microsystems, Ottawa, Ontario, Canada.

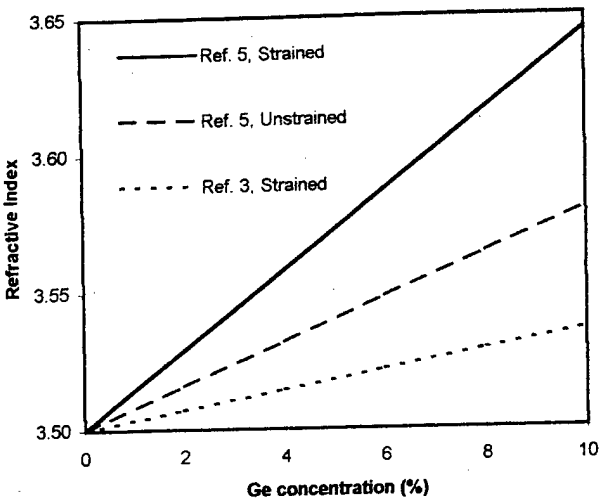


FIG. 1. Refractive index values of silicon–germanium for 1.3 μm light. The values from Ref. 5 are estimated theoretically for strained and unstrained silicon–germanium layers, while the data from Ref. 3 has been experimentally determined for the TE polarization.

departures from the ideal case described above are sufficiently small to ignore for the purposes of this calculation. A more rigorous theoretical examination of the strain-optic relationship in this structure would require numerical modeling of the strain field in the vicinity of the core layer.

In the $Si_{1-x}Ge_x$ core layer of the waveguides, there is an in-plane biaxial stress such that $\sigma_1=\sigma_2$, where σ is the stress tensor, and 1, 2, and 3 refer to the principal axes of the crystal. σ_3, the stress perpendicular to the growth plane, is zero. The optical propagation directions are $\langle 110\rangle$. However, since the in-plane stresses, σ_1 and σ_2, are equal, a calculation that assumes propagation along a principal axis will yield the correct result.

Because we assume that the $Si_{1-x}Ge_x$ layer is perfectly lattice matched to the (unstrained) silicon substrate and cladding layers, it is possible to calculate the strain in the plane of the film, e_1 and e_2. It is simply calculated by assuming that the lattice constant of unstrained $Si_{1-x}Ge_x$ varies linearly with x. Since the lattice mismatch in between silicon and germanium is 4.16%, the strain, $e_1(=e_2)$, is $e_1=-0.0416x$. The strain is related to the stress by the compliance tensor, S_{ij}:

$$e_i=S_{ij}\sigma_j. \tag{1}$$

It is therefore possible to generate a relationship between e_1 (or e_2) and e_3:

$$e_1=e_2=(S_{11}+S_{12})\sigma_1, \tag{2}$$

$$e_3=2S_{12}\sigma_1=2S_{12}e_1/(S_{11}+S_{12}), \tag{3}$$

since for silicon's compliance tensor, $S_{13}=S_{23}=S_{12}$. Substituting the numerical values[8] $S_{11}=0.7691\times10^{-12}$ cm^2/dyne, and $S_{12}=-0.2142\times10^{-12}$ cm^2/dyne into Eq. (3), gives $e_3=-0.772e_1$.

For the relationship between the strain and refractive index, we use the dielectric impermeability tensor, B, and the strain-optic tensor, p. The strain-optic tensor relates the strain in the crystal to a change in the dielectric impermeability tensor

$$\Delta B_i=p_{ij}e_j. \tag{4}$$

The change in refractive index is then derived from the dielectric impermeability tensor

$$\delta n_i=-0.5n^3\Delta B_i. \tag{5}$$

The coefficients for the strain-optic tensor for silicon are[7]

$$p_{11}=-0.094, \qquad p_{12}=0.0017, \qquad p_{44}=-0.051.$$

Using $n=3.5$, an appropriate approximation for low germanium content $Si_{1-x}Ge_x$ at $\lambda=1.3$ μm, we find that $\delta n_1=\delta n_2=+1.932e_1$ and $\delta n_3=-2.286e_1$.

The strain-induced refractive index changes for TE and TM guided light are given by δn_1 and δn_3, respectively. Substituting the value for e_1, as a function of germanium content, x, we find

$$\delta n_{\mathrm{TE}}=-0.080x, \qquad \delta n_{\mathrm{TM}}=0.095x,$$

or

$$n_{\mathrm{TM}}-n_{\mathrm{TE}}=0.175x.$$

This simple model makes two important predictions. First, it implies that the TM mode in a strained $Si_{1-x}Ge_x$ waveguide is more tightly bound that the TE mode due to the higher refractive index experienced by the TM polarization. Second, the model predicts that the birefringence in the material is quite large. Experimentally, the presence of such a strong polarization dependence should be simple to confirm.

TABLE I. Slab waveguide samples used in this study. Each sample consisted of a coherently strained SiGe layer grown on a silicon substrate. A silicon cladding layer was then grown over the SiGe core.

% Germanium	Core thickness (±0.05 μm)	Cladding thickness (±0.05 μm)
1.2	0.97	1.06
1.8	1.05	0.98
2.6	1.00	0.85
3.6	0.50	1.48
6.8	0.28	1.50
9.0	0.19	1.50

III. EXPERIMENT

As shown in Table I, this study examined slab waveguide samples with germanium concentrations ranging from 1.2% to 9% and thickness ranging from 0.19 to 1.0 μm. In each case, the core layer germanium content and thickness were chosen so that the samples would be below the metastable critical thickness for pseudomorphic growth. Wafers were grown by the National Research Council of Canada and an external vendor by both molecular beam epitaxy (MBE) and chemical vapor deposition (CVD) techniques. In all cases, the composition of the samples was verified using x-ray techniques and the thicknesses of the core and cladding layers

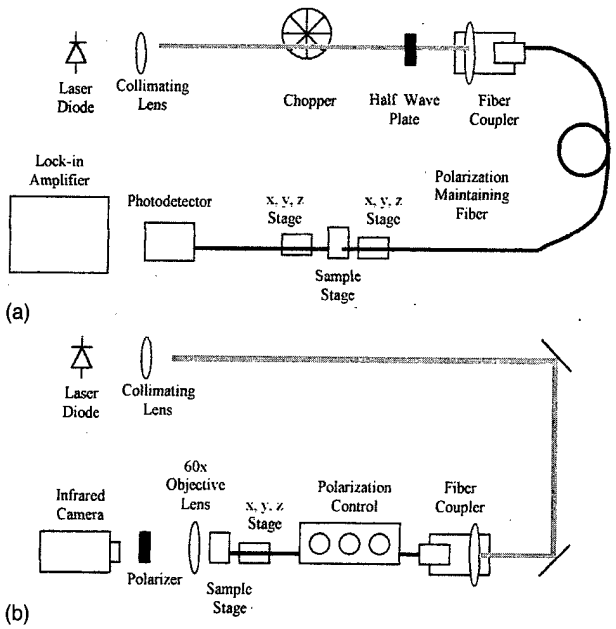


FIG. 2. Two experimental setups used to measure the mode profiles of silicon–germanium slab waveguides in this study. The configuration shown in (a) was used at the NRC, while the one depicted in (b) was used at McMaster University.

were confirmed using scanning electron microscopy (SEM). Results from the x-ray measurements confirmed that the $Si_{1-x}Ge_x$ layer was not relaxed in any of the samples.

For each of the waveguide samples, the refractive index of the $Si_{1-x}Ge_x$ core layer was determined by comparing measured profiles of the bound mode to theoretically generated eigenmode solutions. The theoretical solutions were calculated analytically using the transfer matrix technique.[9] Rather than fitting to the entire mode, exponential curves of the form $y=Ae^{Bx}$ were fit to the tail of the mode extending into the substrate. This region was chosen because it generally has a gradual decay, and thus tends to be less sensitive to noise than other areas. In addition, the decay of the substrate tail displays a wide variation as a function of core index, making it an ideal indicator for this investigation.

The mode profiles were measured using two different techniques, as illustrated in Figs. 2(a) and 2(b) (carried out at NRC and McMaster University, respectively). Both configurations allow the measurement of a mode profile, and both are capable of controlling and monitoring the polarization of the light.

The first configuration [Fig. 2(a)] uses a tapered fiber to introduce light into the waveguide, and a second tapered fiber mounted on a piezo-electrically controlled stage to scan across the output facet of the guide. The power in the output fiber is measured at each point in the scan, constructing a one- or two-dimensional image of the near field intensity. Typical image resolution is 1.0 μm, limited by the fiber tip focal spot size. Using two fibers—one for input and one for output—the transfer function associated with the output of the setup can be estimated. This is accomplished by scanning the output fiber across the input fiber directly. The transfer function thus obtained can be used to deconvolve the measured near field of the waveguide in order to reconstruct the actual mode. For simplicity, the measured transfer function was fit to a Gaussian, which was then used in the deconvolution.

To ensure that the output fiber is positioned at the correct distance from the output facet, it is necessary to perform a series of scans with the output fiber positioned at various distances from the facet. To accurately reproduce the near field of the guided mode, the waveguide facet must be located at the focal point of the fiber tip. This minimizes the width of the transfer function of the apparatus, and decreases the uncertainty associated with the deconvolution of the measured data. In order to select the scan that most accurately reproduces the near field of the mode, the scan exhibiting the narrowest mode profile is used. Since the spot size of the tapered fiber increases when not properly focused, the narrowest profile indicates the sharpest focus.

The second configuration [Fig. 2(b)] uses a cleaved fiber for the input, and images the output of the waveguide onto an infrared camera. The alignment of the input fiber is optimized by maximizing the intensity of the image on the camera. A 60× microscope objective focuses the near field of the mode with appropriate magnification. In order to digitize the vertical profile of the mode, the camera is turned on its side and an appropriate scan line is extracted from its output and recorded by a digital oscilloscope. The transfer function for this system has been estimated by deconvolving measured modes with Gaussian curves of various widths. Since the side of the mode that faces into the air–silicon interface decays very rapidly, it is easy to place an upper limit on the width of the Gaussian by finding the point at which the deconvolution produces an unreasonably sharp decline in this region. Fortunately, the portion of the curve with which we attempt to determine the index of the core is the large portion of the mode extending into the substrate. This region is often quite insensitive to deconvolution with transfer functions for the mode widths that were encountered. For this setup, proper focus of the near field onto the camera is not a significant problem. Because the entire mode is displayed on a monitor, it is a simple job to obtain optimum focus through the adjustment of the microscope objective.

The linearity of the camera response was checked by recording the same mode profile a number of times with a series of neutral density filters used to attenuate the beam by varying amounts. For high intensities the camera output voltage saturated at around 800 mV. For the profiles that were used to determine the refractive index, the peak signal was always kept below 350 mV and the fitting was restricted to the exponentially decaying substrate tail, between 20% and 80% of the peak height. At these signal levels the camera response is sufficiently linear that no significant additional error is introduced into the measurement.

Although both configurations are able to utilize laser sources of various wavelengths, the majority of the measurements were taken at a wavelength of 1.3 μm. For the first configuration, a tunable diode laser was used and for the

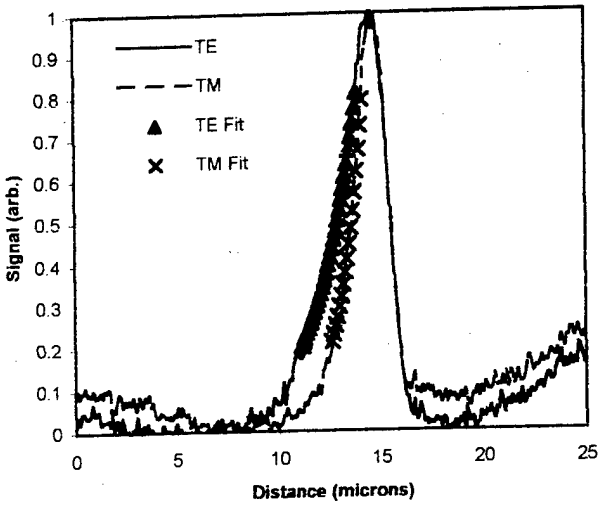


FIG. 3. Mode profiles for the fundamental TE and TM modes of a 2.6% silicon–germanium waveguide. Also shown are the best-fit exponential functions for the substrate tail of the mode.

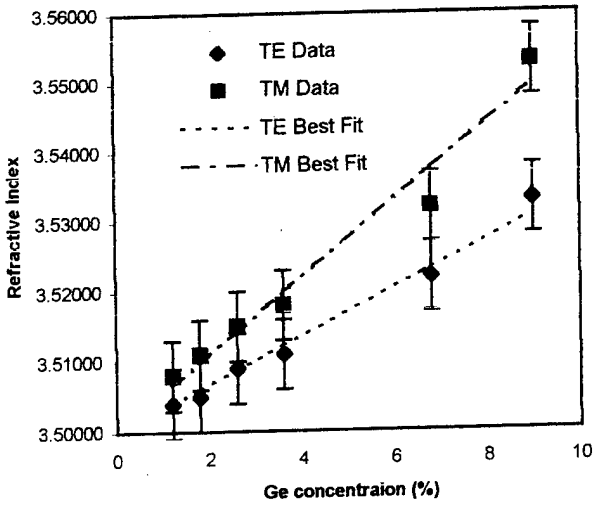


FIG. 4. Refractive index of strained silicon–germanium for TE and TM polarized light at 1.3 μm as a function of Ge concentration.

second, a distributed feedback diode laser was used. For the second setup, some measurements were also taken using a 1.52 μm HeNe laser at McMaster, but since many of the waveguides were found to be incapable of guiding this wavelength, the 1.3 μm sources were chosen as the primary source for this study. In some cases, a 1.15 μm HeNe was also used, but the strongly confined mode that resulted was more difficult to fit reliably.

IV. RESULTS

Significant birefringence was found in each of the $Si_{1-x}Ge_x$ waveguides studied. In each case, the TM mode was found to be more tightly confined to the core layer, indicating a higher refractive index. Figure 3 shows TE and TM mode profiles for a 2.6% germanium sample guiding 1.3 μm light. Also shown are the calculated substrate tails of the modes using the best fit refractive index values. The fit was determined using a least squares algorithm on the substrate tail of the modes. The curves were fit between 20% and 80% intensity levels to minimize the effects of noise.

Figure 4 summarizes the results of this study. Indicated are the refractive indices for TE and TM light as measured by mode profile analysis for each of the samples studied. Also indicated are the best fit lines matched to the experimental data for both polarizations. By fitting the TE and TM refractive indices as a function of germanium content to a straight line, the strain induced component of the refractive index was found to vary as $(0.34\pm0.05)x$ for TE polarized light, and $(0.55\pm0.05)x$ for TM polarized light. These numbers are in good agreement with some previous experimental estimates that examined only TE light.[3] The degree of birefringence found in the measurements closely matches that predicted by our theoretical model. Experimentally, we found that $\delta n_{TM}-\delta n_{TE}=(0.21\pm0.07)x$, while our theoretical prediction was $0.175x$.

V. CONCLUSIONS

We have examined the refractive index dependencies of strained $Si_{1-x}Ge_x$ waveguides over a range of germanium compositions using mode profile analysis. The alloy has been found to be highly birefringent, with the TM modes of the slab waveguides being more tightly confined than the TE modes. The existence of the birefringence, as well as its magnitude, is consistent with theoretical predictions based on bulk photoelastic theory.

[1]J. C. Bean, L. C. Freeman, A. T. Fiory, S. Nakahara, and I. K. Robinson, J. Vac. Sci. Technol. A **2**, 436 (1984).
[2]R. People and J. C. Bean, Appl. Phys. Lett. **47**, 3 (1985).
[3]A. Splett, B. Schupert, K. Petermann, E. Kasper, H. Kibbel, and H.-J. Herzog, *Integrated Photonics Research,* OSA Technical Digest Series, Vol. 10 (Optical Society of America, Washington, DC, 1992), pp. 122–123.
[4]Z. Yang, B. L. Weiss, G. Shao, and F. Namavar, J. Appl. Phys. **77**, 2254 (1995).
[5]R. A. Soref, F. Namavar, and J. P. Lorenzo, Opt. Lett. **15**, 270 (1990).
[6]S. F. Pesarcik, G. V. Treyz, S. S. Iyer, and J. M. Halbout, Electron. Lett. **28**, 159 (1992).
[7]M. J. Weber, *CRC Handbook of Laser Science and Technology* (CRC, Boca Raton, FL, 1986), p. 325.
[8]*Properties of Silicon*, Inspec Emis Datareviews Series No. 4 (Inspec, London, 1988), p. 19.
[9]L. M. Walpita, J. Opt. Soc. Am. A **2**, 595 (1985).

Relaxation of strained, epitaxial $Si_{1-x}Sn_x$

M. F. Fyhn,[a] J. Chevallier, and A. Nylandsted Larsen
Institute of Physics and Astronomy, University of Aarhus, DK-8000 Aarhus C, Denmark

(Received 19 February 1998; accepted 15 May 1998)

The thermal stability of high quality, strained $Si_{1-x}Sn_x$, $2.5\% \leq x \leq 5\%$, grown by molecular beam epitaxy on Si ⟨001⟩ substrates has been investigated by Rutherford backscattering spectrometry and transmission electron microscopy. As a result of annealing at temperatures in the range 400–950 °C for 1 h, both relaxation by precipitation of Sn and generation of misfit dislocations and dislocation loops were found, consistent with the low solubility of Sn in Si and the large lattice mismatch between Si and $Si_{1-x}Sn_x$. In the epitaxial, strained $Si_{1-x}Sn_x$ layers, where the strain is proportional to the Sn concentration, the threshold temperature for generation of precipitates and misfit dislocations was found to decrease with increasing Sn concentration. Above the threshold temperature the influences of the different relaxation channels change considerably with temperature and composition; an orthogonal interfacial misfit dislocation network is seen at temperatures close to threshold while at higher temperatures mainly dislocations in the bulk are found coexisting with large precipitates. The compositional metastability that leads to the precipitation process, reduces the Sn concentration in the matrix, however, neither this depletion of Sn from the matrix of $Si_{1-x}Sn_x$ nor the relaxation due to misfit dislocations cause a fully relaxation of the $Si_{1-x}Sn_x$ layers. © *1998 American Vacuum Society.* [S0734-211X(98)07204-7]

I. INTRODUCTION

Motivated by the possibility of making heterostructure devices which are compatible with conventional Si technology, alloy layers of $Si_{1-x}Ge_x$ have gained considerable interest. To further extend the possibilities of tailoring material properties of Si alloy layers, the two group IV elements C and Sn have come into focus.[1–4]

The group IV alloy containing Sn which has been most intensively studied is $Ge_{1-x}Sn_x$,[5,6] while only recently $Si_{1-x}Sn_x$ of high structural quality was grown.[4,7] The interest in Si and Ge alloy layers containing Sn is due to the fact that Sn in the diamond structure (α-Sn) has a very low (0.08 eV) and direct bandgap making a band-gap reduction possible in an alloy with Sn.[8] Recently the band gap of strained $Si_{0.96}Sn_{0.04}$ was found to be about 0.1 eV smaller than that of Si.[9]

The main difficulties in producing high quality $Si_{1-x}Sn_x$ by epitaxial methods like molecular beam epitaxy (MBE) are due to two factors: the solubility of Sn in Si is less than[10] 0.1% and the difference in lattice constants between α-Sn and Si is large (the lattice constants of Si and α-Sn are 5.43 and 6.49 Å, respectively). These factors necessitate low growth temperatures and low growth rates to suppress the segregation and precipitation of Sn during growth. Ion implantation of Sn into Si substrates is another method that has been tried to produce metastable $Si_{1-x}Sn_x$, however, the required annealing step and the defects left after annealing are serious drawbacks of this method.[11]

Investigations of the structural properties of MBE grown $Si_{1-x}Sn_x$ alloys have only been reported in a few publications and only for as-grown alloys.[4,7,12,13] Recently, the surface morphology of strained and strain-free $Si_{1-x}Sn_x$ layers were investigated and a {113} facetting was identified; the facetting was related to the columnar-structure contrast found in transmission electron microscopy (TEM) micrographs.[12] The columnar structure was interpreted as resulting from small composition fluctuations, but no estimate of the variation in concentration was possible. Mössbauer spectroscopy investigations of enriched $Si_{1-x}Sn_x$ layers revealed that within the sensitivity limit, all Sn atoms are occupying substitutional positions,[13] and Rutherford backscattering (RBS)/channeling spectrometry confirmed that the layers had a high structural quality.[4]

The aim of the present study is to characterize the relaxation channels of strained $Si_{1-x}Sn_x$ as a function of temperature. The strained $Si_{1-x}Sn_x$ are found to be metastable both with respect to the strain and the composition. A knowledge of the temperatures which are harmless during thermal processing is important both for MBE growth purposes as well as for *ex situ* thermal treatments which will take place during the processing of a $Si_{1-x}Sn_x$-based device. Another aspect of the present studies is the possibility of creating metal and semiconductor nanoclusters in Si.

II. EXPERIMENTAL PROCEDURE

Strained, epitaxial $Si_{1-x}Sn_x$ layers of two different compositions, $x=2.5\%$ and $x=5\%$, were grown by MBE on Si ⟨001⟩ substrates. Prior to the growth of $Si_{1-x}Sn_x$, the Si substrates were cleaned following a procedure described in Ref. 14. The $Si_{1-x}Sn_x$ layers were grown at[15] 295 °C with a growth rate of 0.3 Å/s to a thickness of 1700 Å for both compositions. This thickness is many times the estimated equilibrium Matthews–Blakeslee critical thickness[16] (<300 Å), and the layers are therefore metastable with respect to strain. For comparison, a 1700 Å thick Si layer was grown on a Si substrate using the same growth parameters; in the

[a]Electronic mail: fyhn@dfi.aau.dk

following we use the term "Si on Si" when referring to this material. The terms "strained $x=2.5\%$" and "strained $x=5\%$" refer to the $Si_{1-x}Sn_x$ layers of the specified compositions.

The annealings of the $Si_{1-x}Sn_x$ layers were done in a furnace, where a continuous flow of dry N_2 was maintained. The annealing temperatures were in the range 400–950 °C, and the annealing time was always 1 h.

To test the crystalline quality, both RBS/channeling and TEM were performed. The RBS spectra were obtained using 2 MeV He^+ ions and the backscattered ions were detected by a solid state detector placed in a position corresponding to a scattering angle of 161°. The channeling minimum yield,[17] χ_{min}, and the critical angle for channeling,[17] $\Psi_{1/2}$, were determined from angular-yield curves for both the Si and the Sn part of the backscattering spectra. In obtaining the angular-yield curves, the angle of the cone spanned by the surface normal of the sample and the beam direction was increased in steps of 0.1° for angles close to the channeling direction (angle=0°); for each angle 2 μC was accumulated during a rotation of 360°.

TEM analysis was done using a Philips CM20 microscope operating at 200 keV. Plan-view and cross-section samples were prepared following standard techniques.

III. RESULTS

A. RBS/channeling

Figure 1(a) displays the angular-yield curves for three as-grown samples: Si on Si and strained $x=2.5$ and $x=5\%$. The curves are obtained from a depth window, which had a width corresponding to the entire $Si_{1-x}Sn_x$ thickness, positioned in the Si part of the backscattering spectrum. A tilt angle of zero corresponds to the ⟨001⟩ channeling direction. The figure shows that the channeling dip narrows and the minimum yield increases for larger Sn concentrations. A similar narrowing of the angular-yield curves of as-grown $Si_{1-x}Sn_x$ compared to Si grown on Si was found in the ⟨011⟩ channeling direction (not shown). The angular-yield curve obtained from a bulk Si crystal was identical to the angular-yield curve for Si on Si both for the ⟨011⟩ and the ⟨001⟩ directions indicating a high crystalline quality of the Si on Si sample despite the low temperature used during the production of this layer. A comparison of the Si and Sn angular-yield curves obtained from the same layer reveals that these curves are identical for as-grown layers, and therefore no phase separation or preferential shift of either the Si or the Sn atoms have taken place.

The angular-yield curves of Si and Sn from strained $Si_{1-x}Sn_x$ are identical for samples annealed at temperatures $T\leq 600$ °C; Figure 1(b) shows the Si and Sn angular-yield curves for strained $x=5\%$ annealed at 600 °C. In Figs. 1(c) and 1(d) the Si and Sn angular-yield curves are displayed for strained $x=5\%$ annealed at 700 and 800 °C, respectively. Two conclusions can be drawn from a comparison of Figs. 1(b)–1(d): (1) the Si and Sn angular-yield curves deviate more and more for higher annealing temperatures, and (2) both the Si and Sn angular-yield curves narrow and the minimum yield values increase at higher temperatures. A difference between the Si and the Sn angular-yield curves have been found for both compositions after annealing at $T\geq 700$ °C; this indicates that at these temperatures a relatively larger number of Sn atoms than Si atoms are displaced from substitutional positions.

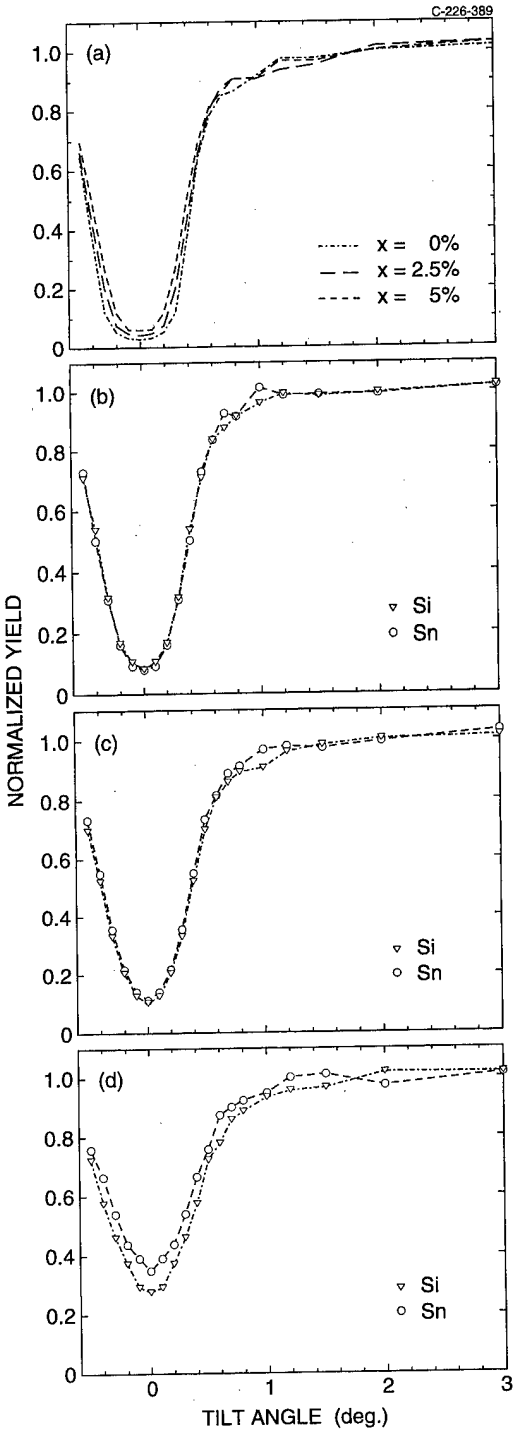


FIG. 1. Channeling angular-yield curves for (a) as-grown, strained $Si_{1-x}Sn_x$ with $x=0\%$, 2.5%, and 5%, and strained $Si_{0.95}Sn_{0.05}$ annealed at (b) 600 °C, (c) 700 °C, and (d) 800 °C.

The change in χ_{min} with annealing temperature is displayed in Fig. 2(a) for strained $x=2.5\%$ and in Figs. 2(b)

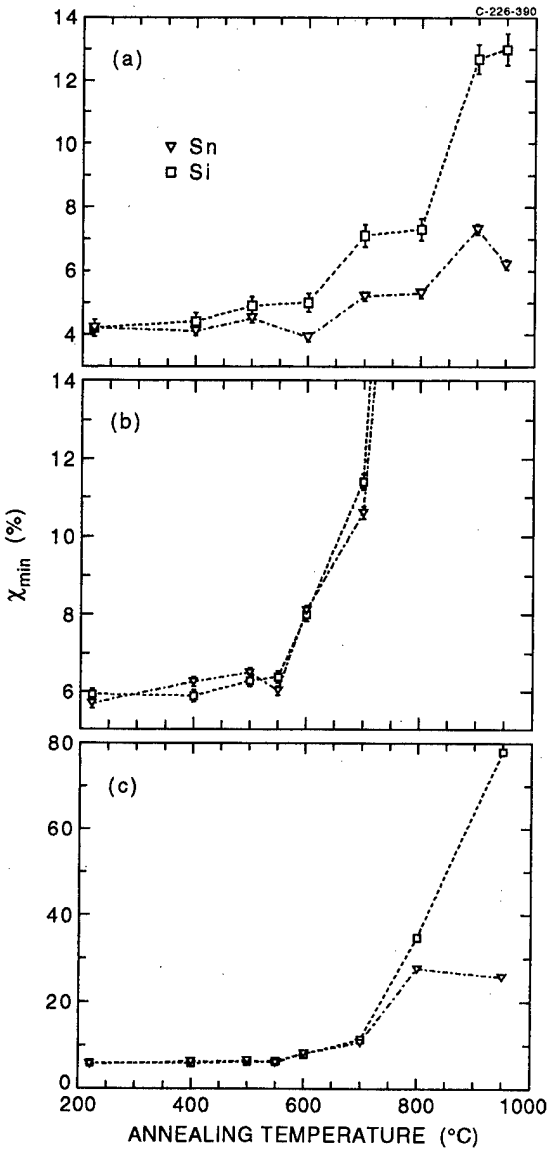


FIG. 2. Minimum-channeling yield, χ_{min}, for (a) strained $Si_{0.975}Sn_{0.025}$, and (b), (c) strained $Si_{0.95}Sn_{0.05}$ as function of annealing temperature. The error bars are only statistical.

and 2(c) for strained $x=5\%$. Note the different scales on the y axis of Figs. 2(b) and 2(c); Fig. 2(b) has the same scale as Fig. 2(a) while Fig. 2(c) shows all the χ_{min} values for strained $x=5\%$. For strained $x=2.5\%$, both χ_{min}(Si) and χ_{min}(Sn) change significantly at temperatures between 600 and 700 °C, and drastically at temperatures above 900 °C. In the case of strained $x=5\%$, the χ_{min}(Si) and χ_{min}(Sn) values increase significantly at temperatures in the range from 550 to 600 °C. After annealing at 950 °C the strained $x=5\%$ has very bad channeling properties as seen in Fig. 2(c).

Only values of the angular width, $\Psi_{1/2}$, obtained from the Si and the Sn part of the channeling spectra of the same sample are compared because of the high sensitivity of the angular width to the position of the random level. For strained $x=2.5\%$, a comparison of $\Psi_{1/2}$ for the Si and Sn angular-yield curves showed no significant deviation between these values at any temperature, whereas for the

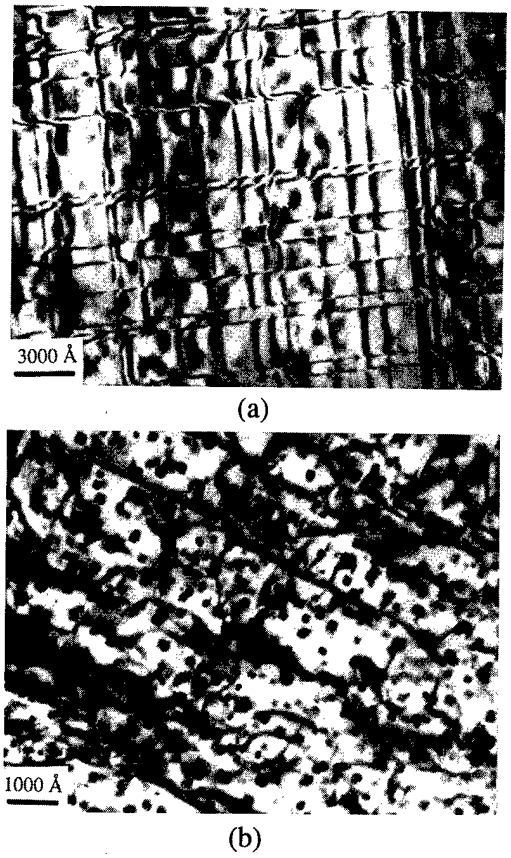


FIG. 3. Plan-view micrographs of strained $Si_{0.95}Sn_{0.05}$ annealed at (a) 600 °C and (b) 950 °C. The micrographs correspond to an area on the TEM specimen which have a thickness larger than the thickness of the $Si_{1-x}Sn_x$ layer and therefore can interfacial misfit dislocations be seen.

strained $x=5\%$, $\Psi_{1/2}$(Si) and $\Psi_{1/2}$(Sn) differs significantly at temperatures above 700 °C; above this temperature the Sn angular-yield curve narrows considerably (see Fig. 1).

B. Transmission electron microscopy

Plan-view and cross-section samples were investigated for all as-grown and annealed samples. In cross-section micrographs of all the as-grown $Si_{1-x}Sn_x$ layers, a columnar structure similar to that reported in Ref. 12 is seen, but no precipitates, defects or dislocations could be observed.

The change in the density of interfacial misfit dislocations and dislocations in the bulk for the strained 5% alloy for two different anneal temperatures are shown in Fig. 3. The (220) bright field, plan-view micrographs corresponds to annealing temperatures of 600 and 950 °C. At 600 °C [Fig. 3(a)] a regular, orthogonal misfit dislocation network appears, whereas for lower annealing temperatures (500 and 550 °C) only very low concentrations of misfit dislocation were found (not shown). The misfit dislocation networks seen at 600 °C are similar to the network generated during relaxation of strained $Si_{1-x}Ge_x$; the network consists of 60° misfit dislocations oriented in the ⟨110⟩ directions, and they are located at the interface between the Si substrate and the

TABLE I. Presence of misfit dislocations and precipitates in strained $Si_{1-x}Sn_x$ layers annealed at various temperatures for 1 h. The term "small" is used when the interfacial misfit dislocation density was too small for an estimate of the degree of relaxation, R, from TEM micrographs, and "perturbed network" indicates that a regular, orthogonal interfacial misfit dislocation network was changed by dislocations and loops in the bulk of the $Si_{1-x}Sn_x$ layer.

Material	Annealing temperature (°C)	Concentration of misfit dislocations	Average size of precipitates (Å)	Density of precipitates (10^{15} cm^{-3})	Fraction of Sn in precipitates (%)
$x=2.5\%$	400	none	...	...	...
	500	none	...	...	...
	600	none	...	...	...
	700	small	30	3.8	0.11
	800	small	55	2.6	0.8
	900	perturbed network	25[a]	0.1	0.005
			80	2.6	2.8
	950	perturbed network	115	1.5	4.3
$x=5\%$	400	none	...	...	...
	500	small	25	...	
	550	regular network, R= 3%	30	...	...
	600	regular network, R=12%	30	5	0.15
	700	regular network, R=16%	40	20	0.8
	800	perturbed network	95	11	8.6
	950	perturbed network	55[a]	12	1.8
			190	5	32

[a]Precipitates situated at the surface.

strained top layer. After annealing at 700 °C the misfit dislocations network was similar to the one found at 600 °C; however, at temperatures above 700 °C the misfit dislocation network becomes less regular, i.e., the network of orthogonal, straight dislocations is to some extend substituted by curved dislocations which are situated in the bulk of the $Si_{1-x}Sn_x$ layer. Figure 3(b) shows a plan-view micrographs of strained $x=5\%$ annealed at 950 °C and here a large density of curved dislocations and loops are seen, however, the density of interfacial misfit dislocations is reduced compared to the density at 700 °C. An estimate of the degree of relaxation of the layers with a regular dislocation network is given in Table I. Thus, at low temperatures ($T=700$ °C) the strained $x=5\%$ layers relaxes mainly by generation of interfacial misfit dislocations, whereas for higher temperatures dislocations in the bulk of the $Si_{1-x}Sn_x$ layer are dominating. The precipitation process further assists the strain relaxation at high temperatures as discussed below.

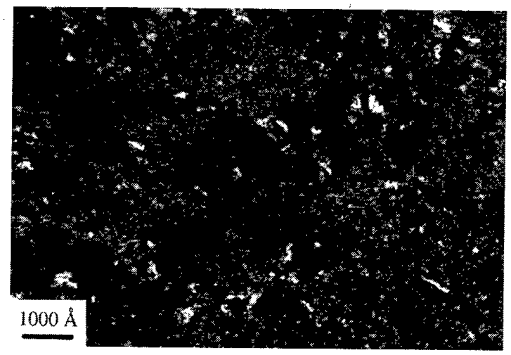
In the strained $x=2.5\%$ alloy, the dislocation network (not shown) never becomes so regular as in the strained $x=5\%$ samples; strained $x=2.5\%$ annealed at 700 and 800 °C have a low density of interfacial misfit dislocations, and for temperatures above 900 °C, the network is perturbed by the dislocations and precipitates in the bulk.

Focusing only on the bulk of the $Si_{1-x}Sn_x$ layers, Fig. 4 shows weak-beam and off-Bragg condition (OBC) micrographs for the strained $x=5\%$ alloy. For strained $x=5\%$ shown in Fig. 4, a comparison of the weak-beam micrographs after annealing at 700 and 800 °C reveals that the concentration of bulk dislocations change very fast in this temperature interval; at 700 °C, where there is a large density of interfacial-misfit dislocations as mentioned above, the concentration of bulk dislocations is small, whereas at 800 °C the concentration of this type of dislocations is high. We note that in Fig. 4(c) the thickness of the investigated layer was about 1200 Å and therefore thinner than the entire $Si_{1-x}Sn$ layer (1700 Å), thus none of the dislocations in Fig. 4(c) are situated at the interface. The thickness of the layers was estimated using the contamination-spot separation method.[18]

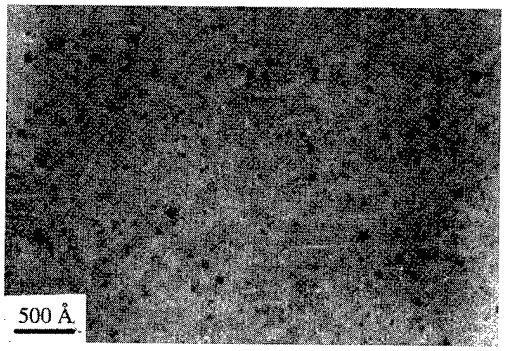
Figure 5 shows weak-beam and OBC micrographs of strained $x=2.5\%$ taken at 800 and 900 °C, which is the temperature range where the density of bulk dislocations change rapidly. At 800 °C the number of dislocations is small but after annealing at 900 °C the density of loops is high [Fig. 5(c)]. Apparently the loops in Fig. 5(c) are decorated with precipitates. The dislocations seen as long dark lines in Fig. 5(d) are misfit dislocations situated at the interface.

The OBC micrographs in Fig. 4 shows how the size of the precipitates increases with annealing temperature in strained 5%. The threshold for precipitation in the strained $x=5\%$ alloy is 500 °C, however, the precipitates are small (average sidelength <40 Å) for temperatures up to 700 °C. Table I gives the average size of the precipitates, the density of the precipitates and an estimate for the fraction of the total number of Sn atoms contained in precipitates. The sidelength of the precipitates is defined as the average side-length of the square-like precipitates found in plan-view TEM micrographs.

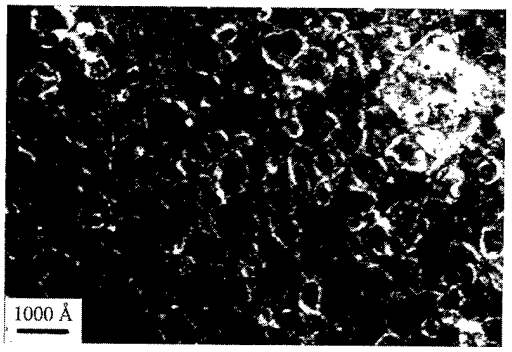
The size of the precipitates in the strained $x=2.5\%$ alloy do also increase with temperature (see Fig. 5 and Table I). The threshold temperature for precipitation in $Si_{1-x}Sn_x$, x



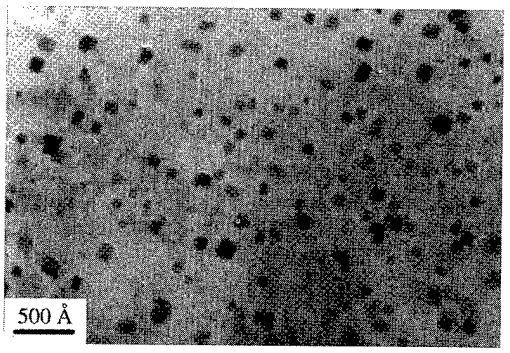
(a)



(b)

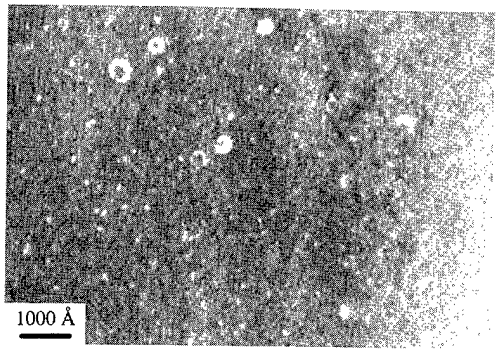


(c)

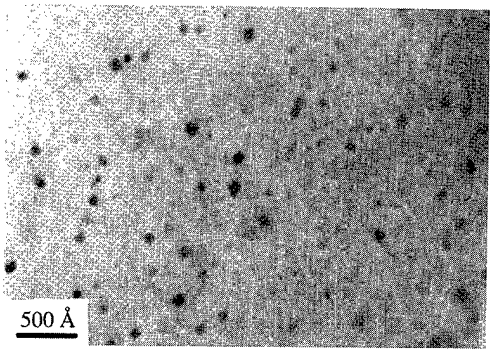


(d)

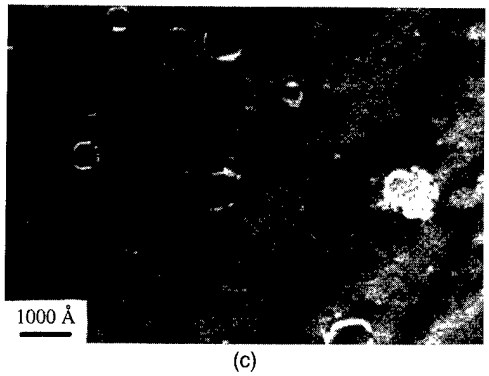
FIG. 4. Plan-view micrographs of strained $Si_{0.95}Sn_{0.05}$ annealed at 700 and 800 °C. (a) and (c) are weak beam (220) micrographs taken after annealing at 700 and 800 °C, respectively; (b) and (d) are micrographs taken away from any Bragg conditions after annealing at 700 and 800 °C, respectively.



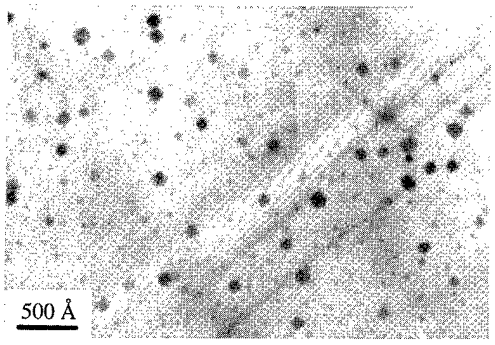
(a)



(b)



(c)



(d)

FIG. 5. Plan-view micrographs of strained $Si_{0.975}Sn_{0.025}$ annealed at 800 and 900 °C. (a) and (c) are weak beam (220) micrographs taken after annealing at 800 and 900 °C, respectively; (b) and (d) are micrographs taken away from any Bragg conditions after annealing at 800 and 900 °C, respectively.

=2.5% is found to be 700 °C. Investigations of the diffraction patterns showed that at high temperatures where the size of the precipitates is sufficient to give visible diffraction spots, spots from both α- and β-Sn are present. For bulk Sn, the phase transition from the diamond-like α-Sn structure to the metallic, tetragonal β-Sn structure takes place at 13.2 °C;[19] the presence of both phases is probably related to interaction with the surrounding matrix and pressure effects.

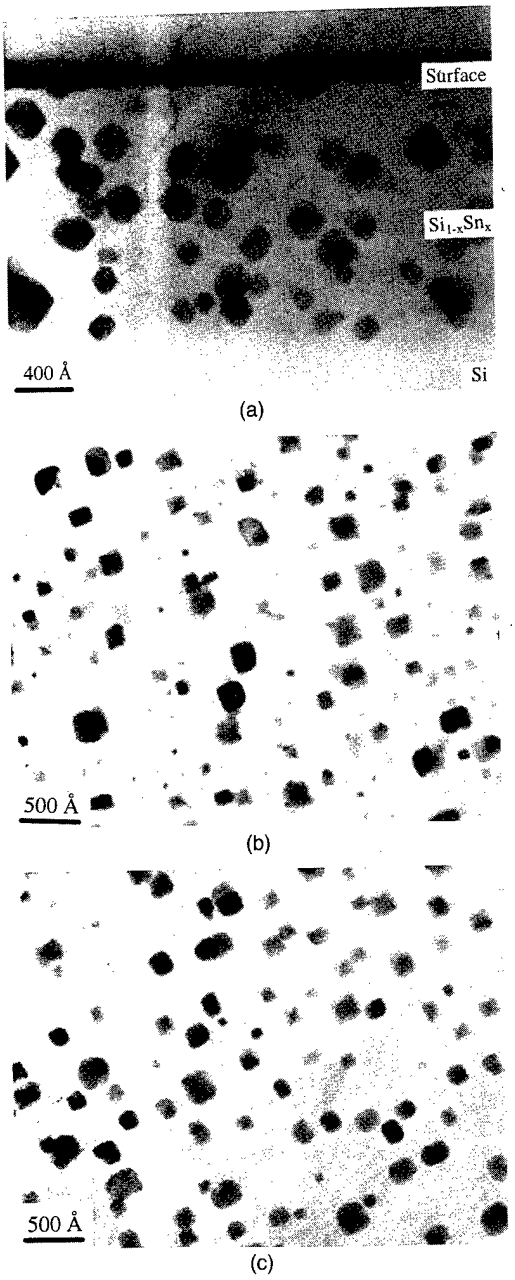


FIG. 6. Strained $Si_{0.95}Sn_{0.05}$ annealed at 950 °C; (a) cross-section micrograph, (b) plan-view micrograph for as-annealed, and (c) plan-view micrograph after surface ion milling which removed a fraction of the original $Si_{1-x}Sn_x$ layer.

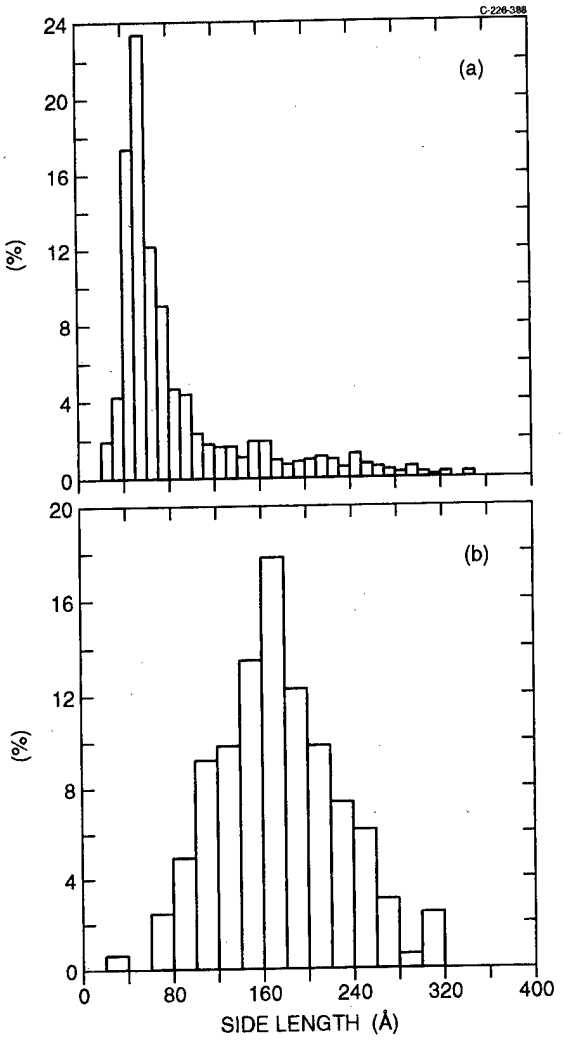


FIG. 7. Size distribution function of the precipitates in $Si_{0.95}Sn_{0.05}$ annealed at 950 °C (a) before and (b) after surface ion milling.

In Fig. 6(a) a cross-section micrograph of strained $x=5\%$ annealed at 950 °C is shown. Several features can be seen in this micrograph: (a) the precipitates in the bulk are obviously bounded by planes of the same orientation, (b) there is a reduced number of precipitates in the near-surface layers (a few hundreds of Å thick), and (c) there are precipitates, smaller than in the bulk, situated on the surface. The shape of the precipitates seen in cross-section and plan-view micrographs makes it likely that the precipitates have an octahedral-like shape; the precipitates are bounded by (111) planes of the $Si_{1-x}Sn_x$ matrix. The reduced concentration of precipitates seen in the surface-near region and the precipitates on the surface is due to segregation of Sn during annealing. In Figs. 6(b) and 6(c) plan-view micrographs of strained $x=5\%$ annealed at 950 °C are shown before and after removal of the precipitates on the surface. This was done by ion-milling the surface for a short time. A comparison of the two figures reveals that the dense cloud of small precipitates seen in Fig. 6(b) is actually situated on the surface of the annealed layers; this can also been seen in the cross-section micrograph in Fig. 6(a).

The size distributions of the precipitates in Figs. 6(b) and 6(c) are shown in Figs. 7(a) and 7(b), respectively. In Fig. 7(a) a large peak corresponding to a precipitate size of about 50–60 Å can be seen and on the right side of this peak a decreasing probability for larger and larger precipitates is found. Figure 7(b) shows the size distribution when the small precipitates on the surface have been removed. The peak in this figure corresponds to the typical size of the precipitates situated in the bulk of the $Si_{1-x}Sn_x$ layer, and therefore Fig. 7(a) is a superposition of two different distribution functions each with a separate peak. The division into two typical sizes

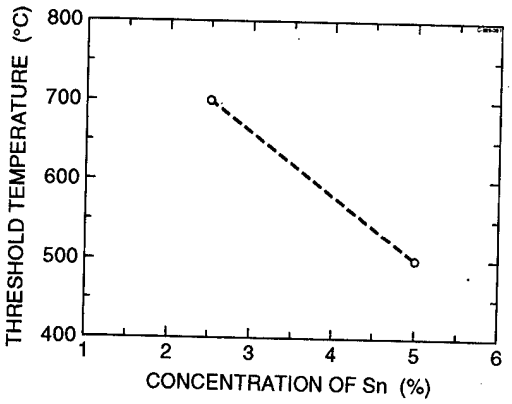


FIG. 8. Threshold temperature for both interfacial misfit dislocations and precipitates for the strained $Si_{1-x}Sn_x$ layers.

of the precipitates was also found in strained $x=2.5\%$ annealed at 900 °C [see Fig. 5(d)].

IV. DISCUSSION

As demonstrated in Sec. III the relaxation of strained, epitaxial $Si_{1-x}Sn_x$ alloys is dominated by several different mechanisms: Generation of (1) interfacial misfit dislocations, (2) line dislocations in the bulk, (3) dislocation loops, and (4) precipitates. Interfacial misfit dislocations and precipitates are the first to appear as a function of temperature, and the threshold temperatures for these relaxations processes are shown in Fig. 8. Line dislocations in the bulk and dislocation loops were only found at temperatures much higher than the threshold temperature for generation of interfacial misfit dislocations; for both alloy concentrations temperatures higher than 800 °C were needed. The density of interfacial misfit dislocations was found to increase very fast in the temperature range 500–600 °C for $x=5\%$ and in the range 800–900 °C for $x=2.5\%$. The minimum yield, $\chi_{\min}$, has been found to increase at the same temperatures where the relaxation by misfit dislocations goes very fast; this indicates that the crystalline quality in the bulk of the strained layers also changes at these temperatures.

The angular-yield curves give information about the structural quality of the as-grown $Si_{1-x}Sn_x$ layers. The angular width $\Psi_{1/2}$ is proportional to $\sqrt{Z_{\text{eff}}}$ (Z_{eff} is the average atomic charge)[17] and, therefore, it should be larger for $Si_{1-x}Sn_x$ than for Si; a theoretical estimate of $\chi_{\min}$ for $Si_{1-x}Sn_x$ and Si shows that these values should be almost similar. None of these expectations corresponds to the experimental data; the angular curves of as-grown $Si_{1-x}Sn_x$ are narrower than those of Si, and the $\chi_{\min}$ value is considerably larger for $Si_{1-x}Sn_x$ than for Si. The increase in $\chi_{\min}$ cannot be due to precipitates because the angular curves for Si and Sn are identical for the as-grown alloys, demonstrating that is there is no preferential scattering on Sn atoms. Threading dislocations (originating at the interfacial misfit dislocations) change the $\chi_{\min}$ value for both the Si and Sn part of the spectra simultaneously, but since the density for the threading dislocations in the as-grown alloys is small ($\leqslant 10^6$ cm^{-2} estimated from the dislocation free TEM micrographs), these dislocations have a negligible effect[20] on $\chi_{\min}$. Monakhov and Nylandsted Larsen[20] found in $Si_{1-x}Ge_x$ a similar increase in $\chi_{\min}$ compared to a simple interpolation between $\chi_{\min}$ for the constituents of the alloy. They explained this effect as being due to a mosaic structure and a static displacement of the atoms. In the relaxed $Si_{1-x}Ge_x$ alloys contributions from both effects imply that $\chi_{\min}$ should increase whereas the influence of the mosaic spread and the static displacement counteracts each other so $\Psi_{1/2}$ should be left unchanged. Following this analysis the channeling in strained $Si_{1-x}Sn_x$ is dominated by static displacement of Si and Sn atoms. Whether this is correlated to the columnar structure seen in all as-grown $Si_{1-x}Sn_x$ layers is not known; however, computer simulations of the channeling spectra are necessary to understand the channeling spectra in better detail.

In the annealed layers the angular-yield profiles change due to precipitates and dislocations situated in the bulk of the $Si_{1-x}Sn_x$ layers. The dislocations in the bulk are misfit dislocations related to precipitates, dislocation loops generated from an excess concentration of point defects and threading dislocations. For temperatures below 800 °C for strained $x=5\%$ and below 900 °C for strained $x=2.5\%$, no dislocations were found in the bulk of the alloy, and therefore, it is primarily the precipitates that change the angular-yield curves at these temperatures. If we consider the change in $\chi_{\min}$ to be due solely to precipitates below these temperatures, the fraction of Sn contained in precipitates can be estimated using the following first order approximation of the substitutional fraction[17] S:

$$S=\frac{1-\chi_{\min}^{\text{annealed}}}{1-\chi_{\min}^{\text{as-grown}}}. \tag{1}$$

The nonsubstitutional fraction, which is assumed to be the fraction in precipitates, can be calculated as 1-S. Both dislocations and precipitates will cause an increase in the backscattering yield, and therefore this formula cannot be used to calculate the fraction of Sn in precipitates when both dislocations and precipitates are present in the bulk of the annealed layers. Taking the difference between the S values calculated for Sn and Si to be due to precipitates, the fraction of Sn in precipitates for strained $x=2.5\%$ is 2% after annealing at both 700 and 800 °C. For strained 5%, the fraction in precipitates is 0.6%, 2% and 1% for 550, 600, and 700 °C, respectively. The reduction at 700 °C is due to a significantly higher non-substitutional fraction of Si, indicating that the bulk dislocations start to develop at this temperature. When estimating the fraction of Sn atoms in precipitates from backscattering spectra, it is important to remember that the yield is sensitive to the alignment of the precipitates relative to the matrix.[21] X-ray and TEM investigations of the phase and orientations of the crystalline precipitates revealed that most of the incoherent precipitates look as amorphous clusters for the ⟨001⟩ channeling direction, so no channeling effect in the precipitates is expected.

Using the density and average size of the precipitates given in Table I the fraction of Sn in precipitates can be

estimated from the TEM results. For strained $x=2.5\%$ annealed at 700, 800, 900, and 950 °C, the volume fraction of precipitates are approximately 0.1%, 0.8%, 2.8%, and 4.3%, respectively; these values were found taking the octahedral shape of the precipitates into account. When estimating the fraction of the Sn atoms in the precipitates, it is important to consider the actual density of atoms in the precipitates; nevertheless, the difference in density of $Si_{1-x}Sn_x$ and α-Sn or β-Sn as well as the lack of knowledge about the fraction in each of these phases implies only a relatively small uncertainty on the Sn fraction in the precipitates. The RBS investigations give that in strained $x=2.5\%$ annealed at 700, 800, 900, and 950°C about 0.6%, 1.2%, 2.5%, and 7% of the Sn are in precipitates, respectively. Considering the uncertainty on both RBS and TEM results the correspondence between the fraction of Sn in precipitates found by the two methods is good.

Based on the size distribution of precipitates in strained $x=5\%$ annealed at 950 °C from Fig. 7(b), the volume fraction of the Sn precipitates was found to be 1.5%, whereas a calculation using the average size and density specified in Table I gives 1.6%. A volume fraction of 1.5% corresponds to about 30% of the Sn in precipitates (neglecting a different atomic density in the precipitates compared to the matrix). The small precipitates situated on the surface corresponds to approximately 6% of the total volume of precipitates. Thus, even after annealing at very high temperatures less than half of the Sn is contained in the precipitates.

The transport of Sn from the matrix to the precipitates is expected to have only a minor influence on the strain relaxation of the matrix. This is because of the difference in atomic density in the precipitates and the matrix; using Vegards law and the elastic constants of Si and Sn the density of strained $x=5\%$ was found to be 4.95×10^{22} cm^{-3}, whereas the density of α-Sn and β-Sn is 2.93×10^{22} and 3.70×10^{22} cm^{-3}, respectively. Therefore, a fixed number of Sn atoms takes up a larger volume in these phases than in the matrix, probably vacancies relax the strain related to this volume difference. Only the surface precipitates induce a depletion of Sn from the matrix, but due to the small fraction of Sn in these precipitates this effect has only a very small influence on the strain relaxation.

The question of how the precipitates are nucleated and how they grow with time cannot be answered from the present study; however, the diffusion constants for Sn in $Si_{1-x}Sn_x$ can be estimated from the size of the precipitates. Assuming that the precipitation is diffusion limited, the diffusion coefficient, D, is approximately given by[22]

$$D=\frac{C_p}{C_0}\frac{r^2}{2t}, \tag{2}$$

where C_p is the Sn concentration in a precipitate, C_0 is the initial alloy concentration of Sn, r_p is the average size of the precipitates and t is the mean growth time. This expression, which includes only the average size and not the number density of the precipitates, is based on an analysis given in Ref. 23 and it has been derived assuming that less than ~20% of the Sn atoms are contained in precipitates so the interaction between precipitates can be neglected. Using a mean growth time equal to half the annealing time (as in Ref. 22) and an average radius of 57.5 Å for $Si_{0.975}Sn_{0.025}$ annealed at 950 °C, a diffusion coefficient of 4×10^{-15} cm^2/s is estimated. In obtaining this result we assumed that the precipitates consist of only Sn, which is in agreement with the observation that the distance between the atomic planes as determined by diffraction fits the tabulated values for α-Sn and β-Sn. In tracer diffusion studies of Sn in Si a preexponential factor of 5×10^3 and an activation energy of 4.91 eV has been found,[24] which implies a diffusion coefficient $D=3\times10^{-17}$ cm^2/s at 950 °C. Therefore at 950 °C the diffusion of Sn in $Si_{1-x}Sn_x$ is two orders of magnitude faster than dilute Sn diffusion in Si, and this conclusion holds also for lower temperatures. Studies of the diffusion of Sb in strained $Si_{1-x}Ge_x$ of about the same strain as in the present case have also demonstrated an increase in the diffusivity which could be correlated both to the strain and to a chemical effect of the Ge.[25] The diffusion of Sn in Si (and in $Si_{1-x}Ge_x$ alloys) is expected to be similar to that of Sb in Si;[26] hence an increase in the diffusivity of Sn in a strained $Si_{1-x}Sn_x$ alloy relative to that in Si is in agreement with expectations. However, the observed increase of two orders of magnitude is significantly larger than that observed in strained $Si_{1-x}Ge_x$.

The change from only interfacial misfit dislocations to a situation with a large density of bulk dislocations takes place at the same temperatures as where the precipitates become large. One suggestion is, that the precipitates are blocking the movement and the multiplication processes of the interfacial misfit dislocations, and hence the curved misfit dislocation network in the bulk develops. Similar blocking of misfit dislocations by precipitates has been found in annealing of strained $Si_{1-x}C_x$.[27] The observation that the interfacial misfit dislocation density never implies very high degree of relaxation and that less than half of the Sn is included in precipitates reveals that the strained $Si_{1-x}Sn_x$ layers are far from being fully relaxed at any of the investigated annealing temperatures.

V. CONCLUSIONS

The thermal stability of strained $Si_{1-x}Sn_x$ has been studied for a fixed annealing time of 1 h. The observed relaxation channels are generation of precipitates, dislocation loops and misfit dislocations; the misfit dislocations are either interfacial or situated in the bulk of the $Si_{1-x}Sn_x$ layer. The threshold temperature was found to be the same for the formation of interfacial misfit dislocations and precipitates: 700 °C for a $Si_{0.975}Sn_{0.025}$ alloy and 500 °C for a $Si_{0.95}Sn_{0.05}$; higher temperatures were necessary for the generation of dislocation loops and misfit dislocations in the bulk. The sizes of the precipitates increase fast, whereas their densities are found to decrease with temperature. Estimates of the fraction of Sn in precipitates reveal that less than half of the Sn is in precipitates even after annealing at high temperatures. Neither this depletion of Sn nor the interfacial misfit dislocations are enough to completely relax the strained $Si_{1-x}Sn_x$ layers.

ACKNOWLEDGMENTS

This work was supported by the Materials Development Program MUPII through the "Center for Nanostructures" (CNAST). Pia Bomholt is acknowledged for the preparation of the TEM samples and J. Lundsgaard Hansen for the MBE growth of the layers investigated in this study. The authors are grateful to Dr. J. Duyster, University of Bochum, Germany, for his assistance on the computer image analysis.

[1]K. Eberl, S. S. Iyer, S. Zollner, J. C. Tsang, and F. K. LeGoues, Appl. Phys. Lett. **60**, 3033 (1992).
[2]G. He, M. D. Savellano, and H. A. Atwater, Appl. Phys. Lett. **65**, 1159 (1994).
[3]P. Mei, M. T. Schmidt, E. S. Yang, and B. J. Wilkens, J. Appl. Phys. **69**, 8417 (1991).
[4]S. Y. Shiryaev, J. Lundsgaard Hansen, P. Kringhøj, and A. Nylandsted Larsen, Appl. Phys. Lett. **67**, 2287 (1995).
[5]W. Wegscheider, J. Olajos, U. Menczigar, W. Dondl, and G. Abstreiter, J. Cryst. Growth **123**, 75 (1992).
[6]G. He and H. A. Atwater, Appl. Phys. Lett. **68**, 664 (1996).
[7]F. J. Guarin, S. S. Iyer, A. R. Powell, and B. A. Ek, Appl. Phys. Lett. **68**, 3608 (1996).
[8]R. A. Soref and C. H. Perry, J. Appl. Phys. **69**, 539 (1991).
[9]A. T. Khan, P. R. Berger, F. J. Guarin, and S. S. Iyer, Appl. Phys. Lett. **68**, 3105 (1996).
[10]T. B. Massalski, H. O. Kamoto, P. R. Subramanian, and L. Kacprzak, *Binary Alloy Phase Diagrams* (ASM International, Metals Park, OH, 1990).
[11]M. R. Sardela, Jr., R. Turan, M. Willander, G. V. Hansson, and L. Hultman, J. Appl. Phys. **77**, 1411 (1995).
[12]M. F. Fyhn, S. Y. Shiryaev, J. Lundsgaard Hansen, and A. Nylandsted Larsen, Appl. Phys. Lett. **69**, 394 (1996).
[13]M. Fanciulli, H. C. Vestergaard, G. Weyer, M. Fyhn, S. Yu. Shiryaev, and A. Nylandsted Larsen, in *Proceedings of the 23rd International Conference on the Physics of Semiconductors,* edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 1059.
[14]J. Lundsgaard Hansen, S. Y. Shiryaev, and E. V. Thomsen, J. Cryst. Growth **157**, 317 (1995).
[15]In previous articles of Refs. 4 and 12, the growth temperature was specified to be 225 °C during the growth of the $Si_{1-x}Sn_x$ layers; this temperature was obtained from an extrapolation of pyrometer data. However, careful calibration done by J. Lundsgaard Hansen using a thermocouple mounted on a Si substrate revealed that the growth temperature is offset by 70 °C, and, therefore, the growth temperature was 295 °C during production of the layers used in this study.
[16]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[17]L. C. Feldman, J. W. Mayer, and S. Thomas Picraux, *Materials Analysis by Ion Channeling* (Academic, New York, 1982).
[18]The thickness of a plan-view sample can be determined using the distance between carbon spots produced on both surfaces of the sample by a focused e-beam when the sample is tilted a given angle; see, for example, G. W. Lorimer, G. Cliff, and J. N. Clark, in *Developments in Electron Microscopy and Analysis,* edited by J. A. Venables (Academic, London, 1976), p. 153.
[19]G. A. Busch and R. Kern, in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1961), Vol. 11, p. 1.
[20]E. V. Monakhov and A. Nylandsted Larsen, Nucl. Instrum. Methods Phys. Res. B **117**, 71 (1996).
[21]K. K. Bourdelle, V. A. Khodyrev, A. Johansen, E. Johnsen, and L. Sarholt-Kristensen, Phys. Rev. B **50**, 82 (1994).
[22]S. J. Pennycook, J. Narayan, and O. W. Holland, J. Appl. Phys. **55**, 837 (1984).
[23]P. Hasen, *Physical Metallurgy* (Cambridge University, London, 1978).
[24]P. Kringhøj and R. G. Elliman, Appl. Phys. Lett. **65**, 324 (1994).
[25]P. Kringhøj, A. Nylandsted Larsen, and S. Y. Shiryaev, Phys. Rev. Lett. **76**, 3372 (1996).
[26]U. M. Gösele and T. Y. Tan, in *Materials Science and Technology, A Comprehensive Treatment, Volume 4: Electronic Structure and Properties of Semiconductors,* edited by W. Schröter (VCH, Weinheim, 1991), p. 197.
[27]A. R. Powell, F. K. LeGoues, and S. S. Iyer, Appl. Phys. Lett. **64**, 324 (1994).

Layer by layer growth mode stabilization and step-edge smoothing of cation-stabilized $In_{1-x}Ga_xAs$ strained layers grown on InP

Y. Robach,[a),b)] A. Soléré,[b),c)] M. Gendry,[c)] and L. Porte[b),c)]
Ecole Centrale de Lyon, BP 163, 69131 Ecully Cedex, France

(Received 21 January 1998; accepted 29 May 1998)

Scanning tunneling microscopy has been used to study the surface morphology of strained $In_{1-x}Ga_xAs$ layers (either 2% compressively or 2% tensilely strained for $x=0.18$ or $x=0.75$, respectively) grown by molecular beam epitaxy on (001) InP substrate under cation-stabilized conditions. Under such growing conditions a smooth two-dimensional (2D) surface morphology is well preserved until the onset of plastic relaxation. This behavior is completely unlike the 2D/3D growth mode transition observed under As-stabilized conditions of growth. Along with the 2D growth mode stabilization, a step-edge smoothing is also observed. These results are assigned to a high value of step formation free energy on (4×2) reconstructed surfaces. The resultant increase in surface tension delays the onset of coherent 3D island formation beyond the onset of plastic relaxation and stabilizes 2D growth. © *1998 American Vacuum Society.*
[S0734-211X(98)13504-7]

I. INTRODUCTION

Recently significant efforts have been devoted to controlling the epitaxial growth mode of strained semiconductor layers. High performances of strained layer based electronic and photonic devices are intimately related to the realization of morphologically smooth and compositionally sharp interfaces. It was demonstrated for instance that the use of surfactants can drastically modify the epitaxial growth mode of strained systems such as Ge/Si (Ref. 1) or (In, Ga)As/GaAs,[2] changing it from three-dimensional (3D) island formation to layer by layer growth. Concerning III-V semiconductor heteroepitaxy, the cation to anion flux ratio was also recognized as another parameter influencing the growth mode, without resorting to the addition of foreign species. Recent experimental results have highlighted the contrasting growth behavior of epitaxial highly strained systems such as InAs/InGaAs[3,4] or (In, Ga)As/GaAs,[5] under two limiting flux conditions: Under As-rich conditions (As-stabilized surfaces) the growth is characterized by a 2D/3D morphological transition in a dislocation free film (strain relaxation through coherent islanding); on the contrary under reduced arsenic pressure (cation-stabilized surfaces) 3D islanding can be inhibited and a 2D growth preserved up to the onset of plastic relaxation linked to the formation of dislocations.

In this work, a scanning tunneling microscopy (STM) morphological characterization of highly strained $In_{1-x}Ga_xAs$ epitaxial layers grown on InP (001) surfaces under cation-stabilized conditions is presented. The purpose of our study was to examine the intimate relation between the layer by layer growth mode stabilization of the strained film and its morphology. As the strained layer thickness increases, the persistence of a 2D morphology is concomitant with the formation of small isotropic 2D islands and the smoothing of [1–10] step edges. These morphological features and their relation to the 2D growth mode stabilization are discussed.

II. EXPERIMENTAL PROCEDURE

Epitaxial layers were grown in a RIBER 2300 molecular beam epitaxy system equipped with *in situ* reflection high-energy electron diffraction (RHEED) and connected under ultrahigh vacuum to a STM analysis chamber. A detailed description of this system has been given elsewhere.[6] After thermal desorption of the native oxide from an exactly (001) oriented n^+-type InP substrate, a ~500 nm lattice-matched $In_{0.53}Ga_{0.47}As$ buffer layer was first grown at a substrate temperature of 525 °C under As-stabilized conditions with a 2×1 surface reconstruction, in two steps: for the first 400 nm, a growth rate of 1 μm/h and corresponding As/element III beam equivalent pressure ratio of 20 were used. For the last 100 nm, these growth parameters were changed to 0.4 μm/h and 40, respectively. The layer was Si doped at a donor level concentration of $\sim 10^{18}$ cm^{-3} except for the last 50 Å. The buffer layer surface was allowed to recover under As flux for a few minutes while the cation effusion cell temperatures were adjusted for subsequent growth of the strained layer. The As shutter was then closed and the RHEED pattern changed immediately from the As-stabilized 2×1 to the cation-stabilized 4×2 reconstruction, while the As/element III beam equivalent pressure ratio decreased to a value of ~0.5 ($P_{As} \sim 5.10^{-8}$ Torr). Highly mismatched strained $In_{1-x}Ga_xAs$ layers (either 2% compressively strained or 2% tensilely strained, for $x=0.18$ and $x=0.75$, respectively) were grown at 525 °C and a slow growth rate of 0.25 μm/h under cation-stabilized conditions without direct As flux (i.e., the As shutter remained closed). The As supply comes from the residual As species present in the surrounding of the growing film. The 4×2 reconstruction characteristic of a cation-stabilized surface was still visible throughout the

[a)]Author to whom correspondence should be addressed; electronic mail: Robach@ec-lyon.fr
[b)]Département de Physicochimie des Matériaux.
[c)]LEAME.

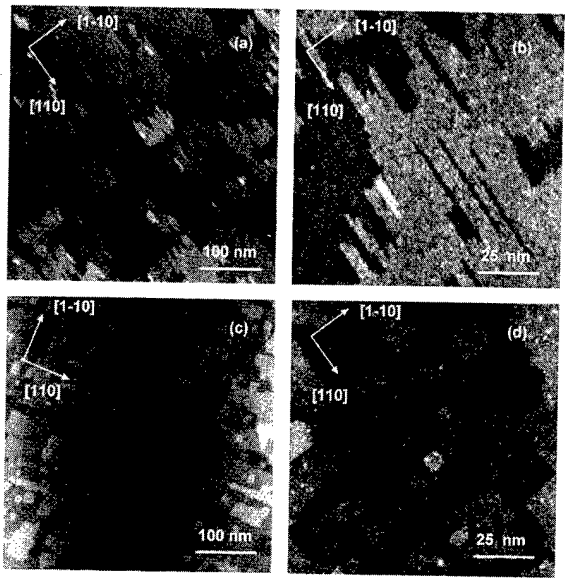


FIG. 1. Surface morphologies of 2% compressively strained $In_{0.82}Ga_{0.18}As$ layers epitaxially grown on (001)InP substrate under cation-stabilized conditions: (a) 480 nm×480 nm and (b) 120 nm×120 nm STM images of a 3 ML strained layer. (c) 480 nm×480 nm and (d) 120 nm×120 nm STM images of a 10 ML strained layer. Step heights of terrace and island edges correspond to 1 ML (~3 Å).

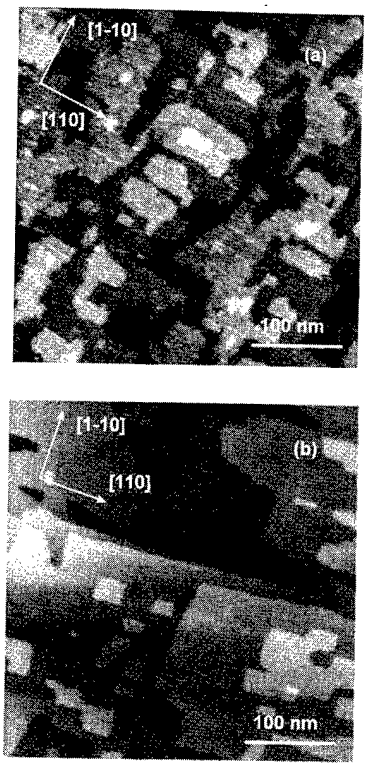


FIG. 2. Surface morphologies of 2% tensilely strained $In_{0.25}Ga_{0.75}As$ layers epitaxially grown on (001)InP substrate under cation-stabilized conditions: 480 nm×480 nm STM images of (a) a 3 ML strained layer and (b) a 17 ML strained layer.

strained layer growth. Immediately after growth, the samples were rapidly cooled to ~450 °C, transferred to the STM chamber and allowed to cool to room temperature. Atomic compositions of strained layers have been checked by high resolution x-ray diffraction (HRXRD). STM images were typically acquired at a tunneling current of 0.2 nA, a sample tunneling voltage of ~2 V (either at positive or negative bias), using electrochemically etched $Pt_{0.8}Ir_{0.2}$ tips. Characteristic images were acquired at different places on the sample to get a representative picture of the layers.

III. RESULTS

The 2D/3D growth mode transition of strained $In_{1-x}Ga_xAs$ layers epitaxially grown on InP(001) under As-stabilized conditions was previously studied in our laboratory, with the STM,[7,8] giving the following main results: for compressively strained $In_{0.82}Ga_{0.18}As$ layers, a layer by layer growth is observed until 4 monolayers (ML) were deposited. Large anisotropic 2D platelets, elongated in the [1–10] direction, develop on complete terraces and act as precursors to 3D islands formation which appears at the onset of the 2D/3D transition, when the fifth monolayer was deposited. For tensilely strained $In_{0.25}Ga_{0.75}As$ layers the behavior is slightly different; a more gradual roughening is observed, until at a deposited thickness of ~13 ML well defined 3D structures developed; the resulting morphology could be described as 3D "zigzag" chains surrounding holes.

Under cation-stabilized conditions, the growth behavior is completely different: a 2D growth mode is preserved until the onset of plastic relaxation as described below. A typical STM image of the surface after deposition of 3 ML of compressively strained $In_{0.82}Ga_{0.18}As$ is shown in Fig. 1(a). The smooth 2D morphology exhibits about four level changes (if levels introduced by the local miscut of the sample are not taken into account) on a large scale of ~4500 Å, with slightly anisotropic islands (average aspect ratio of ~4) elongated in the [110] direction. On the filled-state image shown in Fig. 1(b), at higher magnification, bright lines running along the [110] direction are apparent, with a regular spacing of 17 Å in the [1–10] direction; they correspond to the fourfold periodicity of the 4×2 reconstructed surface observed under reduced As pressure. This reconstruction characteristic of a cation-terminated surface has been well documented especially for GaAs.[9,10] It is described by cation dimers rows along the [110] direction with a regular array of dimer vacancies in the [1–10] direction. The step edges running parallel to the cation dimer rows (As-terminated [110] steps) are straight, whereas the perpendicular step edges (In-terminated [1–10] steps) show many kinks. One characteristic feature of the surface morphology is the frequent occurrence of very narrow terrace holes due to [110] missing rows. They have mostly a width of one or two unit cells of the reconstructed surface. A partial relaxation of strain can occur at these steps and hinder the completion of the last missing row of size one unit cell of the reconstructed surface.

After growth of a thicker 10 ML compressively strained $In_{0.82}Ga_{0.18}As$ layer, a smooth 2D morphology is still preserved [Fig. 1(c)]. This contrasts clearly with the growth morphology observed under As-stable conditions, where 3D islanding has formed since the fifth deposited layer.[7] The

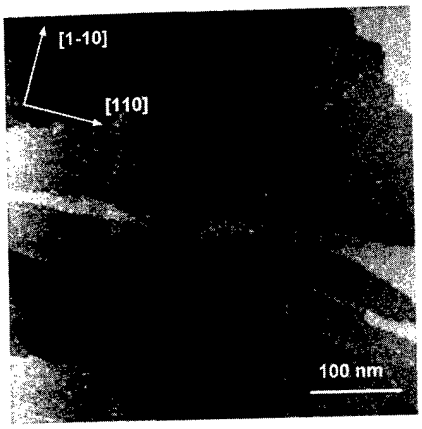


FIG. 3. Surface morphology of a 2% compressively strained $In_{0.82}Ga_{0.18}As$ layer after a 5 min annealing at 525 °C: 480 nm×480 nm STM image.

evolution of the growth front morphology is mainly characterized by a more isotropic growth and a very significant smoothing of the [1–10] step edges [Fig. 1(d)]. Numerous small quasi-square 2D islands of 1 ML height and size typically less than ~200 Å have formed. Bigger islands result from the coalescence of smaller ones. There is practically no overgrowth on top of the existing islands. The numerous 2D islands provide a high density of steps. Incoming adatoms can easily be incorporated at steps and further growth of the layer proceeds by lateral expansion and coalescence of these 2D islands.

The surface morphologies of 2% tensilely strained $In_{0.25}Ga_{0.75}As$ layers show the same characteristic evolution, as the strained layer thickness increases from 3 to 17 ML [Figs. 2(a) and 2(b)]. A smooth 2D morphology is still observed for the thicker deposit. At the same time, the growth becomes more isotropic and the [1–10] step edges smoothen, even more than for the compressive case. In that case too, under As-stabilized conditions, 3D structures would have developed for an equivalent deposited thickness.

We have also checked the stability of the 2D morphology. A 10 ML compressively strained $In_{0.82}Ga_{0.18}As$ layer was annealed just after growth for 5 min at a growth temperature of 525 °C. The surface morphology (Fig. 3) exhibits only three 1 ML high level changes on a large scale of 480 ×480 nm^2. The morphological evolution, by coalescence of 2D platelets, has led to a large decrease of the total step edge length, and thus of the corresponding step edge free energy. From a rough estimation, we measured a step length of ~5600 Å per 1000 $Å^2$ for the as grown sample and ~1700 Å per 1000 $Å^2$ for the annealed one, i.e., a decrease by a factor of 3–4. This evolution of the strained layer from a kinetically restricted morphology towards an equilibrium surface configuration still preserves and ameliorates the 2D morphology.

IV. DISCUSSION

In highly strained systems, the growth usually changes from an initial layer by layer to a coherent 3D island growth (Stranski–Krastanov growth mode). This strain induced coherent islanding is driven by a competition between elastic energy and surface free energy of the system. Strain energy is reduced by elastic deformation of island edges and substrate, while surface energy is increased by formation of steps or new facets. It has been shown that the decrease of elastic energy can outweigh the increase of surface energy for sufficiently large islands.[11–13] This competition is at the origin of the formation of 3D islands on the surface. However, this roughening instability may be delayed by kinetic restrictions associated with a low surface diffusion,[14] which limits the approach to the thermodynamically stable configuration. Kinetic effects are actually manifested in our observations: a high nucleation rate, particularly obvious for the 10 ML compressively strained layer, and an evolution of the as grown morphology during annealing; when reaching the highly reactive cation-stabilized surface, As atoms would be incorporated rapidly in the growing film, leading to an effective short diffusion length of the As adatoms.[3,4] It would prevent the system from reaching its equilibrium state. However, we think energetic considerations are mainly responsible for the preservation of a layer by layer mode when growing the strained layer under cation-stabilized conditions. In fact, as the strained layer is allowed to evolve towards its energetically stable configuration along with annealing, the 2D morphology is still preserved. The morphological evolution is associated with a large decrease of the total step length per unit area, and hence with the corresponding step free energy. These results may be understood if we assume a high step formation energy on (4×2) reconstructed surfaces under As deficient conditions. During its initial stage of development, strain induced roughening requires the formation of steps, thereby costing additional surface energy. With a high step energy, the energetic barrier to form 3D islands is thus increased and it would be necessary to strain relieve a larger volume for the energy gain due to elastic relaxation at step edges to exceed the step energy increase. Thus, the transition to coherent 3D islanding may not be allowed and metastable elastically strained layers are expected to remain planar until the occurrence of plastic relaxation by formation of dislocations. For deposits of equivalent thickness grown under As-stabilized conditions, step formation energy is low,[15] the energetic balance leads to the formation of 3D islands. Such a high step formation energy on (4×2) reconstructed surfaces has been also suggested to explain the step bunching observed on vicinal GaAs and InAs (001) surfaces[16,17] annealed at high temperatures under reduced As pressure. The experimental observation of smooth step edges is also coherent with a high step formation energy. Since the presence of numerous kinks increases the total length of steps, the step free energy is lowered by decreasing the number of kinks. The smoothing effect mainly concerns the steps parallel to the [1–10] direction and is associated with a decrease in the step length in the [110] direction. It expresses an increased [110] step free energy with time, which may be explained as follows: during the growth of the strained layer

under cation-stabilized conditions, the residual As species present in the surrounding of the growing film are consumed, leading to a gradual decrease of the As pressure (the As shutter was closed at the beginning of the strained layer growth). Theoretical calculations by Zhang *et al.*[18] have shown that step free energy is dependent on the As chemical potential, i.e., on As pressure. Consequently, as the thickness of the strained layer increases, i.e., with growing time, step free energy may evolve. The growth evolution towards the formation of isotropic 2D platelets, for the thicker strained layers, tends to indicate an increase of the [110] step free energy up to a value close to the [1−10] step energy.

In sight of our results, step energetics play a large contributory factor to the morphological evolution of strained layers grown under As deficient conditions, particularly as regards stabilization of a 2D growth mode.

V. CONCLUSION

The growth front morphological evolution of 2% compressively and 2% tensilely strained $In_{1-x}Ga_xAs$ layers (with x equal to 0.18 and 0.75, respectively) grown on (001) InP surfaces under cation-stabilized conditions has been investigated using STM. Contrary to As-stabilized growth conditions, a layer by layer growth mode is well preserved as the strained layer thickness increases. At the same time, the growth becomes more isotropic and [1−10] step edges smoothen. These morphological features were assigned to a high step formation free energy on (4×2) reconstructed surfaces, which play a crucial role in the 2D morphology stabilization.

[1]M. Copel, M. C. Reuter, E. Kaxiras, and R. M. Tromp, Phys. Rev. Lett. **63**, 632 (1989).
[2]J. Massies, N. Grandjean, and V. H. Etgens, Appl. Phys. Lett. **61**, 99 (1992).
[3]E. Tournié, O. Brandt, K. H. Ploog, and M. Hohenstein, Appl. Phys. A: Solids Surf. **56**, 91 (1993).
[4]E. Tournié and K. H. Ploog, J. Cryst. Growth **135**, 97 (1994).
[5]C. W. Snyder, B. G. Orr, and H. Munekata, Appl. Phys. Lett. **62**, 46 (1993).
[6]P. Krapf, J. P. Lainé, Y. Robach, and L. Porte, J. Phys. III **5**, 1871 (1995).
[7]L. Porte, P. Krapf, Y. Robach, M. Phaner, M. Gendry, and G. Hollinger, Surf. Sci. **352–354**, 60 (1996).
[8]P. Krapf, Y. Robach, M. Gendry, and L. Porte, Phys. Rev. B **55**, R10229 (1997).
[9]D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L. E. Swartz, Phys. Rev. B **41**, 5701 (1990).
[10]Q. Xue, T. Hashizume, J. M. Zhou, T. Sakata, and T. Sakurai, Appl. Surf. Sci. **87–88**, 364 (1995).
[11]C. Ratsch and A. Zangwill, Surf. Sci. **293**, 123 (1993).
[12]J. Tersoff and F. K. LeGoues, Phys. Rev. Lett. **72**, 3570 (1994).
[13]Y. Chen and J. Washburn, Phys. Rev. Lett. **77**, 4046 (1996).
[14]C. W. Snyder, J. F. Mansfield, and B. G. Orr, Phys. Rev. B **46**, 9551 (1992).
[15]E. J. Heller, Z. Y. Zhang, and M. G. Lagally, Phys. Rev. Lett. **71**, 743 (1993).
[16]S. L. Skala, S. T. Chou, K. Y. Gheng, J. R. Tucker, and J. W. Lyding, Appl. Phys. Lett. **65**, 722 (1994).
[17]H. Yamaguchi, M. R. Fahy, and B. A. Joyce, Appl. Phys. Lett. **69**, 776 (1996).
[18]S. B. Zhang and A. Zunger, Phys. Rev. B **53**, 1343 (1996).

Reduced 980 nm laser facet absorption by band gap shifted extended cavities*

P. G. Piva,[a] R. D. Goldberg, and I. V. Mitchell
Department of Physics, University of Western Ontario, London, Ontario N6A 3K7, Canada

S. Fafard, M. Dion, M. Buchanan, and S. Charbonneau[b]
Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

G. Hillier and C. Miner
Nortel Technology, Ottawa, Ontario K1Y 4H7, Canada

(Received 17 April 1998; accepted 22 April 1998)

Reflectance modulation thermography has been used to determine facet temperatures of InGaAs/GaAs double quantum well (QW) GRINSCH ridge-waveguide lasers possessing band gap shifted extended cavities (BSECs). The incorporation of BSECs produced by mega-electron-volt ion-implantation enhanced QW intermixing, significantly decreased the laser facet temperatures and should result in increased device longevity prior to the onset of catastrophic mirror failure. Low energy implants in Al-free InGaAs/InGaAsP/InGaP laser structures exhibited large effective diffusivities of intermixing enhancing defects from the implant damage regions. This latter material system is particularly well suited for the implementation of BSECs as end of range damage from the implant can be kept spatially isolated from the optical mode regions. © *1998 American Vacuum Society.* [S0734-211X(98)01804-6]

I. INTRODUCTION

Semiconductor quantum well (QW) lasers find use in a variety of applications such as optical memory systems, laser printing, and as pump sources for other gain media. In the telecommunications industry, high power 980 nm QW diode lasers are used to pump erbium-doped amplification stages for long distance optical signal transmission. The longevity of these devices is generally limited by photoabsorption induced degradation at the mirror facets.[1,2] A number of different techniques have been devised to counter this effect[3–5] but these involve the addition of complex processing steps such as growth/regrowth, or the deposition of passivation coatings onto the laser facets. While extended cavities (transparent or otherwise) have been used to improve various aspects of laser diode performance,[6,7] to our knowledge, none have reported their effect on the laser facet temperatures during device operation.

Here we report laser facet temperature measurements performed on InGaAs/GaAs/AlGaAs double QW GRINSCH ridge-waveguide lasers possessing integrated band gap shifted extended cavities (BSECs) produced by high-energy ion implantation enhanced QW intermixing. We show that the addition of BSECs lead to important decreases in the laser facet heating coefficients resulting from the combined effects of reduced photoabsorption and distancing of the reflectors from the device's active region. Additionally, we present preliminary results on the disordering of Al-free InGaAs/InGaAsP/InGaP laser structures. The success of low energy ion implantation in significantly interdiffusing QWs at an order of magnitude greater depth than the end of range damage, suggests that this material system would be particularly well suited for the fabrication of BSECs as the end of range damage resulting from the implantation can be kept distanced from the optical mod region.

II. EXPERIMENT

The InGaAs/GaAs/AlGaAs laser structure used in this experiment was grown by molecular beam epitaxy and is described elsewhere.[8] Given the depth of the QWs from the surface (1.98 μm) high energy implants were required to effectively interdiffuse the QWs given the low effective diffusivities of intermixing enhancing defects in this material system.[9] The InGaAs/InGaAsP/InGaP samples used in the second part of this study consisted of (from the GaAs substrate up) a 500 Å GaAs buffer layer, 1 μm of $In_{0.49}Ga_{0.51}P$, 1000 Å of $In_{0.25}Ga_{0.75}As_{0.49}P_{0.51}$, 85 Å of $In_{0.2}Ga_{0.8}As$, 1000 Å of $In_{0.25}Ga_{0.75}As_{0.49}P_{0.51}$, 2000 Å of $In_{0.49}Ga_{0.51}P$, and a 50 Å GaAs cap. The samples were grown by metalorganic chemical vapor deposition.

All implantations were carried out using the 1.7 MV Tandem accelerator facility located at the University of Western Ontario. Samples were implanted at 7° off normal to minimize the effects of ion channeling. For the InGaAs/GaAs/AlGaAs work, one sample was cleaved from the growth wafer, and masked using a 400-μm-wide Cu wire. This defined the geometry of the unimplanted (and therefore unshifted band gap) region of the sample to comprise the active regions of our devices. The sample was then implanted at 200 °C with 8.56 MeV As^{4+} ions to a dose of 2.5×10^{13} cm^{-2}. The expected range of these ions in our structures is 3.2 μm as estimated by TRIM-97,[9] placing the ions well into

*No proof corrections received from author prior to publication.

[a]Electronic mail: pgpiva@nrcphy1.phy.nrc.ca
[b]Also an Adjunct Professor with the Dept. of Physics and Astronomy, University of Western Ontario.

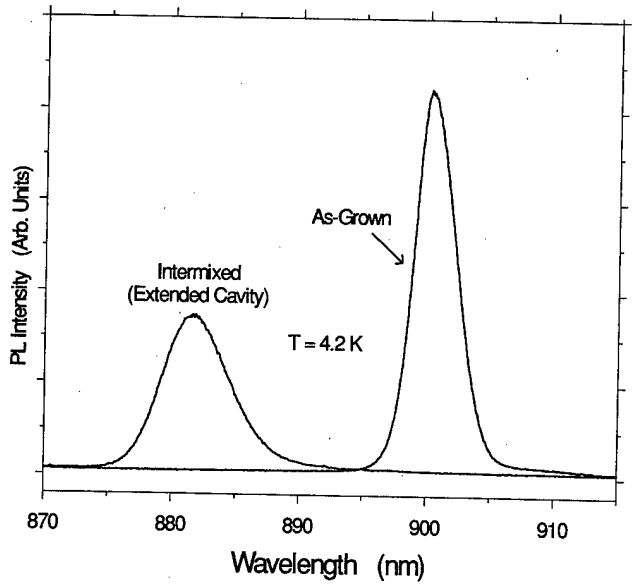


FIG. 1. Low temperature PL spectra for both intermixed (passive cavity material) and nonintgermixed (active device material) regions of the implanted InGaAs/GaAs DQW-GRINSCH laser sample.

the substrate. A second sample taken from an adjacent location on the wafer was used as a control.

For the InGaAs/InGaAsP/InGaP study, the samples were implanted using 30 keV, 80 keV, and 7 MeV P ions at 200 °C. The mean ion ranges were, respectively, 367 Å (the upper third of the topmost InGaP layer), 917 Å (short of the uppermost InGaP/InGaAsP interface), and 3.13 μm (ions penetrating the InGaAs QW and burying themselves well into the GaAs substrate) as determined by TRIM-97. Samples were half masked and implanted with fluences ranging from 1×10^{12} to 1×10^{15} ions/cm^2. The 1×10^{12} ions/cm^2 were implanted using a flux of 6.6 nA/cm^2, all other implants were performed with a flux of 60 nA/cm^2.

Following implantation, all samples were annealed in a N_2 ambient using a Heatpulse 610 rapid thermal annealer (RTA). It is during this stage that defects produced by the ion implantation serve to enhance the interdiffusion of barrier and well species, typically resulting in a blue-shifting of the QW band gap in the implanted region. The InGaAs/GaAs/AlGaAs samples were annealed at 850 °C for 60 s, as these conditions were previously found to saturate interdiffusion in the implanted regions while leaving the unimplanted material unshifted.[10] For the InGaAs/InGaAsP/InGaP samples an anneal temperature of 825 °C for 90 s was selected based on the same criteria. In all cases, during annealing the sample surfaces were protected from arsenic and gallium desorption by means of a GaAs proximity cap.[11]

Continuous-wave photoluminescence (PL) measurements were used to monitor the extent of the quantum well disordering achieved in the implanted regions. These measurements were performed at 4.2 K using weak HeNe excitation. The PL spectra were collected by a charge-coupled device coupled to a spectrometer (system resolution was 1.2 meV) or a Fourier transform infrared fitted with an InGaAs detector (resolution setting was 4 cm^{-1}) in the case of the InGaAs/InGaAsP/InGaP samples.

Figure 1 shows the low temperature PL spectra of the implanted and the as-grown regions after annealing for AlGaAs based 980 nm material. The position of the nonintermixed strip (determined by PL) on the implanted sample was recorded from a micrometer fitted to the cryostat's translation stage. The samples were then processed into ridge waveguide lasers. Details regarding processing of the laser diodes have been presented elsewhere.[8] Lasers with equivalent front and back facets were prepared with symmetrically extended cavities on both sides of the 400 μm central active region. Both intermixed and nonintermixed passive sections varied between $d=0$ μm (no extended cavity), and $d=500$ μm in length. No further treatments or coatings were applied to the resulting front and back facets.

The laser diodes were driven in a constant current mode with a 50% duty-cycle square wave at 80 Hz, with a constant base temperature of 20 °C. The current dependence of the wavelength-integrated power (L-I curves) for lasers without extended cavities and for lasers with band gap shifted extended cavities were similar with threshold currents of $\approx$12 mA, and output powers of $\approx$35 mW at 125 mA. The spectral outputs of these devices were typically comparable. In contrast, if any significant length of nonintermixed material is used to form an extended cavity, detrimental increases in cavity losses and threshold current occur. Lasing was not observed for devices with $d\geqslant200$ μm. This demonstrated the suitability of high energy implantation induced QW intermixing in producing transparent windows to distance the laser facets from the active regions of the ridge-laser devices.

Temperature measurements of the facets were made using the technique of reflectance modulation.[12] Details regarding this measurement have been presented elsewhere.[8] Heating coefficients characterizing the various devices were calculated by normalizing the peak temperature (found at the center of the optical mode) by the diode laser output. These values are plotted for the varying BSEC lengths in Fig. 2. The poorest performances (not plotted) were exhibited by the nonintermixed extended cavity lasers possessing heating coefficients in excess of 10 °C/mW. While low temperatures for these devices were measured, normalization to their corresponding reduced laser output levels revealed very high heating coefficients. For the bare cleaved facets without extended cavities, the heating coefficients reach slightly above 1 °C/mW. Dramatic decreases in the heating coefficients are observed, however as the length of the intermixed facets are increased. These values are $\approx$0.3 °C/mW, $\approx$0.15 °C/mW, and $\approx$0 °C/mW for $d=100$ μm, 200 μm, and 300 μm, respectively. That the heating coefficients for the facets at the center of the optical mode decrease gradually for BSEC lengths in excess of 100 μm indicate that in addition to reduced photoabsorption occurring in the disordered QWs, distancing of the facets from the electrical contact regions may be contributing to the observed decreases in facet temperatures.

Preliminary investigations of the evolution of laser facet

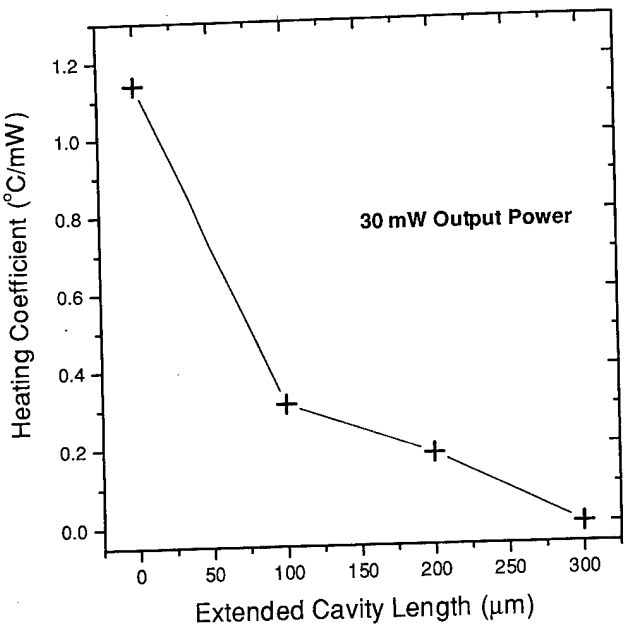


FIG. 2. Plot of measured laser facet heating coefficients as a function of BSEC length. Length of zero indicates a device without symmetrically extended BSEC sections.

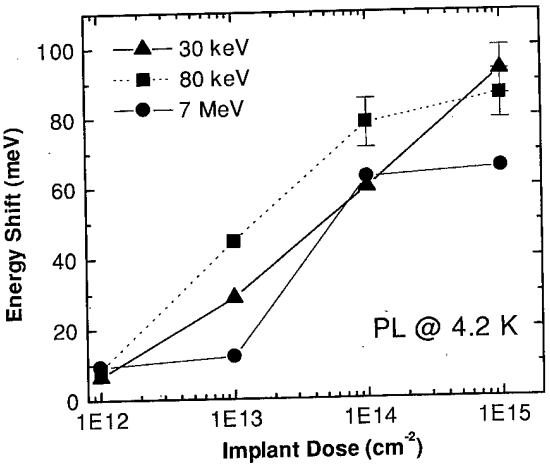


FIG. 3. Plot of achieved blue-shifts for the 980 nm InGaP based laser structure as a function of various P implant conditions. PL peak emission energy of the as-grown QW structure at 4.2 K was 1405.7 meV.

temperatures and lifetimes for these devices indicate associated increases in life-span and persistence of reduced facet temperatures for those devices possessing band gap shifted facets. A more detailed study with devices having lower contact resistances and coatings on the high reflector and output coupler facets is currently under way.

Concerns that the gains achieved by reduced photoabsorption in the facet region may be offset by nonradiative recombination at defects remaining in these structures following implantation and annealing[8] initiated a parallel investigation using a partial InGaAs/InGaAsP/InGaP laser structure. InGaAs/InGaAsP/InGaP laser diode lasers possess intrinsically more stable facets due to the lower surface recombination velocities[13] in these materials and an absence of aluminum which is both reactive with ambient oxygen, and associated with the formation and propagation of dark line defects.[2,14] Bare facet Al-free laser diodes have already demonstrated longevities in excess of 100-fold that of uncoated Al-based devices in lifetime testing.[15]

Figure 3 shows a plot of the achieved blue-shifts in the PL peak emission energies for the InGaAs/InGaAsP/InGaP samples at 4.2 K. Energy shifts were measured with respect to the emission energies of the unimplanted halves of each sample following annealing. Error bars present in the case of the higher fluence samples resulted from the superposition of the intermixed QW emission signal with emission lines from the underlying GaAs. Curve fitting was required to deconvolve the data, and margins of confidence were correspondingly assigned. From the data, only a small dependence on implant energy is apparent.[16] TRIM-97 yields the vacancy production efficiencies of 460, 870, and 6800 vacancies/ion for the 30, 80, and 7 MeV implants, respectively. As the initial defect concentrations are roughly confined to the penetration depths stated earlier for these ions, these results indicate that comparable fractions of the interdiffusion enhancing defects from the 30 and 80 keV implant sites are reaching the 85 Å InGaAs QW. In the case of the 7 MeV implants, clearly only a much smaller fraction of the total defects produced is participating in the enhancement of the QW interdiffusion (and most probably due in part to the much greater depths over which these defects are initially distributed following implantation).

A room temperature comparison between the integrated PL intensities of the implanted and masked regions of each of the samples was made. At a dose of 1×10^{13} cm^{-2} a relative decrease by factor of 5 (10) in the integrated PL intensities for the 30 keV (80 keV) implanted samples was noted. For the 7 MeV implanted samples, the losses were an order of magnitude greater. For higher fluence samples, persistence of PL emission was observed for the 30 and 80 keV implants, whereas the 7 MeV implanted material exhibited extinction of QW luminescence for fluences exceeding 1×10^{13} cm^{2}. From this we may infer that the 7 MeV implantation has deposited a far greater number of nonradiative recombination centers in the QW region (consistent with the penetration of these ions through the QW region during implantation) than those produced by the two lower energy implants (in which cases, both the ion ranges and corresponding damage profiles are confined to the overlying InGaP layer, and kept removed from the InGaAs QW prior to annealing). We expect the crystallinity of the InGaAsP confinement layers to be similarly affected.

III. CONCLUSIONS

In conclusion, we have shown the feasibility of band gap shifted extended cavities BSECs produced by high energy ion implantation in reducing the facet temperatures of 980 nm AlGaAs laser diodes. Results based on attempts to intermix Al-free 980 nm laser structures indicate that BSECs may be produced utilizing low energy ion implantation. Conse-

quently, BSECs may be produced in the latter material system without compromising the crystallinity of the epitaxy in the optical mode region.

ACKNOWLEDGMENT

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada.

[1]D. P. Cooper, C. H. Gooch, and R. J. Sherwell, IEEE J. Quantum Electron. **QE-2**, 329 (1966).
[2]C. H. Henry, P. M. Petroff, R. A. Logan, and F. R. Meritt, J. Appl. Phys. **50**, 3721 (1979).
[3]M. Uchida, S. Ishikawa, N. Takado, and K. Asakawa, IEEE J. Quantum Electron. **QE-24**, 2170 (1988).
[4]H. Blauvelt, S. Margalit, and A. Yariv, Appl. Phys. Lett. **40**, 1029 (1982).
[5]T. Shibutani, M. Kume, K. Hamada, H. Shimizu, K. Itoh, G. Kano, and I. Teramoto, IEEE J. Quantum Electron. **QE-23**, 760 (1987).
[6]J. Werner, E. Kapon, N. G. Stoffel, E. Colas, S. A. Schwarz, C. L. Schwartz, and N. Andreadakis, Appl. Phys. Lett. **55**, 540 (1989).
[7]J. H. Marsh, Semicond. Sci. Technol. **8** (1993), and references therein.
[8]P. G. Piva, S. Fafard, M. Dion, M. Buchanan, S. Charbonneau, R. D. Goldberg, and I. V. Mitchell, Appl. Phys. Lett. **70**, 1662 (1997).
[9]P. J. Poole, S. Charbonneau, G. C. Aers, T. E. Jackman, M. Buchanan, M. Dion, R. D. Goldberg, and I. V. Mitchell, J. Appl. Phys. **78**, 2367 (1995).
[10]J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Ion Range of Ions in Matter* (Pergamon, New York, 1985).
[11]S. Charbonneau, P. J. Poole, P. G. Piva, G. C. Aers, E. S. Koteles, M. Fallahi, J.-J. He, J. P. McCaffrey, M. Buchanan, M. Dion, R. D. Goldberg, and I. V. Mitchell, J. Appl. Phys. **78**, 3697 (1995).
[12]P. W. Epperlein, Jpn. J. Appl. Phys., Part 1 **32**, 5514 (1993).
[13]M. Razeghi, Nature (London) **369**, 631 (1994).
[14]D. Z. Garbuzov, N. Yu. Antonishkis, A. D. Bondarev, A. B. Gulakov, S. N. Zhigulin, N. I. Katsavets, A. V. Kochergin, and E. V. Rafailov, IEEE J. Quantum Electron. **27**, 1531 (1991).
[15]M. Razeghi, H. Yi, J. Diaz, S. Kim, and M. Erdtmann, Proc. SPIE **3001**, 243 (1997).
[16]This is in contrast to an energy dependence study performed on InGaAs/GaAs/AlGaAs 980 nm laser structures presented in Ref. 9 which indicated that for this system, ion implantation through the QW was necessary to initiate any sizable PL emission blue-shift. The relatively lesser differences apparent in Fig. 3 will be addressed in a subsequent publication.

Localized interface states and the optical spectra of AlSb/InAs heterostructures

M. J. Shaw,[a] G. Gopir, P. R. Briddon, and M. Jaros
Department of Physics, The University of Newcastle upon Tyne, NE1 7RU Newcastle upon Tyne, United Kingdom

(Received 7 November 1997; accepted 29 May 1998)

The existence of localized states at InSb-like interfaces of AlSb/InAs superlattices is predicted using empirical pseudopotential calculations. These predictions are shown to be in agreement with those of *ab initio* pseudopotential calculations performed using the local density approximation of density functional theory, demonstrating the ability of the empirical approach to describe the microscopic features of the interface. The frequency dependence of the absorption coefficient is calculated for a series of AlSb/InAs superlattices with differing interface configurations, and the role of the interface localization in determining the optical response is identified. © *1998 American Vacuum Society.* [S0734-211X(98)13404-2]

I. INTRODUCTION

There is a growing body of evidence that the optical and transport properties of many semiconductor heterostructures may be largely governed by the existence of states localized at the interfaces.[1,2] For the development of high performance optoelectronic devices the link between the microscopic features of the interfaces and the key parameters which determine the heterostructure response (e.g., lifetimes) must be established. In recent years heterostructure systems based on antimonide compounds and alloys have emerged as some of the most promising for optical device applications,[3] leading to the demonstration of a variety of infrared optical devices, including lasers,[4,5] detectors[6] and optical signal processing components.[7,8] A particular system for which the interfaces have become the subject of considerable attention is that of [001]-AlSb/InAs superlattice structures,[9,10] the combination of deep conduction wells with extremely high electron mobilities making them particularly attractive.[11,12] However, in order that such structures can be exploited to their full potential it is important to understand the microscopic processes which underlie the optical response, and specifically, to relate the properties of the interfaces to the key parameters which affect the optical spectra of these structures.

Localized interface states in AlSb/InAs heterostructures were first proposed by Kroemer *et al.*[13] in response to a series of experimental studies on AlSb–InAs–AlSb quantum wells.[10] These indicated that in the absence of intentional doping, and taking into account conventional electron sources (e.g., surface effects, deep donors in the barriers), a significant excess of electron concentration was observed. A possible explanation for this, consistent with the high measured mobility, was what Kroemer *et al.*[13] describe as a "generalized Tamm state." Since there are no ion species common to both well and barrier it is possible to envision two different atomic interface arrangements: AlAs-like and InSb-like. Experimentally, molecular beam epitaxy techniques have been shown to be capable of growing a particular combination of interface types.[10] The localized interface states proposed were states formed at InSb-like interfaces, most easily visualized by assuming the interface layer to be an extremely narrow bulk InSb-like quantum well. The valence band edges in such a system are shown schematically in Fig. 1. Subsequent studies suggest that while such states may indeed exist, they might not in fact be the source of the carriers.[14–16] In the present work we are not concerned with the source of electrons. Rather, we address the question of whether such localized states exist, and in particular consider the effect of these states on the optical characteristics of the structures.

In a recent study of ours[17] we used *ab initio* pseudopotential calculations to demonstrate interface-induced localization in a number of AlSb/InAs superlattices. However, while the *ab initio* calculations perform very well in the description of very short range features such as the interface localizations, they are not so well suited to the description of optical spectra. The evaluation of optical sums by this method is neither practical, due to the high computational demands, nor particularly desirable since the effects which dominate the optical spectra occur on a length scale comparable to the quantum well widths. Such longer-range effects are more appropriately described by a number of simpler models such as semiempirical pseudopotentials. These calculations give a reliable description of the longer-range features which determine the optical properties, and provide a practical tool for evaluation of optical sums. Other simpler models, such as the $k.p$ theory, are unable to provide such a reliable microscopic description of the states in complex systems such as the strained superlattices being studied here. In this study, we apply semiempirical pseudopotential calculations to the superlattice studied in Ref. 17 and find that, in addition to the accurate representation of the long-range features, this simpler model predicts qualitatively similar localized interface states to those predicted by the *ab initio* pseudopotential approach. The *ab initio* calculations of ours[17] are extended to provide a complete comparison of the two theoretical methods and the available experimental data. The success of the

[a]Electronic mail: M.J.Shaw@newcastle.ac.uk

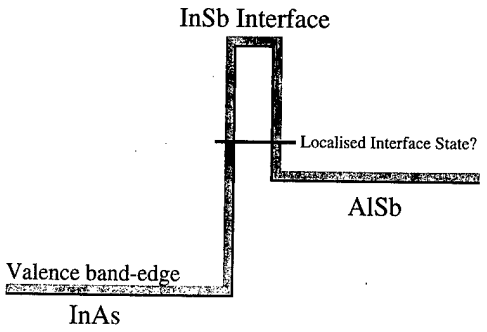


FIG. 1. Schematic diagram of the valence band edge at an InSb-like interface between AlSb and InAs as proposed by Kroemer *et al.* (1992).

empirical approach enables us to use it to perform full-scale evaluations of the optical coefficients, and we study how the absorption is affected by the composition of the interfaces in the superlattice.

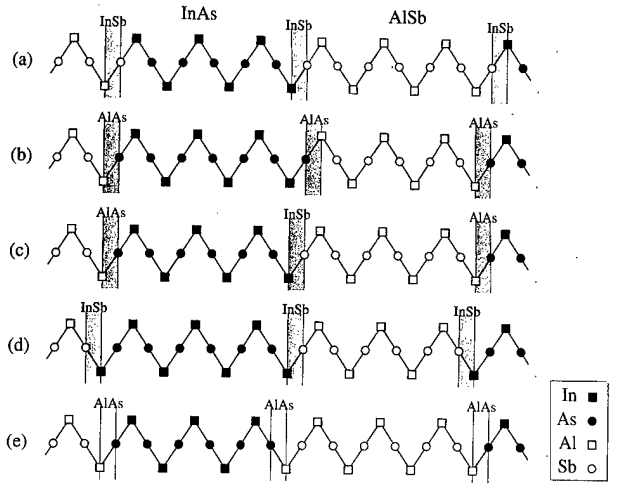


FIG. 2. Atomic composition of the five AlSb/InAs superlattices described in Table I are illustrated schematically. Different interface types are indicated.

II. THEORY OF ELECTRONIC STRUCTURE CALCULATIONS

In this article we present the results of semiempirical pseudopotential calculations for a number of [001]-AlSb/InAs superlattices. A detailed mathematical formulation of the calculations is available in the literature, and will not be reproduced here (see, for example, Ref. 18). In this method the two constituent materials (AlSb and InAs) are described by pseudopotentials empirically fitted to give a good description of key bulk band structure properties. The superlattice is described as a perturbation from one of the constituent materials (the "host") due to the presence of layers of the other (the "dopant"). Such an approach accounts for the microscopic effects of the individual atomic potentials, including the effect of the spin-orbit interaction.

The performance of such calculations is well established for superlattices of a number of material systems (see Ref. 19, and references within). However, the properties of the superlattice structures which have generally been of interest are those associated with the long-range order of the systems, i.e., effects arising through potential variations on the scale of the quantum well and barrier widths. How well might we expect these calculations to describe the microscopic interface features?

In order to answer this question we must consider more carefully the nature of the pseudopotentials involved. The pseudopotential for each atomic species is chosen to reproduce empirical band structure data for a given bulk compound. The net contribution to the potential from all many-body interactions, which will depend on the local environment of the atoms, is implicitly included in this pseudopotential. As a result, the atomic pseudopotential is only strictly applicable for atoms in that environment, i.e., in the particular bulk material for which the fit is carried out. For example, the Al pseudopotential obtained by a fit to AlSb is not transferable to an Al atom in AlAs. Where such pseudopotentials are used to describe a superlattice system, the potential in the bulklike well and barrier layers will be well represented. In contrast, the microscopic variation in the potential across the interface is not as well accounted for. First, an abrupt steplike potential is introduced to model the valence band offset between the two bulk materials. Second, the potentials assigned to the atoms forming the interface bond are the pseudopotentials associated with the well and barrier layer bulk constituents. As described above these are not rigorously applicable to the new material represented by the interface bonds. So, where the interest in the superlattice lies in long-range order effects, which is usually the case in studies of quantum-confined states, the empirical pseudopotential scheme is expected to perform well. In the present study, however, where the focus of attention is the interfaces themselves, the reliability of the calculations must be further established.

In order to verify the interface description provided by the empirical pseudopotential calculations, we present the results of *ab initio* pseudopotential calculations for comparison. Such calculations, using the local density approximation (LDA) of density functional theory, were reported by us[17] in a study of a variety of AlSb/InAs structures, and we include in this article an extended analysis of the results of these calculations. We described the full details of the *ab initio* calculations in Ref. 17 (and references within), and these will not be reproduced here. The key feature of these calculations, with regard to the present discussion, is the use of the first principles norm-conserving pseudopotentials of Bachelet *et al.*[20] These are derived from the exact atomic charge densities and are fully transferable. The method includes many-body effects explicitly and does not involve any reference to the bulk material properties. The pseudopotentials used are applicable regardless of the environment in which the atoms are placed. Such potentials are thus ideal for the description of microscopic interfacial properties. The problems facing the empirical method concerning the interface potentials simply do not arise.

The empirical pseudopotential calculations presented in this article assumed that the positions of the atoms are fixed

TABLE I. Interface types and structural properties of the superlattices studied. Number of each atom type in one unit cell is listed as n_{In}, n_{As}, n_{Al} and n_{Sb}.

Structure	Interface types	n_{In}	n_{As}	n_{Al}	n_{Sb}
(a)	InSb+InSb	6	5	6	7
(b)	AlAs+AlAs	6	7	6	5
(c)	InSb+AlAs	6	6	6	6
(d)	InSb+InSb	7	6	5	6
(e)	AlAs+AlAs	5	6	7	6

to those of an unstrained zinc-blende lattice. The effect of lattice mismatch between the AlSb and InAs layers is accounted for simply through a change to the valence band offset, in accordance with simple elastic theory.[21] Since the mismatch between InAs and AlSb is relatively small (≈1.35%), the freezing of atom positions should not greatly affect the description of well width scale features. However, the formation of the hybrid materials AlAs and InSb at the interfaces results in highly strained bonds. It is not clear that the neglect of strain remains a valid approximation in the case of these interfaces. When analyzing the results of the empirical calculations we must consider the role strain may play in affecting the interface properties. In contrast, the LDA calculations allow a complete relaxation of all atom positions, implicitly incorporating the full effect of strain in the layers and at the interfaces.

In this article we compare the results obtained using empirical pseudopotentials with those of the *ab initio* approach.

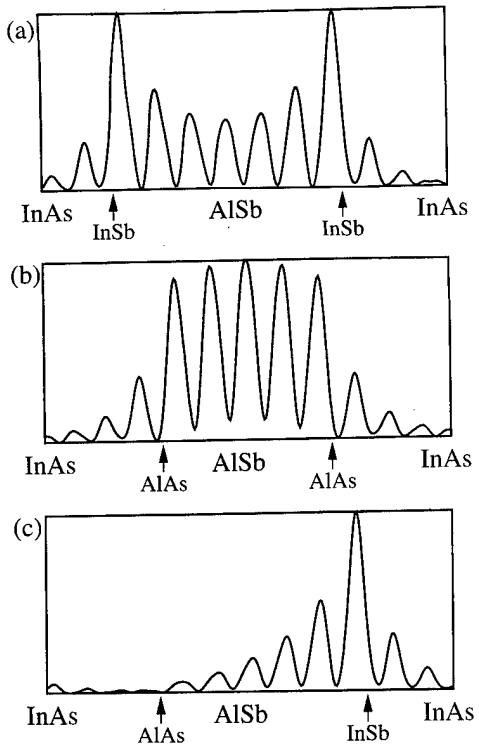


FIG. 3. Comparison of the charge density of the uppermost valence band in structures (a), (b) and (c) as calculated using the empirical pseudopotential method. Interface types are indicated.

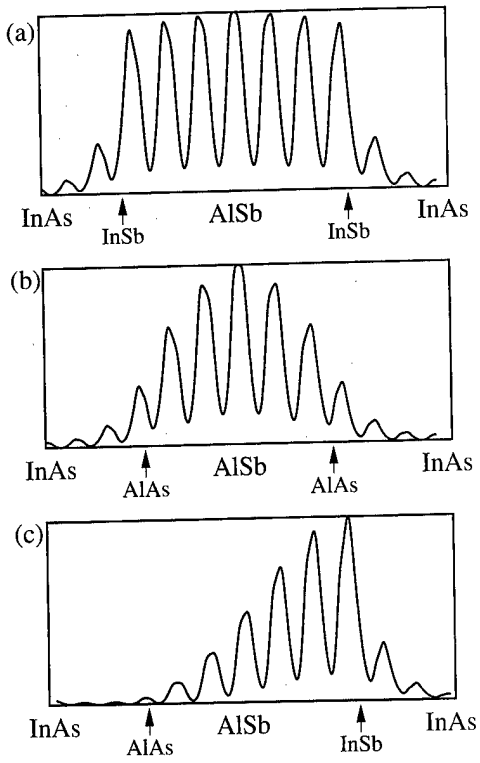


FIG. 4. Comparison of the charge density of the uppermost valence band in structures (a), (b) and (c) as calculated using the *ab initio* pseudopotential method. Interface types are indicated.

This enables us to determine how well the empirical method is able to describe the microscopic interface features which have been identified with the *ab initio* calculations. Having verified the performance of the empirical method we are able to proceed to full-scale calculations of the optical absorption coefficient, using the empirical band structure calculations.

III. RESULTS OF ELECTRONIC STRUCTURE CALCULATIONS

A. Localized interface states

The structures which we have studied are based on a 3 AlSb/3 InAs superlattice (where 3 AlSb indicates 3 lattice constants, 6 monolayers of AlSb), where one interface is AlAs-like and one InSb-like. A series of structures is then defined in which the interface types have been changed: two with both interfaces AlAs-like, two with both InSb-like, and the original structure with one of each interface. These are illustrated schematically in Fig. 2, and their atomic composition is listed in Table I. Empirical pseudopotential calculations were performed for these five structures. As described in Sec. II, the effects of strain were neglected, with an average lattice constant used for the calculations. The zone center charge density of the uppermost valence state is plotted for the first three structures of Table I in Fig. 3. For comparison, the charge densities obtained by the *ab initio* method are shown in Fig. 4. Comparison of Figs. 3 and 4 shows that the charge densities predicted by the two methods are very simi-

TABLE II. Changes in band-gap energies for structures with different interface configurations. Calculated gaps, ΔE_g^{EPP} (empirical pseudopotential) and ΔE_g^{LDA} (*ab initio* pseudopotential), are compared with experimental data of Spitzer *et al.* (Ref. 2).

Structure	Number[a]	ΔE_g^{EPP} (meV)	ΔE_g^{LDA} (meV)	$\Delta E_{AlSb}^{Expt.}$ (meV)
(a)	1	+57	-2	+22
(b)	2	+10	+112	+120
(c)	5	0	0	0
(d)	*	−64	−103	*
(e)	*	+123	+209	*

[a]Reference 2.
*No experimental results available.

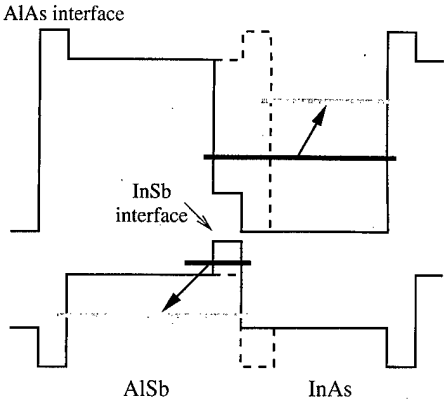


FIG. 5. Band-edge profile associated with the simple particle-in-a-box picture of the interfaces. Solid line shows structure (c) with one of each interface, while the dashed line shows the profile for (e). Ground conduction and valence states are indicated by dark shaded lines for (c) and light shaded lines for (e).

lar. In particular, the localization at the InSb-like interfaces, which was reported in our previous paper,[17] is well represented in the empirical results of Fig. 3. Structure (c) shows charge localized at the InSb interface while structure (a), with two InSb-like interfaces, shows charge drawn away from the center of the well. In contrast, the structure with two AlAs interfaces shows a conventional symmetrical ground state envelope function. The empirical pseudopotential approach is clearly able to successfully predict the InSb interface localizations which are obtained using the *ab initio* approach.

In our earlier paper[17] we provide a comparison between the results of the *ab initio* calculations and experimental data available in the literature. Specifically the dependence of the fundamental gap on interface structure was compared with the photoluminescence data of Spitzer *et al.*[2] for the structures listed (a)–(c) in Table I. Since the LDA is well known to underestimate the absolute value of the band gap, changes in gap were compared instead [using structure (c) as a reference]. In Table II we present the changes in the band gaps determined from the empirical pseudopotential (EPP) calculations, and compare these to both the LDA results and, where available, the experimental data. Where Spitzer *et al.* report spectra for more than one sample with a given nominal structure (e.g., where they compare the effects of the order of interface growth) we have compared them with the sample shown to exhibit the highest quality.

Examination of Table II shows that the band-gap changes predicted by the empirical pseudopotential method are not consistent with the experimental data. While the LDA calculations give a good agreement with the available data, the EPP method does not predict the correct order for the gaps. Clearly, while the qualitative features of interface localization are described by the empirical method, the quantitative predictions of energies are not so well described. This is not surprising given the nature of the potentials used to describe the interfaces.

When making a comparison with the experimental results it is important to acknowledge the practical difficulties presented by the huge interface mismatch. There is evidence that the growth quality is particularly poor for lower AlAs-like interfaces,[2] and indeed studies of phonon spectra call into question the existence of AlAs interface bonds.[22] However, the theoretical treatment of the possible interface reconstructions and growth imperfections which might result lies beyond the scope of the present study. The calculations presented here assume a pseudomorphic growth.

The changes in energy gaps involve a combination of two effects, namely the formation of localized states, and the changes in conduction and valence well widths accompanying them. To analyze further the discrepancies between the EPP results and the LDA and experimental gap changes it is useful to invoke a simple particle-in-a-box picture of the superlattices. In addition we extend the analysis of the LDA results to extract the changes in the conduction and valence band edges individually, to enable direct comparison with the EPP method.

B. Particle-in-a-box interpretation

The simplest interpretation of the localization effect is that presented by Kroemer *et al.*,[13] where the InSb- and AlAs-like interfaces are considered as very narrow bulk layers in a particle-in-a-box picture. For structure (c) with one of each interface this leads to the band-edge diagram shown in Fig. 5. Viewing the InSb-like interface as a valence band quantum well, the localized interface state is regarded as being confined in this well. While such a description serves to help understand the qualitative features of the localization, it is stressed that such a particle-in-a-box picture is not suitable for quantitative calculations. In particular, one must bear in mind that to assign bulk concepts such as band edges to the interface layers belies the fact that these are just single highly strained bonds of the material. Despite the limitations of this simple model, it is possible to make predictions concerning the changes in band gaps with the interface configuration based upon it.

As stated previously, the band-gap change is a combination of the change in the well widths and the change in energy associated with the confinement of a state in the narrow InSb well (if such an interface exists). For example consider the change in band gap as we go from structure (c) to (e), as illustrated schematically in Fig. 5. The interface band edges

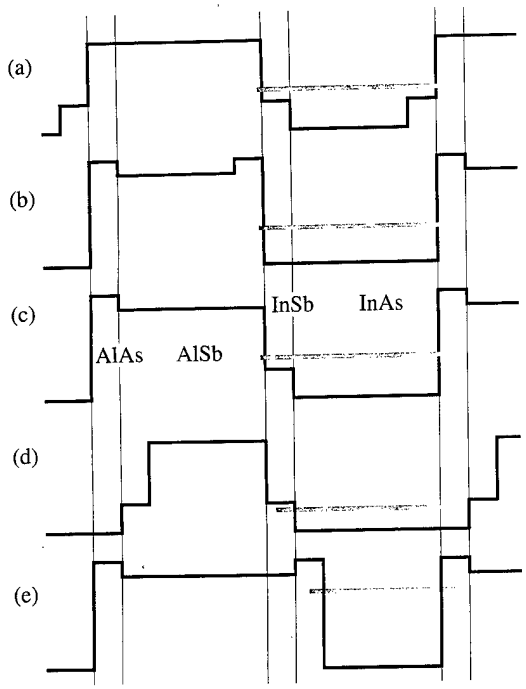


FIG. 6. Comparison of the particle-in-a-box conduction band-edge profiles for the structures in Table I. Positions of the lowest conduction state in each structure is indicated by the shaded lines. Vertical lines show the positions of the interfaces in the reference structure, (c).

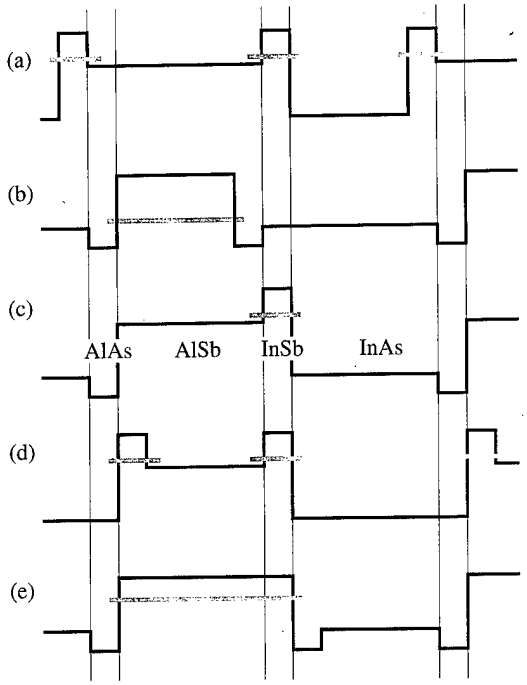


FIG. 7. Comparison of the particle-in-a-box valence band-edge profiles for the structures in Table I. Positions of the lowest valence state in each structure is indicated by the shaded lines. Vertical lines show the positions of the interfaces in the reference structure, (c).

are estimated from the model solid predictions for the bulk materials concerned. The energy of the conduction band edge is expected to increase since the width of the effective conduction well has been decreased. The valence band edge jumps down in energy since the localized state, whose energy is relatively insensitive to the AlSb width, is no longer present and the valence well quantum confinement energy lowers the energy of the valence band edge state. Both effects result in an increase in the band gap so we expect the band gap to increase considerably as seen in Table II.

Similar analysis can be repeated to compare the band-edge changes and band-gap changes in all five structures. Figure 6 shows the schematic conduction band-edge profiles for the five structures studied. Approximate positions of the lowest conduction energy levels are indicated. Since the conduction wells in structures (a), (b) and (c) differ only by a monolayer of InSb (which presents a small barrier potential) we expect the conduction edges in these to be at approximately the same energy. In contrast, the well width in (d) is significantly greater than (c) so we expect the energy in (d) to be lower, while for structure (e) the narrower well should result in an increased confinement energy. The expected conduction band-edge changes may be summarized as follows:

$$E_C^{(d)} < E_C^{(a)} \approx E_C^{(b)} \approx E_C^{(c)} < E_C^{(e)}, \tag{1}$$

where $E_C^{(i)}$ is the energy of the conduction band-edge in structure (i).

The valence band-edge profiles are shown in Fig. 7. Assuming that the localized interface states are of approximately the same energy regardless of the other layer widths in the superlattice, it is clear that the edges in structures (a), (c) and (d), each with at least one InSb interface, should be approximately the same. Where no localized state exists the edges are lowered by the AlSb well confinement, depending on the thickness of the AlSb layer. Thus the simple band-edge model predicts:

$$E_V^{(b)} < E_V^{(e)} < E_V^{(a)} \approx E_V^{(c)} \approx E_V^{(d)}. \tag{2}$$

Of course, experimentally it is not possible to measure changes in the individual band edges. Rather one measures changes in the gap. Combining the above inequalities it can be seen that

$$E_g^{(d)} < E_g^{(a)} \approx E_g^{(c)} < E_g^{(b)} < E_g^{(e)}, \tag{3}$$

where $E_g^{(b)} < E_g^{(e)}$ is deduced from the fact that the conduction band effective mass is smaller than that of the valence band.

C. Interface dependence of band-edges

In order to determine how well this intuitive picture explains the behavior of the system we must compare the predictions above with the results of the LDA calculations. Examination of Table II in conjunction with Eq. (3) shows a good correlation between the ordering predicted by the particle-in-a-box model and the band-gap changes calculated from the *ab initio* pseudopotential method. To compare the changes in the individual band-edges, it is necessary to relate the one-electron eigenvalues obtained from the calculations for the different structures to a common energy scale. How-

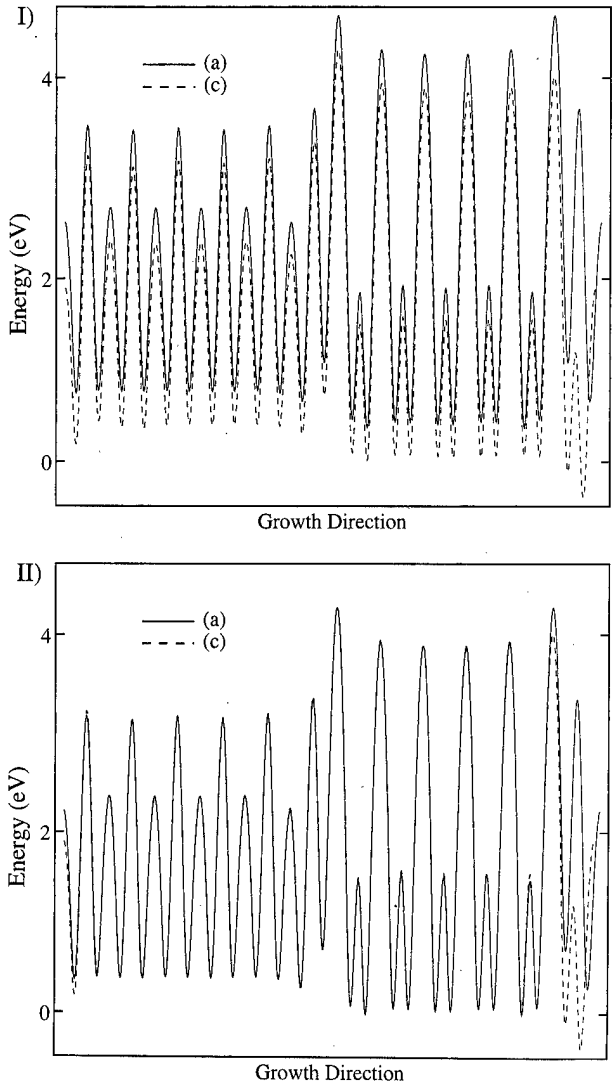


FIG. 8. Total self-consistent potential, integrated in the plane parallel to the interfaces, is plotted against position along the growth axis for structures (a) and (c). In I the potentials in the two calculations are compared directly. In II the profile for (a) has been adjusted to align the potentials in the central regions of the layers.

ever, the one-electron energies which one obtains directly from the LDA calculations are not comparable between different unit cells.

To relate the energy scales of the different structures we need to find a reference point for the self-consistent potentials in the different unit cells. Since the structures differ only in the immediate interface region, the potential near the center of the InAs and AlSb layers should be common to all structures (i.e., the "bulk" potentials). To compare one-electron energies between two structures it is simply necessary to calculate the shift in the total potential required in order to align the potential in regions away from the interfaces. This gives the energy shift which must be applied to the eigenvalues to place them on a common energy scale.

As an example, consider the case of the structures (a) and (c). The total potentials for these structures, integrated in the plane parallel to the interfaces, are plotted in Fig. 8(I). Except for the interface at the edge of the diagram, it is clear that the form of the potentials in the center of the layers is almost identical. It is therefore relatively straightforward to determine the shift required to map the curves onto each other. Figure 8(II) shows the adjusted potentials, demonstrating the extent of the healing of the potential within a few angstroms of the changed interface. By computing the change in potential required to align the profiles as in Fig. 8 one is able to derive a common energy scale. It is interesting to note how well the potential matches at the center of the two layers simultaneously, demonstrating that the interface valence band offset is not sensitive to the interface type, i.e., that the offset is an intrinsic property for these structures.

TABLE III. Conduction band-edge energies determined from the local density (LDA) and empirical pseudopotential (EPP) calculations, compared to the ordering predicted from the simple particle-in-a-box model.

Prediction	$E_C^{(d)}$	$<$	$E_C^{(a)}$	$\approx$	$E_C^{(b)}$	$\approx$	$E_C^{(c)}$	$<$	$E_C^{(e)}$
LDA (meV)	−117		8		9		0		111
EPP (meV)	−47		73		−68		0		53

Using the method described above, the conduction and valence band edges for structures (a), (b), (d) and (e) were related to those of structure (c) [with the edge in (c) set to zero]. The conduction band edges are presented in Table III, in comparison with the predictions of the simple particle-in-a-box model. Also given are the conduction edge energies obtained from the empirical pseudopotential method. The equivalent results for the valence band are given in Table IV. The remarkable agreement between the calculated LDA band edges and the ordering predicted from the simple particle-in-a-box picture suggests that this simple intuitive model does indeed describe well the behavior of the system. The formation of the localized interface states can be understood in keeping with the original proposal of Kroemer *et al.*

We may also now make a more detailed evaluation of the performance of the empirical pseudopotential method by directly comparing the band-edge energies with those obtained from the LDA approach. The EPP calculations reproduce quite well the variations in the valence band edges, which of course include the interface localization effects in which we are primarily interested. The discrepancy in the prediction of band gaps, seen in Table II, is therefore seen to arise through a significant difference in the calculated conduction edges for which there is almost total disagreement. The reason for the breakdown of the EPP in the conduction band may be understood by considering the effect of neglect on the strain

TABLE IV. Valence band-edge energies determined from the local density (LDA) and empirical pseudopotential (EPP) calculations, compared to the ordering predicted from the simple particle-in-a-box model.

Prediction	$E_V^{(b)}$	$<$	$E_V^{(e)}$	$<$	$E_V^{(a)}$	$\approx$	$E_V^{(c)}$	$\approx$	$E_V^{(d)}$
LDA (meV)	−103		−97		10		0		−14
EPP (meV)	−79		−69		16		0		17

in these calculations. By straining the interface InSb and AlAs bonds to the zinc-blende structure of the (assumed unstrained) AlSb–InAs layers we have introduced large but spurious hydrostatic shifts in the conduction band edges. Since the InSb has been compressed its conduction band edge is increased in energy, the converse occurring in the AlAs. As a result, the effective conduction barrier potentials of the InSb and AlAs layers assumed in the intuitive model (see Fig. 6), and confirmed by comparison to the LDA results (where of course strain is fully accounted for), are not recreated in the empirical pseudopotential approach. The conduction edge variations obtained from the EPP calculations may still be understood in terms of the particle-in-a-box model, but where the relative effective conduction barriers presented by the InSb and AlAs layers are reversed.

IV. OPTICAL SPECTRA OF AlSb/InAs SUPERLATTICES

Having established that the empirical pseudopotential calculations correctly describe the essential features of the AlSb/InAs structures, namely the interface localization effects, we now turn our attention to the optical properties of these structures. As mentioned previously, the principle motivation behind the development of these structures lies in their optoelectronic device potential. In order to realize this potential there is a clear need to understand the optical response characteristics of these structures, and in particular the role which the localized interface states play.

A. Calculation details

The empirical pseudopotential method has been used extensively to perform full-band structure calculations of the linear and nonlinear optical properties of a variety of superlattice systems,[23–25] providing an efficient scheme for identifying the microscopic processes involved. Here we apply our EPP calculations to the evaluation of the linear absorption spectra of the AlSb/InAs superlattices whose electronic structure has been described in Sec. III. The pseudopotential calculations provide the band structure information required to calculate the frequency dependence of the first-order susceptibility from the following expression, derived from density matrix theory[26]:

$$\chi^{(1)}_{\mu\alpha}(-\omega;\omega) = -\frac{e^2 N}{Vm\epsilon_0\omega^2}\delta_{\mu\alpha} + \frac{e^2}{Vm^2\epsilon_0\hbar\omega^2} \times \sum_k \sum_a f_a \sum_b \left[\frac{p^{\mu}_{ab} p^{\alpha}_{ba}}{(\Omega_{ba}-\omega)} + \frac{p^{\alpha}_{ab} p^{\mu}_{ba}}{(\Omega_{ba}+\omega)} \right]. \quad (4)$$

Here, ω is the frequency of the incident and response photons of polarization α and μ, respectively. The summations over k, a and b represent summations over pairs of states with wave vector k and band indices of a and b. The numerators are products of momentum matrix elements, p^{δ}_{ab}, of polarization δ, and are multiplied by f_a, the Fermi factor of state a. The frequency denominators include terms $\Omega_{ij} = (E_i - E_j)/\hbar$, where E_i is the energy of the state i. The imaginary part of the first-order susceptibility is related to the absorption coefficient by the equation

$$\alpha_{\mu\alpha}(\omega) = \frac{\omega}{n_0 c}\mathrm{Im}[\chi^{(1)}_{\mu\alpha}(-\omega;\omega)]. \quad (5)$$

Here, n_0 is the background refractive index, and $\chi^{(1)}$ is in SI units.

The summation over all wave vectors in the Brillouin zone tends, in the limit of a crystal of infinite extent, to an integral over wave vector space. To obtain a practical evaluation of the imaginary part of this integral we use the linear tetrahedron method, sampling a tetrahedral grid of wave vectors with our empirical pseudopotential calculations. To reduce the number of sampling points required, the sampling volume is restricted to a relatively small region of the zone in the x/y plane (the plane parallel to the interfaces) around the zone center, Γ. Since the magnitude of the energy gap increases rapidly as we move away from Γ, the contributions to $\chi^{(1)}$ decrease sharply at the frequencies of interest. The size of the sampling region is chosen to ensure that the susceptibility converges. The whole extent of the zone is sampled in the growth (z) direction. A further reduction in computation is achieved by sampling only the irreducible segment of the Brillouin zone, applying the appropriate transformations to exploit the point group symmetry of the crystal.

B. Calculated absorption spectra

The absorption coefficients for the structures labeled (a), (b) and (c) in Table I were calculated for photon energies between 0.8 and 1.3 eV, energies around the fundamental gap of these superlattices. These three structures exhibit the interface configurations of one of each type, two AlAs-like, and two InSb-like interfaces, and enable us to assess the effect of interface configuration on absorption. The spectra of the other two structures do not show any qualitatively different effects. For each structure the α_{xx} and α_{zz} components of the absorption, representing normal and parallel incidence absorption, respectively, were calculated.

The band structure of the reference structure, structure (c), with one of each interface type, is plotted in Fig. 9 along the [001] and [100] symmetry axes. From this diagram it is clear that for the 0.8–1.3 eV energy range of interest, only the lowest conduction band can be involved in interband absorption processes. In contrast, there are a number of valence bands lying relatively close to the gap and which might be featured in the interband processes. In particular we shall concentrate on the two highest valence bands which are expected to dominate the absorption spectrum around the fundamental gap. While the band structures of the three structures differ in detail, the principal features discussed above remain the same.

In Fig. 10 we plot the total absorption α_{xx}, together with the contributions from the individual processes $V1 \rightarrow C1$ and $V2 \rightarrow C1$ (where $V1$ and $V2$ are the highest and second high-

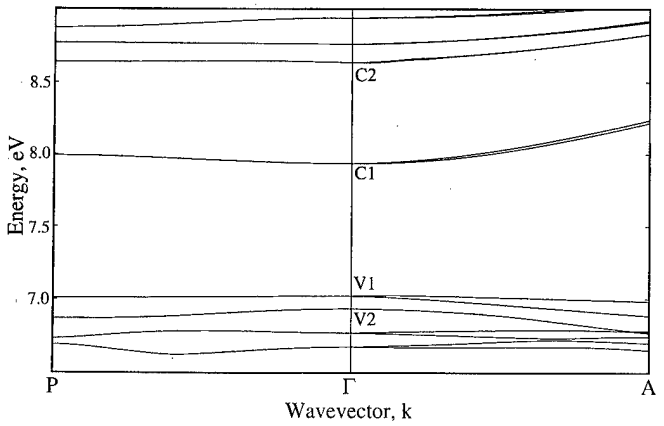


FIG. 9. Electronic band structure for the reference superlattice, (c), is plotted along symmetry lines, [001] and [100].

est valence states and $C1$ is the lowest conduction state) against photon energy for the three structures. Each of the spectra exhibit a strong absorption edge corresponding to the zone-center $V1 \rightarrow C1$ transition edge, and at higher energies a second, shallower, edge originating in the $V2 \rightarrow C1$ transitions. In addition, each structure shows a sharp spectral feature associated with transitions from $V2$ at higher energies (originating from wave vectors lying away from the zone

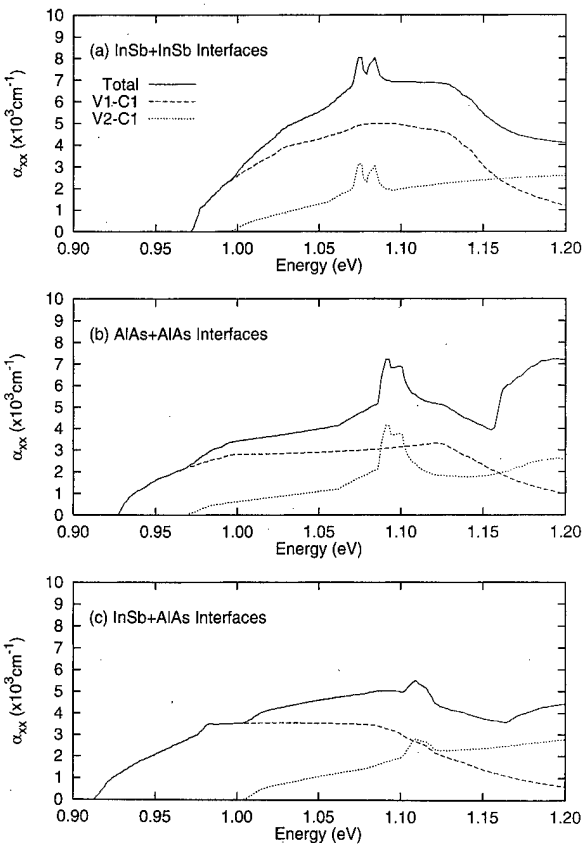


FIG. 10. Absorption coefficient α_{xx} (cm^{-1}), corresponding to normal incidence, is shown as a function of photon energy (eV) for structures (a), (b) and (c). Solid line shows the total absorption while the individual contributions from $V1 \rightarrow C1$ and $V2 \rightarrow C1$ are plotted separately.

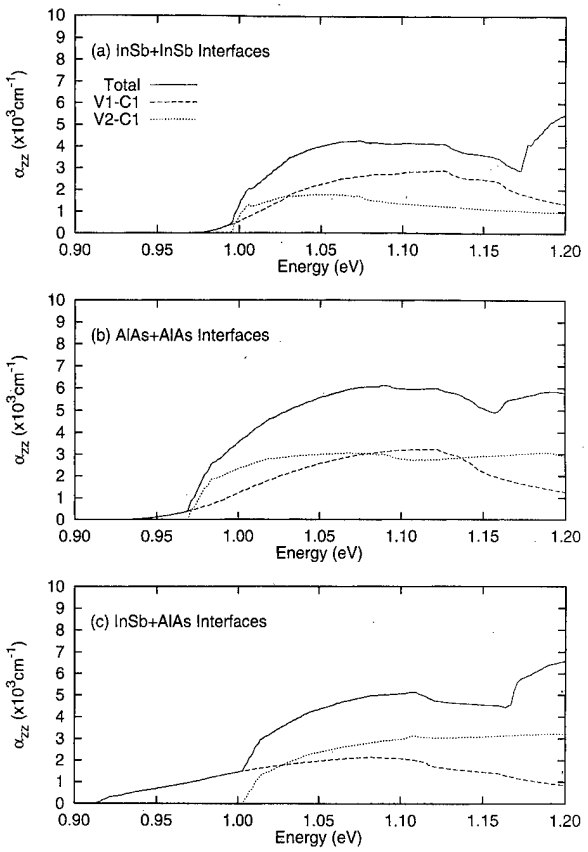


FIG. 11. Absorption coefficient α_{zz} (cm^{-1}), corresponding to parallel incidence, is shown as a function of photon energy (eV) for structures (a), (b) and (c). Solid line shows the total absorption while the individual contributions from $V1 \rightarrow C1$ and $V2 \rightarrow C1$ are plotted separately.

center), and finally, towards the upper end of the photon energy range, contributions from lower-lying valence bands. Significantly, then, the qualitative features of all three structures are very similar, indicating that the presence of localized interface states does not dramatically affect the form of the linear absorption spectrum.

The parallel incidence absorption spectra for the three structures are shown in Fig. 11. Again the individual contributions from $V1 \rightarrow C1$ and $V2 \rightarrow C1$ transitions are shown. For α_{zz}, the $V1$ absorption edge is considerably weaker than the higher energy $V2$ edge—in particular, for the structures with two AlAs and two InSb interfaces the absorption edge at the fundamental gap is very shallow indeed. The reason for these very weak absorption edges can be understood by examining in detail the wave vector dependence of the momentum matrix elements. Figure 12 shows the squared magnitudes of the z component of the momentum matrix elements against the wave vector for the $V1 \rightarrow C1$ and $V2 \rightarrow C1$ transitions in the three structures. For structures (a) and (b) it is clear that the $V1 \rightarrow C1$ transition is forbidden at the zone center, increasing as the parallel component of the wave vector increases. For structure (c), whose point group symmetry is reduced by virtue of the inequivalent interfaces, the transition is no longer strictly forbidden at the zone center. However, it remains very much weaker than $V2 \rightarrow C1$.

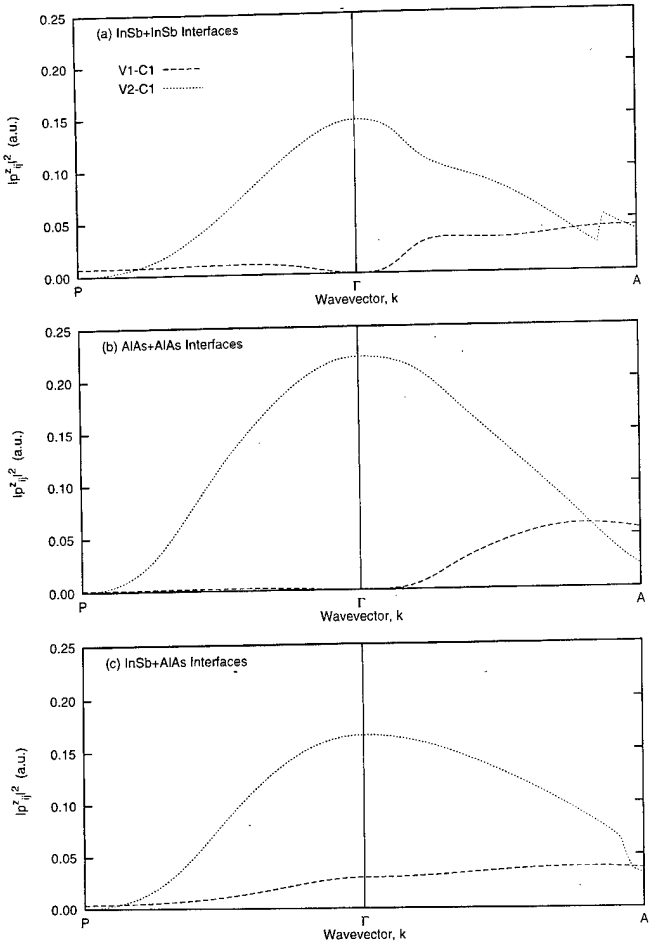


FIG. 12. Wave vector dependence of the squared magnitude of the optical matrix elements p_{ij}^z is plotted for structures (a), (b) and (c). Dashed lines represent transitions $V1 \rightarrow C1$ while the dotted lines represent transitions $V2 \rightarrow C1$. P is the point on the edge of the Brillouin zone in the [001] direction and A lies on the [100] axis, one-tenth of the way to X.

Since the $V1 \rightarrow C1$ transitions are either forbidden or weak at the zone center, the absorption at the zone center energy separation is very weak. Thus the absorption edge associated with $V1$ is considerably weaker than that of $V2$, from which relatively strong transitions to $C1$ can occur at the zone center. Although the differences between the α_{zz} spectra for the three structures are more significant than they were for α_{xx}, the largest differences exist between structure (c) and the other two. These differences originate in the change of symmetry in going to structures with two identical interfaces, not in the formation of localized interface features.

Finally, it is interesting to consider how the effective mass of the localized valence states may differ from the conventionally confined states in the structures with no interface states. The dispersion of the miniband at the valence band edge along the [001] direction is plotted for all five structures in Fig. 13. While the different structures clearly show significant differences in miniband curvature, particularly in the immediate vicinity of the zone center, once again there is no direct correlation between the masses and the existence of the localized interface states.

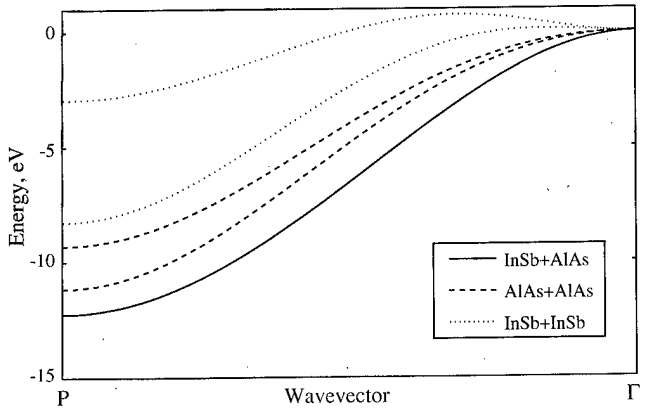


FIG. 13. Variation of the energy of the uppermost valence miniband (eV) with wave vector along the [001] axis for the five structures in Table I.

V. CONCLUSION

In conclusion, we have demonstrated the success of the empirical pseudopotential method at describing the electronic structure of AlSb/InAs superlattices through comparison with the results of *ab initio* pseudopotential calculations and existing experimental data. The empirical calculations have been shown to give a good description of the principle features of these structures, in particular reproducing the interface-induced localizations which have been shown to occur at InSb-like interface in these structures. In addition, detailed analysis of the results of the *ab initio* calculations confirms that the simple intuitive particle-in-a-box picture originally proposed by Kroemer *et al.*[13] does indeed provide a realistic qualitative description of the behavior of the system.

Having established the ability of the empirical pseudopotential method to describe the interfacial features, we used this approach to study the optical properties of a series of AlSb/InAs superlattices. It is clear from the calculations of absorption spectra presented in this article that the qualitative features of the linear absorption spectra are not affected by the existence of the states localized at the InSb-like interfaces. However, the method outlined in this article is easily extended to the study of optical nonlinearities, and it remains for future study to determine the effect which the interface localization might play in higher-order optical processes.

ACKNOWLEDGMENTS

The authors would like to thank the Computational Science Initiative of the U. K. Engineering and Physical Science Research Council, and the Office of Naval Research (U.S.A.) for financial support.

[1]R. J. Turton and M. Jaros, Appl. Phys. Lett. **69**, 2891 (1996).
[2]J. Spitzer, A. Höpner, M. Kuball, M. Cardona, B. Jenichen, H. Neuroth, B. Brar, and H. Kroemer, J. Appl. Phys. **77**, 811 (1995).
[3]T. C. McGill and D. A. Collins, Semicond. Sci. Technol. **8**, S1 (1993).
[4]R. H. Miles, D. H. Chow, Y.-H. Zhang, P. D. Brewer, and R. G. Wilson, Appl. Phys. Lett. **66**, 1921 (1995).
[5]J. I. Malin, J. R. Meyer, C. L. Felix, J. R. Lindle, L. Goldberg, C. A. Hoffman, F. J. Bartoli, C.-H. Lin, P. C. Chang, S. J. Murry, R. Q. Yang, and S.-S. Pei, Appl. Phys. Lett. **68**, 2976 (1996).

[6]L. A. Samoska, B. Brar, and H. Kroemer, Appl. Phys. Lett. **62**, 2539 (1993).
[7]H. Xie, W. I. Wang, J. R. Meyer, C. A. Hoffman, and F. J. Bartoli, J. Appl. Phys. **74**, 2810 (1993).
[8]J. R. Meyer, I. Vurgaftman, and C. A. Hoffman, Appl. Phys. Lett. **81**, 7662 (1997).
[9]B. Brar, J. Ibbetson, H. Kroemer, and J. H. English, Appl. Phys. Lett. **64**, 3392 (1994).
[10]G. Tuttle, H. Kroemer, and J. H. English, J. Appl. Phys. **67**, 3032 (1990).
[11]G. Tuttle, H. Kroemer, and J. H. English, J. Appl. Phys. **65**, 5239 (1989).
[12]J. Wagner, J. Schmitz, H. Obloh, and P. Koidl, Appl. Phys. Lett. **67**, 2963 (1995).
[13]H. Kroemer, C. Nguyen, and B. Brar, J. Vac. Sci. Technol. B **10**, 1769 (1992).
[14]S. Ideshita, A. Furukawa, Y. Mochizuki, and M. Mizuta, Appl. Phys. Lett. **60**, 2549 (1992).
[15]D. J. Chadi, Phys. Rev. B **47**, 13478 (1993).
[16]J. Shen, H. Goronkin, J. D. Dow, and S. Y. Ren, J. Appl. Phys. **77**, 1576 (1995).
[17]M. J. Shaw, P. R. Briddon, and M. Jaros, Phys. Rev. B **52**, 16341 (1995).
[18]M. A. Gell, D. Ninno, M. Jaros, and D. C. Herbert, Phys. Rev. B **34**, 2416 (1986).
[19]M. Jaros, Semicond. Semimet. **32**, 175 (1990).
[20]G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[21]C. Van de Walle, Phys. Rev. B **39**, 1871 (1989).
[22]I. Sela, C. R. Bolognesi, L. A. Samoska, and H. Kroemer, Appl. Phys. Lett. **60**, 3283 (1992).
[23]M. J. Shaw, K. B. Wong, and M. Jaros, Phys. Rev. B **48**, 2001 (1993).
[24]M. J. Shaw and M. Jaros, Phys. Rev. B **50**, 7768 (1994).
[25]S. A. Hosseini, M. J. Shaw, and M. Jaros, Phys. Rev. B **53**, 6988 (1996).
[26]P. N. Butcher and T. P. MacLean, Proc. Phys. Soc. London **81**, 219 (1963).

Chemical beam epitaxial growth of $GaAs_{1-x}P_x$ on GaAs (100) substrates

D. Wildt, B. J. García,[a] J. L. Castaño, and J. Piqueras
Laboratorio de Microelectrónica, Departamento de Física Aplicada C-XII, Universidad Autónoma de Madrid, Cantoblanco 28049 Madrid, Spain

C. J. Pastor
Servicio Interdepartamental de Investigación (S.I.d.I.) C-IX, Universidad Autónoma de Madrid, Cantoblanco 28049 Madrid, Spain

(Received 15 October 1997; accepted 17 April 1998)

Phosphorus incorporation during chemical beam epitaxial (CBE) growth of $GaAs_{1-x}P_x$ from triethylgallium, tertiarybutylarsine, and tertiarybutylphosphine is investigated. Reflection high-energy electron diffraction intensity oscillations are used to measure the phosphorus incorporation during the As-limited and the (As+P)-limited growth on a Ga-rich surface. The resulting phosphorus mole fraction is compared with the phosphorus composition measured by x-ray rocking curves on the strained $GaAs_{1-x}P_x$ layers grown by conventional CBE with a simultaneous supply of group V and group III elements. The phosphorus incorporation rate during CBE growth is lower than that measured during the group V controlled growth but is still much higher than the incorporation rate reported for molecular beam epitaxy growth using elemental sources. © *1998 American Vacuum Society.* [S0734-211X(98)02104-0]

I. INTRODUCTION

Phosphorus based semiconductors, such as $In_{1-x}Ga_xAs_yP_{1-y}$, as well as the end ternary alloys $GaAs_{1-x}P_x$ and $InAs_xP_{1-x}$, have many applications in optoelectronics because their fundamental band gaps are suitable for both infrared detectors and emitters operating between 0.7 and 2.0 μm.[1]

A solid effusion cell of red phosphorus[2] was used for the growth by molecular beam epitaxy (MBE) of the first alloys of $GaAs_{1-x}P_x$. However, one of the main problems in MBE is the reproducible growth of high quality III–V alloys containing both As and P due to the difficulties in achieving a precise control over the arsenic to phosphorous flow ratio because of cell instabilities. Improvements in the flux control have been achieved by using valved cracked As and P sources[3] or thermally cracked phosphine and arsine in gas-source MBE.[4]

Despite the fact that phosphorus incorporation has been studied extensively by all these techniques, precise and reproducible compositional control still remains a difficult issue since As and P have different sticking coefficients with group III metals.[4–8] There is no simple relationship between the fluxes and the solid phase compositions of the group V elements. Additionally, the use of arsine and phosphine as group V sources causes various problems which are related to the high toxicity of the hydrides and the fact that they are pyrophoric.

In this article, we present the results of a study on the phosphorus incorporation during the chemical beam epitaxial (CBE) growth of $GaAs_{1-x}P_x$ using precracked tertiarybutylarsine (TBAs) and tertiarybutylphosphine (TBP) as group V precursors, while uncracked triethylgallium (TEGa) was used for the Ga supply. These precursors present much lower handling risks than arsine and phosphine because they are liquids with subatmospheric vapor pressures at room temperature.

II. EXPERIMENT

The growth was performed on semi-insulating GaAs(100) substrates. The details of the preparation process and of the CBE system have been published elsewhere.[9,10]

The TEGa was introduced by a cell whose temperature was kept constant at 80 °C. Both group V gases were introduced through the same high temperature cell at a temperature of 700 °C. This cell has been modified to ensure a cross talk between TBAs and TBP gas lines below 2%. Both TBAs and TBP are partly precracked at the cell temperature. The TBP decomposition process takes place in a similar way to the TBAs decomposition.[9,11] We performed some experiments using a quadrupole mass spectrometer showing the complete decomposition of TBP for cracking temperatures in the range of 750 °C.

After thermal oxide desorption in the growth chamber, a GaAs buffer layer of about 500 Å was grown on the substrate at T=550 °C using a simultaneous supply of TEGa and TBAs. Two different growth procedures were used to study the phosphorus incorporation. First, some $GaAs_{1-x}P_x$ samples were grown by classical CBE, utilizing a simultaneous supply of TEGa, TBAs, and TBP, adjusting the V/III ratio to achieve intense and long lasting reflection high-energy electron diffraction (RHEED) intensity oscillations under a 2×4 surface reconstruction.[12] The second method consisted in the growth of thin layers by group V controlled growth,[13,14] in which the TBAs+TBP flux is turned on after the deposition of the equivalent of some Ga monolayers from TEGa on the sample surface. The excess Ga that accumulated on the surface formed Ga droplets, behaving as a Ga

[a]Electronic mail: basilio.javier.garcia@uam.es

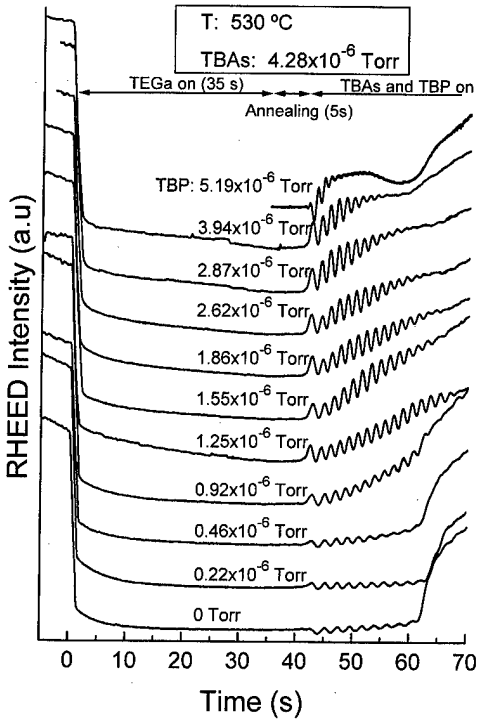


FIG. 1. RHEED intensity oscillations during the (As+P) controlled growth for increasing TBP flux and fixed TBAs flux.

source for two-dimensional growth once the V element flux is turned on.[14]

The group V controlled growth was performed after stopping the TBAs flux to eliminate the As excess over the substrate surface. The resulting surface showed a 2×4 reconstruction with an optimum RHEED specular spot, a required condition to get intense group V controlled RHEED intensity oscillations. A Ga-rich surface was then intentionally formed by supplying TEGa for 35 s. The surface reconstruction changes immediately from an As-rich 2×4 surface to a Ga-rich 4×6 reconstructed surface. After stopping the Ga supply, a 5 s pause enables TEGa decomposition on the sample surface. Then the TBAs and TBP fluxes were turned on simultaneously, so the excess Ga on the surface reacts with the As and P species impinging on the surface giving rise to a two-dimensional growth characterized by RHEED oscillations, whose number depends linearly on the amount of deposited Ga.[9,13,14]

III. RESULTS AND DISCUSSION

The RHEED intensity oscillations obtained during the group V controlled growth for different TBP fluxes and fixed TBAs flux are plotted in Fig. 1. The values of the fluxes are shown in the insets of this figure by its measured beam equivalent pressure (BEP) at a cell temperature of 100 °C, the lowest curve belonging to the pure As-controlled growth. The TEGa deposition time was kept constant at 35 s, while the Ga flux was slightly increased for higher TBP fluxes to increase the number of observed intensity oscillations. An

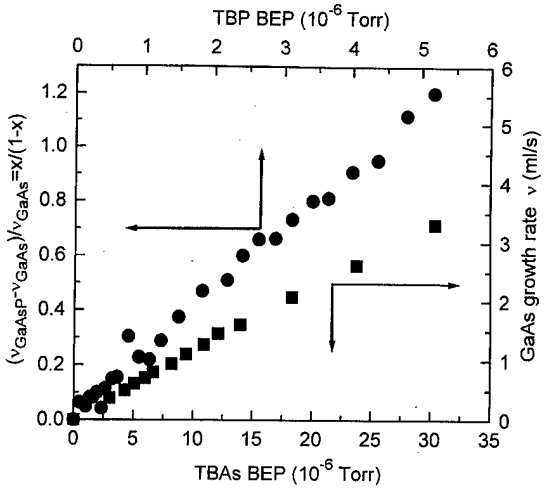


FIG. 2. Growth rate for the As-controlled growth (squares, bottom, and right axes) and the growth rate increase due to additional TBP flux (circles, left, and top axes).

increase of the oscillation frequency is observed as the TBP flux increases. The same behavior has been observed when the TBAs flux was increased.

For TBP fluxes below 0.5×10^{-6} Torr, the 4×6 reconstruction is stable until all Ga atoms are consumed, while for larger fluxes the reconstruction changes to 2×4 during the growth. This change is accompanied by an increase of both the average RHEED intensity and the oscillation amplitude, indicating that the growth is performed on a (As+P)-rich surface, despite the excess Ga on the surface. The best RHEED oscillations are obtained for a total group V BEP (the sum of the TBP and TBAs BEP) around 7×10^{-6} Torr. For higher TBP fluxes, the oscillations become more and more damped due to the group V excess on the surface, and above a total group V BEP of about 1×10^{-5} Torr no appreciable oscillations were obtained for the Ga deposition time used in the experiment. Nevertheless, with shorter Ga deposition doses, we were able to observe oscillations up to group V BEP as high as 3×10^{-5} Torr. The lower limit to observe group V induced oscillations was in our experiments a group V BEP of 2.5×10^{-6} Torr, independent of the Ga dose.

Figure 2 illustrates the growth rate, deduced from the RHEED oscillation frequency, for the As-controlled growth as a function of the TBAs BEP. As expected, a clear linear dependence can be seen, indicating that the growth rate is limited by the As supply.[9,13,14]

The growth rate increase due to an additional flux of TBP as a function of its BEP is also plotted in Fig. 2 for a fixed TBAs flux of 3.0×10^{-6} Torr, different from that used in Fig. 1. A clear linear increase in the growth rate with increasing TBP flux can be observed, also indicating that the growth rate increase is only dependent on the TBP flux. So, it seems reasonable to assume that the increase in the growth rate is only due to the addition of P atoms. Because of the excess Ga on the surface, the competition between As and P atoms incorporation into the lattice positions may be negligible. This means that the difference in the growth rates of the

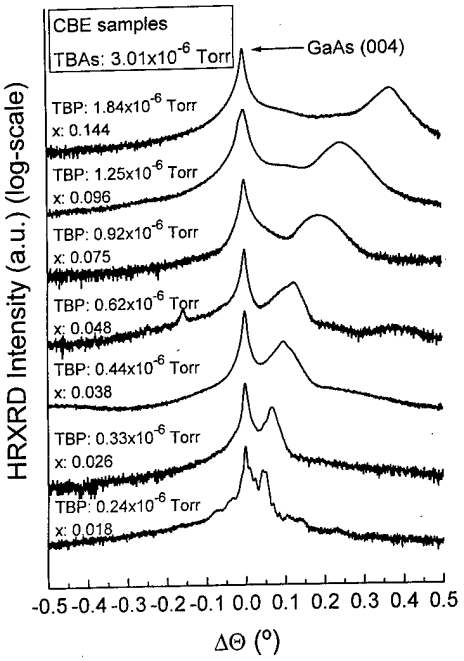


FIG. 3. HRXRD curves of the $GaAs_{1-x}P_x$ layers grown by CBE.

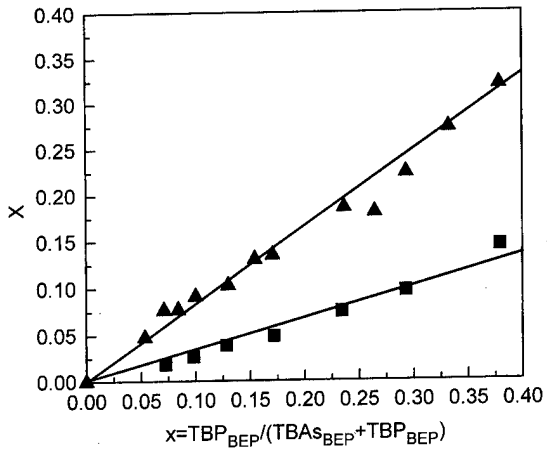


FIG. 4. Comparison of the mole fractions measured by RHEED oscillations during the (As+P)-controlled growth (triangles) and by HRXRD for the CBE grown samples (squares) as a function of the beam composition.

As-controlled and the As+P-controlled growth is a direct measurement of the net incorporation rate of phosphorus atoms and it can be taken as a measure of the phosphorus mole fraction in the grown layer.

In order to compare the incorporation of the P and As atoms on a Ga-rich surface with the incorporation rate during the CBE growth, several $GaAs_{1-x}P_x$ layers have been grown using the same substrate temperature (530 °C) and the same fixed TBAs flux of 3.0×10^{-6} Torr. The thickness of the grown layers ($d\leqslant2000$ Å) and the range of the mole fraction ($0\leqslant x\leqslant0.15$) ensured pseudomorphic growth. The composition of the grown layers was determined by high-resolution x-ray diffraction (HRXRD) measurements.

The obtained HRXRD rocking curves for layers grown by CBE with different TBP fluxes are shown in Fig. 3. For grown thickness within the critical layer thickness and assuming the validity of the Vegard's law, the structural parameters and the mole fraction of the grown layer can be calculated using the elastic stiffness constants.[15]

The phosphorus mole fractions, for both growth methods, group V controlled growth (obtained from RHEED oscillations), and CBE growth (calculated from the HRXRD), are plotted in Fig. 4 versus the phosphorus composition in the gas phase (calculated from the BEP of TBAs and TBP). It is remarkable that in both cases the P mole fraction is proportional to the TBP flow rate fraction over the whole range covered in our experiments. However, the P mole fraction is in both cases different from the TBP flow rate fraction.

This discrepancy can be attributed to two different factors. Concerning the group V controlled growth, the expected value of the slope in Fig. 4 should be close to unity. In this case, an absence of competition between the impinging As and P atoms can be assumed due to the excess of Ga atoms on the surface at growth temperatures that are low enough to avoid reevaporation losses.[16] However, the obtained value is lower than unity (0.83). This can be attributed to the fact that the As, P, or Ga fluxes cannot be directly extracted from BEP measurements because the different species involved in the decomposition in the ionization gauge may have different ionization cross sections. Because the contribution of each precursor to the measured BEP is affected by a different factor, the observed difference is not surprising.

In the case of CBE growth, the above-mentioned difference is also due to the different sticking coefficients of As and P, and to their competition to incorporate into the substrate. The value of the slope (0.34) is smaller than the value obtained for the group V controlled growth. The ratio of both slopes (2.44) is indicative of the competition of both group V species, being the competition process less important than in the case of MBE, where As atoms are 50 times more effective in displacing P than vice versa.[5]

From the results in Fig. 4 we can obtain an accurate measure of the mole fraction for the reproducible growth of phosphorus containing alloys with an accuracy of 0.02. In addition to the mentioned growth experiments, some layers were also grown using atomic layer epitaxy (ALE), showing higher incorporation rates than the CBE grown layers. The details of this work are currently under investigation.

IV. CONCLUSIONS

In summary, the (As+P)-controlled growth of $GaAs_{1-x}P_x$ has been performed, determining the margin where RHEED oscillations are observed. Growth rates and incorporation rates were measured with a very high accuracy in the pseudomorphic regime. HRXRD curves confirm the P incorporation and the possibility of the accurate mole fraction control in CBE growth.

ACKNOWLEDGMENTS

The authors would like to acknowledge T. Vallés for technical assistance. This work was supported by CICYT under the Project No. MAT 97-0291.

[1]H. C. Casey and M.B. Panish, *Heterostructure Lasers* (Academic, New York, 1978).
[2]J. R. Arthur and J. J. Lepore, J. Vac. Sci. Technol. **6**, 545 (1969).
[3]F. G. Johnson and C. E. C. Wood, J. Appl. Phys. **78**, 1664 (1995).
[4]M. B. Panish, Prog. Cryst. Growth Charact. **12**, 1 (1986).
[5]C. T. Foxon, B. A. Joyce, and M. T. Norris, J. Cryst. Growth **49**, 132 (1980).
[6]T. Fukui and N. Kokayashi, J. Cryst. Growth **71**, 9 (1985).
[7]T. Nomura, H. Ogasawara, M. Miyao, and M. Hagino, J. Cryst. Growth **111**, 61 (1991).
[8]L. Samuelson, P. Omling, and H. G. Grimmeiss, J. Cryst. Growth **61**, 425 (1983).
[9]M. Aït-Lhouss, J. L. Castaño, B. J. García, and J. Piqueras, J. Appl. Phys. **78**, 5834 (1995).
[10]M. Aït-Lhouss, J. L. Castaño, and J. Piqueras, Mater. Sci. Eng., B **28**, 155 (1994).
[11]S. H. Li, C. A. Larsen, N. I. Buchan, and G. B. Stringfellow, J. Electron. Mater. **18**, 457 (1989).
[12]L. Samuelson, B. Junno, and G. Paulsson, J. Cryst. Growth **124**, 23 (1992).
[13]F. Lewis, R. Fernandez, A. Madhukar, and F. J. Grunthaner, J. Vac. Sci. Technol. B **4**, 560 (1986).
[14]H. Neave, B. A. Joyce, and B. J. Dobson, Appl. Phys. A: Solids Surf. **34**, 179 (1984).
[15]A. Caridi and J. B. Stark, Appl. Phys. Lett. **60**, 1441 (1992).
[16]R. Fernandez, J. Cryst. Growth **116**, 98 (1992).

Effect of a deep-level trap on hole transport in $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ metal–semiconductor–metal photodetectors*

Kun-Jing Lee,[a] F. G. Johnson, and W. B. Johnson
Laboratory for Physical Sciences, College Park, Maryland 20740

Junghwan Kim and Chi H. Lee
Department of Electrical Engineering, University of Maryland, College Park, Maryland 20742

(Received 20 February 1998; accepted 29 May 1998)

The photoresponse of $In_{0.5}Ga_{0.5}As$ metal–semiconductor–metal photodetectors is related to the presence of a hole trap. Detectors made from material grown with an $In_{0.5}Al_{0.5}As$ buffer layer had no measurable trap density when examined using deep-level transient spectroscopy, and the full width half maximum (FWHM) of the photoresponse was 80 ps at 5 V bias for 3 μm interdigitated fingers and spacings. Detectors made from material grown without an $In_{0.5}Al_{0.5}As$ buffer layer had a hole trap and a FWHM photoresponse of 220 ps. This deep hole trap is likely related to impurities that diffused upward from an interface of an InP substrate and an InGaAs epilayer. © *1998 American Vacuum Society.* [S0734-211X(98)13304-8]

I. INTRODUCTION

InGaAs metal–semiconductor–metal (MSM) photodetectors with interdigitated electrodes show promise for applications in high-speed fiber-optic and atmospheric communications. However, the existence of a long tail on the falling edge of the photoresponse slows the high-speed performance of MSMs. Initially, this long tail was attributed to slow hole transport[1] due to charge pile-up at the abrupt $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ ($In_{0.5}Al_{0.5}As$ as a Schottky barrier height enhanced layer) interface. Subsequently, graded InAlGaAs layers[2] and $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ graded superlattices[3] were proposed as transition layers to improve hole transport and obtain higher bandwidths. There are also reports stating that device performance is improved by inserting an $In_{0.5}Al_{0.5}As$ buffer layer.[4] However, no systematic study of the material properties of layers grown on semi-insulating InP with and without a buffer layer, along with the associated MSM's characteristics has been reported.

In this article, the influence of buffer layers and transition layers on device performance has been explored. Hall–van der Pauw measurements and deep-level transient spectroscopy (DLTS) were used to determine the material properties of the MSMs. Epilayers grown without an $In_{0.5}Al_{0.5}As$ buffer have a higher electron mobility and a background carrier concentration that is three times larger than layers grown with a buffer layer. The photoresponse of detectors with graded or step transition layers is identical at high applied bias, but the detector grown without an $In_{0.5}Al_{0.5}As$ buffer layer has a hole trap which slows the response.

II. EXPERIMENTS

The materials were grown using solid-source molecular beam epitaxy (MBE) on (100), Fe-doped, semi-insulating InP misoriented 4° toward (111)A, and the growth temperature was 485 °C. Two of the structures consist of a 20 nm $In_{0.5}Al_{0.5}As$ ($E_g = 1.48$ eV) buffer layer, a 1 μm $In_{0.5}As_{0.5}As$ ($E_g = 0.79$ eV) absorption layer, and either a 20 nm $In_{0.5}Al_xGa_{0.5-x}As$ graded layer (structure "*A*") or an $InP(E_g = 1.35\text{ eV})/In_{0.73}Ga_{0.27}As_{0.57}P_{0.43}$ ($E_g = 0.954$ eV) step layer (structure "*B*"), then finally a 20 nm $In_{0.5}Al_{0.5}As$ ($E_g = 1.48$ eV) barrier-height-enhanced layer on the top. The third structure (structure "*C*") is identical to structure *B* except that it was grown without the $In_{0.5}Al_{0.5}As$ buffer layer. Ti–Au metallization was followed by a standard lift-off process (chlorobenzene-soaked positive photoresist) to delineate 300 nm thick, 3 μm wide, interdigitated fingers where the finger width and spacing are identical. The area covered by interdigitated fingers was approximately 100 μm×200 μm. Coplanar waveguides with ground plane were designed to exhibit a 50 Ω characteristic impedance. The interdigitated fingers were arranged on the center of the signal line of the coplanar waveguide.

The background carrier concentrations and mobilities were measured by applying the Hall–van der Pauw method to these three structures. For DLTS and net capacitance measurements, two ports from the DLTS equipment were connected to the two ends of the signal line from the coplanar waveguides of the MSMs. The reverse bias of the DLTS measurements was carefully chosen so as not to deplete the MSM completely. The temperature scan range was varied from 30 to 380 K. Data were collected by a computer through a high speed analog–digital converter board. The photoresponse measurements were collected with an HP sampling oscilloscope with a 20 G Hz bandwidth, and the MSM was excited with an ultrashort, 200 fs pulse at a wavelength of 800 nm, which was extracted from a Ti-sapphire mode-locked laser pumped by an Ar-ion laser.

III. RESULTS AND DISCUSSIONS

Table I lists the material properties and the MSM device performance for the three different epitaxial structures. The background carrier concentration of the unintentionally

*No proof corrections received from author prior to publication.
[a]Electronic mail: kunjing@eng.umd.edu

TABLE I. Comparison of material properties and device performance for MSM photodetectors made from structures *A*, *B*, and *C*.

Structure	*A*	*B*	*C*
InAlAs buffer	Yes	Yes	No
Transition layer	InAlGaAs graded	InP/InGaAsP step	InP/InGaAsP step
Carrier concentration	3.25×10^{15} cm^{-3}	2.61×10^{15} cm^{-3}	8.04×10^{15} cm^{-3}
Mobility (cm^2/V s)	1785	2173	5307
Net capacitance	7 pF	4.3 pF	27 pF
Deep-level trap	No	No	Hole trap
Dark current (uA) at 3 V	11	8	200–300
Photoresponse (5 V)	85 ps (FWHM)	74 ps (FWHM)	220 ps (FWHM)

doped $In_{0.5}Ga_{0.5}As$ layer grown on the $In_{0.5}Al_{0.5}As$ buffer layer (structures *A* and *B*) is about 3×10^{15} cm^{-3}, which is similar to data reported by others for MBE,[5] liquid phase epitaxy (LPE),[6] and chemical beam epitaxy (CBE)[7] material. But, the carrier concentration measured for structure *C* is almost three times larger (about 8.04×10^{15} cm^{-3}). One possible explanation for this increase in the carrier concentration is upward diffusion of impurities from the wafer during growth.[8,9] Incorporating a high-band-gap material as a buffer layer is believed to reduce this impurity diffusion.

At the same bias, the detectors with higher background concentrations had smaller depletion widths. Consequently, the net capacitance, which is inversely proportional to the depletion width, is larger for MSMs made from structure *C* than either *A* or *B*. This increase in the carrier concentration and the net capacitance are related to the absence of an $In_{0.5}Al_{0.5}As$ buffer layer and are consistent with previous reports in the literature.[4] Note, however, that the electron mobilities for structures *A* and *B* are lower than those measured for structure *C*. The photoresponse of all three MSM structures, however, is hole transit-time limited.[10] Therefore, the lower electron mobility did not affect the bandwidth of the MSMs from structures *A* and *B*.

Figure 1 shows the DLTS spectra from an MSM made from structure *C* with a varying initial delay (initial delay $=t_1$; $t_2/t_1=10$). One hole trap with a trap density of 6.02×10^{14} cm^{-3} was found, whereas, no traps were detected for detectors made from either *A* or *B*. The difference between structures *A* and *B* as compared to *C* is the presence of the $In_{0.5}Al_{0.5}As$ buffer layer. Figure 2 shows an Arrhenius plot of $\ln(e_{p(\max)}/T^2)$ vs $(1/T)$ for structure *C* where

$$(e_p)_{\max}=\frac{\ln(t_2/t_1)}{t_2-t_1}=BT_m^2\sigma_p\exp\left(-\frac{(E_T-E_V)}{KT_m}\right). \quad (1)$$

Here, e_p represents the hole emission rate, B the constant, T_m the temperature at which the maximum of DLTS signal occurs, σ_P the hole capture cross section, E_V the top of the valence band, E_T the trap energy level, and K the Boltzmann constant. An activation energy of 0.475 eV with respect to the valence band and a hole capture cross section of 1.11×10^{-15} cm^2 were determined from the slope and intercept shown in Fig. 2. Assuming $m_{\rm ph}$ (the effective mass of a heavy hole) of $0.61m_0$ and $m_{\rm pl}$ (the effective mass of a light hole) of $0.0588m_0$,[11] where m_0 is the electron rest mass, N_V (the effective density of states in the valence band) and $v_{\rm th}$ (the thermal velocity of the hole at 300 K) were obtained as 1.23×10^{19} cm^{-3} and 3.4×10^6 cm/s, respectively. The carrier (hole) lifetime is 2 ms from the following equation:

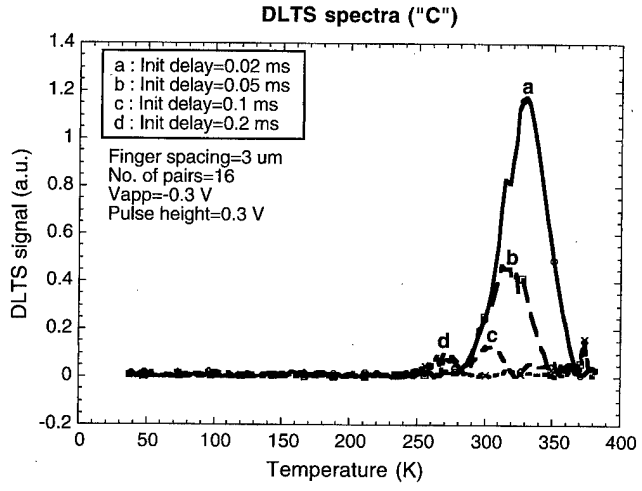


FIG. 1. DLTS spectra from MSMs made from structure *C* with a varying initial delay (initial delay $=t_1$; $t_2/t_1=10$). (Only one hole trap was found.)

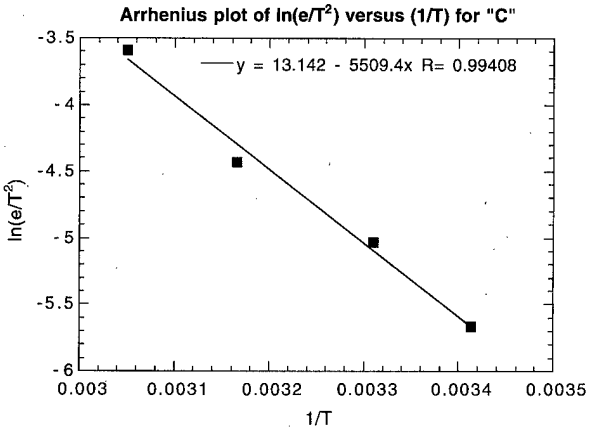


FIG. 2. Arrhenius plot of $\ln(e_p/T^2)$ vs $(1/T)$ for structure *C*.

$$\tau = \frac{1}{e_P} = \frac{\exp[(E_T - E_V)/KT]}{v_{\text{th}} \sigma_P N_V}, \tag{2}$$

where τ is the carrier lifetime. Note that the mechanism of thermally stimulated measurements such as DLTS is similar to that of carriers escaping from deep levels due to the electric field of the photoresponse measurement, especially since both measurements were performed on the same devices. A carrier (hole) lifetime of 2 ms can significantly affect the high-speed response of an MSM.

Without taking into account shielding effects from charge accumulation at the heterojunction, the Schottky barrier height (0.7 eV for Au on $In_{0.5}Al_{0.5}As$)[12] and a reverse bias between 0 and 1 V will create a depletion length in material with a carrier concentration of 8×10^{15} cm^{-3} from 0.37 to 0.57 μm based on Eq. (3):

$$d = \sqrt{\frac{2\epsilon_0 \epsilon_r (V_{\text{bi}} - V_{\text{ap}})}{qN}}, \tag{3}$$

where ϵ_0 is the permittivity of free space, ϵ_r the relative permittivity of the semiconductor, V_{bi} the Schottky barrier height, and N the carrier concentration. Therefore, by changing the reverse bias and keeping the same positive pulse width during the DLTS measurement, a volume consisting of the top 570 nm of the $In_{0.5}Ga_{0.5}As$ layer was profiled. The hole trap appeared at all planes within this region. The origin of the hole trap is still not clear. Because no trap was found in structures *A* or *B*, the hole trap must be related to the absence of the $In_{0.5}Al_{0.5}As$ buffer layer. Defects resulting from impurities diffusing from the substrate–epilayer interface are one of the most likely explanations. In the literature, an electron trap with an activation energy of 0.4 eV in MBE grown $In_{0.5}Ga_{0.5}As$ has been reported,[5] and oxygen-related electron traps with energies ranging from 0.5 to 0.7 eV in $In_{0.5}Al_{0.5}As$ have been found.[13] In this study, the DLTS was performed on real MSM devices grown on Fe-doped InP substrates. This differs from reports in the literature where the same epilayers were grown on *n*-type InP substrates and used as a reference for the material properties of the MSMs. It is likely, however, that the material properties will be affected by the choice of substrate (Fe doped vs S doped). In our case, we have measured the material properties of the actual MSM layers grown on semi-insulating (Fe doped) InP and this may explain why this is the first report of a hole trap seen in these materials.

Figure 3(a) shows the photoresponse at 3 V bias for the three MSM structures. The photoresponse for structures *A* and *B* is nearly identical, which shows that the abrupt heterojunction present in *B* does not affect the response. This is due to the reduction in the band-gap discontinuities at the heterojunctions by incorporating a two-step transition region as compared to the $In_{0.5}Ga_{0.5}As/In_{0.5}Al_{0.5}As$ abrupt heterojunction adopted by other researchers. Device *C*, however, exhibits a significant tail on the falling edge of the photoresponse. We attribute this slow response to the deep-level hole trap that appears in MSM structures grown without the thin $In_{0.5}Al_{0.5}As$ buffer layer. An alternative argument is that

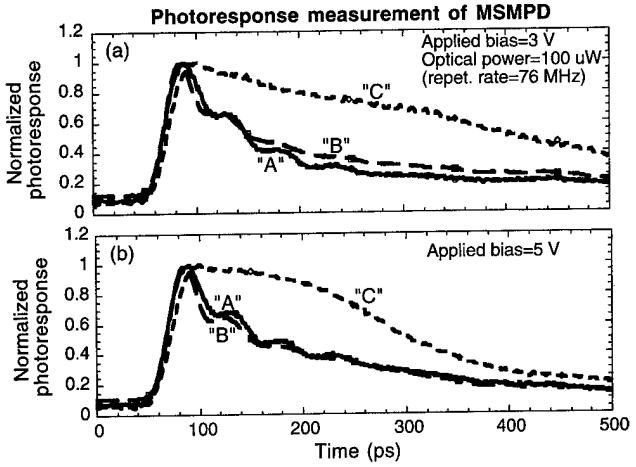


FIG. 3. Photoresponse of MSM photodetectors for *A*, *B*, and *C* at (a) 3 V and (b) 5 V.

structure *C*, which possesses a higher background carrier concentration, has a smaller depletion region under the same bias than either *A* or *B*. Therefore, a slower response might be due to slower diffusion of free electrons and holes excited in the nondepleted region. Figure 3(b) shows the photoresponse measurement at 5 V bias. Compared with Fig. 3(a), the responses of *A* and *B* are still similar but *C* in (b) shows a faster response than *C* in (a). It indicates one or both of the following: (a) the nondepleted region was decreased with increased bias leading to fewer carriers excited in a region of low electric field, (b) the probability of carriers being trapped in deep levels was decreased with the increased kinetic energy of the holes associated with the increasing bias. Structure *C* has the same structure as *B* except for the buffer layer. After taking the carrier concentration differences (the carrier concentration of *C* is nearly three time larger than *B*) into account, the FWHM of *C* at 9 V was found to still be twice that of *B* at 3 V. Therefore, the slow response of *C* cannot simply be explained by excited-carrier diffusion in low electric-field regions, and the hole trap, related to the absence of the $In_{0.5}Al_{0.5}As$ buffer layer, must play a major role in the degradation of the photoresponse.

IV. CONCLUSIONS

The influence of buffer layers and transition layers on the material properties and device performance of MSMs has been explored. The epilayers grown without an $In_{0.5}Al_{0.5}As$ buffer layer have a larger background carrier concentration and a higher electron mobility as compared to layers grown with the buffer layer. The photoresponse of these MSM photodetectors is related to the presence of a hole trap. A detector (3 μm interdigitated fingers width and spacing) made from material grown with an $In_{0.5}Al_{0.5}As$ buffer layer had no measurable hole trap density and a FWHM photoresponse of about 80 ps. But, detectors from material grown without a buffer layer had a hole trap density of 6.02×10^{14} cm^{-3} and a FWHM photoresponse of 220 ps at 5 V bias. This slow

response cannot be simply explained by a diffusion of free electrons and holes excited in nondepleted regions, and the hole trap, related to the absence of the buffer layer, plays a major role in the degradation of the photoresponse.

[1]W. C. Koscielniak, J. L. Pelouard, and M. A. Littlejohn, Appl. Phys. Lett. **54**, 567 (1989).
[2]H. T. Griem, S. Ray, J. L. Freeman, and D. L. West, Appl. Phys. Lett. **56**, 1067 (1990).
[3]O. Wada, H. Nobuhara, H. Hamaguchi, T. Mikawa, A. Tackeuchi, and T. Fujii, Appl. Phys. Lett. **54**, 16 (1989).
[4]S. E. Ralph, M. C. Hargis, and G. D. Pettit, Appl. Phys. Lett. **61**, 2222 (1992).
[5]P. S. Whitney, W. Lee, and C. G. Fonstad, J. Vac. Sci. Technol. B **5**, 796 (1987).
[6]Y. Nashimoto, S. Dhar, W. P. Hong, A. Chin, P. Berger, and P. K. Bhattacharya, J. Vac. Sci. Technol. B **4**, 540 (1986).
[7]N. Debbar, A. Rudra, J.-F. Cardin, and M. Ilegems, Appl. Phys. Lett. **65**, 228 (1994).
[8]H. Roehle, H. Schroeter-Janssen, P. Harde, and D. Franke, J. Electron. Mater. **24**, 1535 (1995).
[9]S. P. Watkins, H. D. Cheung, G. Knight, and G. Kelly, Appl. Phys. Lett. **68**, 1960 (1996).
[10]M. Hargis, S. E. Ralph, J. Woodall, and D. McInturff, Appl. Phys. Lett. **67**, 413 (1995).
[11]K. Brennan, IEEE Trans. Electron Devices **33**, 1502 (1986).
[12]W. Gao, P. R. Berger, R. G. Hunsperger, G. Zydzik, W. W. Rhodes, H. M. O'Bryan, D. Sivco, and A. Y. Cho, Appl. Phys. Lett. **66**, 3471 (1995).
[13]A. Naritsuka, T. Noda, A. Wagai, S. Fujita, and Y. Ashizawa, Jpn. J. Appl. Phys., Part 2 **32**, L925 (1993).

Effect of the forward biasing the source-substrate junction in *n*-metal–oxide–semiconductor transistors for possible low power complementary metal–oxide–semiconductor integrated circuits' applications

F. J. De la Hidalga-W.[a)]
School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 and National Institute of Astrophysics, Optics and Electronics, INAOE, P.O. Box 216&51, Z.P. 72 000 Puebla, Mexico

M. J. Deen[b)]
School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6

E. A. Gutierrez-D.
National Institute of Astrophysics, Optics and Electronics, INAOE. P.O. Box 216&51, Z.P. 72 000 Puebla, Mexico

F. Balestra
Laboratoire de Physique des Composants a Semiconducteurs, ENSERG-INPG, BP257, 38016 Grenoble, France

(Received 17 April 1998; accepted 22 April 1998)

The use of forward biasing the source-substrate junction in metal–oxide–semiconductor transistors, to reduce its threshold voltage is a simple method to realize low power complementary metal–oxide–semiconductor integrated circuits. A 2 μm long *n*-metal–oxide–semiconductor field effect transistor was used as the device under test. The threshold voltage was measured in the temperature range of 300–77 K, using two different methods. The classic long channel threshold voltage model was fitted by optimization to the experimental data of the reverse-biased substrate, and the model with the fitted parameters was used to calculate the threshold voltage under forward-biased substrate conditions. The agreement between the fitted and extrapolated threshold voltage with the experimental values demonstrated the validity of this classic model for a substrate forward bias up to 0.5 V, and for a wide temperature range. © *1998 American Vacuum Society.* [S0734-211X(98)01904-0]

I. INTRODUCTION

The main contribution to the power consumption in complementary metal–oxide–semiconductor (CMOS) logic gates is proportional to the square of the power supply, thus the lowering of the power supply has been proposed as an alternative to realize low power CMOS integrated circuits (ICs). Nevertheless, this reduction can produce an increase in the delay time per gate. In order to avoid this undesirable increase, several approaches have been proposed. Among them is the threshold voltage scaling approach.[1] Generally, this goal is achieved by technological changes due to the dimensional scaling of the metal–organic–semiconductor field effect transistor (MOSFET), but it usually results in a degradation of the subthreshold characteristics, with a higher subthreshold static current, and a degradation of the noise margin.[1]

By using the *n*-MOSFET connected in such a way that the substrate-source voltage, V_{BS}, follows the gate-source voltage, V_{GS}, it is possible to obtain a device with a dynamic threshold voltage, V_{TH}.[2–7] According to the well known body effect, V_{TH} increases for a higher reverse V_{BS}, while a decrease is expected for an increasing forward V_{BS}. This means that in order to switch on the gate, a lower input voltage will be needed, while the off condition can be achieved with just a slight reduction of V_{GS}, as the threshold would have been shifted, avoiding the subthreshold leakage and the degradation of the noise margin.

In this approach, however, the MOSFET operates with a forward biased source-substrate (S-B) junction, with the consequent injection of mobile carriers through the depletion region of this junction. This produces an increase of the leakage current of this junction which eventually may result in the loss of electric isolation between devices by turning on the parasitic bipolar transistor inherent to any MOS transistor and whose emitter-base junction is the S-B junction.[4,7,8] This is why the forward biasing of the S-B junction has never been used in standard bulk CMOS circuits operating at room temperature. Nevertheless, it has found applications in low temperature bulk CMOS,[9,10] where a high V_{TH} is obtained due to both the low operating temperature, and the high substrate concentration characteristic of the modern MOS technology. Some attempts have been made to adapt both bulk CMOS,[2] silicon on insulator SOI-CMOS[3] and BiCMOS[11,12] technologies to support MOSFETs working as dynamic threshold devices.

[a)]Electronic mail: jhidalga@inaoep.mx
[b)]Electronic mail: jamal@cs.sfu.ca

The high quality and high doping concentration of modern substrates suitable for submicron devices, produce *S-B* junctions with a very high turn-on voltage, typically around 0.9 V. So it is thought that by keeping the *S-B* forward bias well below that value ($V_{BS}<0.5$ V), the number of mobile carriers injected will not be too large, and so will not have an appreciable effect on the electrical characteristics of the MOSFET.

As a contribution to this threshold scaling approach, we investigated the behavior of the *n*-MOS threshold voltage under a forward-biased *S-B* junction over a wide temperature range, from 300 to 77 K. First, we discuss the V_{TH} model to be considered in this temperature range. Next, the experimental procedure followed to obtain V_{TH} at two different surface potential conditions is presented. The well known behavior of V_{TH} for $V_{BS}<0$ was used to obtain, the effective value of the parameters of the V_{TH} model from an optimization routine. Finally, the V_{TH} determined with the classic model is compared to the experimental data for $V_{BS}>0$.

II. THEORETICAL BACKGROUND

A. *n*-MOS long channel threshold voltage

The 2 μm channel length device used in this study is from a 0.1 μm CMOS technology. The V_{TH} is adjusted by a boron ion implantation giving a surface concentration of $\sim 1.7\times10^{17}$ cm^{-3} and a peak concentration of $\sim 4\times10^{17}$ cm^{-3} at 155 nm from the surface. The nominal substrate concentration, obtained from a box approximation, is $\sim 3\times10^{17}$ cm^{-3}. This effective concentration is high enough to prevent the short channel effect. From our optimization described later, N_A values were in the $(2.4–2.7)\times10^{17}$ cm^{-3} range, very close to the nominal substrate concentration obtained from the box approximation. Furthermore, in order to avoid the drain-induced barrier lowering (DIBL),[13–15] the drain voltage used was 50 mV. The channel width was 14 μm, so narrow channel effects were not considered.

As can be easily found elsewhere in the literature,[16] the long channel threshold voltage can be expressed as

$$V_{TH}(V_{BS})=V_{T0}+\gamma(\sqrt{\phi_B-V_{BS}}-\sqrt{\phi_B}), \tag{1}$$

where γ, the body effect parameter, is given by

$$\gamma=\frac{\sqrt{2q\epsilon_s N_A}}{C_{ox}}. \tag{2}$$

The effective substrate doping concentration is given by N_A, ϵ_s represents the dielectric permittivity of silicon, and C_{ox} is the oxide capacitance per unit area, determined by the oxide permittivity ϵ_{ox} and the oxide thickness t_{ox} and is

$$C_{ox}=\frac{\epsilon_{ox}}{t_{ox}}. \tag{3}$$

V_{T0} is defined as V_{TH} evaluated at $V_{BS}=0$, and is given by

$$V_{T0}=V_{FB}+\phi_B+\gamma\sqrt{\phi_B}. \tag{4}$$

The total surface band bending, ϕ_B, is strongly dependent on V_{GS},[9] and it is usually defined as

$$\phi_B=2\phi_F+\alpha\phi_t \tag{5}$$

with ϕ_t being the thermal voltage (kT/q), α an empirical fitting parameter, and ϕ_F the Fermi potential in the substrate, given by

$$\phi_F=\phi_t \ln\left(\frac{P}{n_i}\right). \tag{6}$$

In Eq. (6), P represents the majority carrier concentration in the bulk and n_i is the intrinsic concentration in silicon. The flatband voltage which appears in Eq. (4) is determined by the effective oxide charge density Q_0 and the difference in work functions between the gate (n^+ polysilicon) and the silicon, and is

$$V_{FB}=-\phi_F-\phi_g-\frac{Q_0}{C_{ox}}. \tag{7}$$

The V_{TH} expressed by Eq. (1) was derived from a charge balance equation using both the depletion and charge sheet approximations. These approximations might be violated by the presence of mobile carriers injected as a consequence of the forward V_{BS} only if their concentration is as high as the doping level, which determine the amount of charge in the depletion region.

B. Low temperature behavior of the threshold voltage

It is well known that the increment of V_{TH} with the lowering of temperature and the simple model of Eq. (1) is good enough to explain such dependence.[17] Nevertheless, care has to be taken in correctly interpreting the temperature dependence of the terms involved in Eq. (1). The only parameter whose temperature dependence is not well understood is the gate contact potential ϕ_g. Some authors[17] assume that the commonly used n^+-doped polysilicon can be modeled as *n*-type degenerate silicon, so that the contact potential is equal to half of the silicon band gap. Even though the substrate suffers from freeze-out at low temperatures, and the majority carrier concentration P is very different from the doping level, the electric field in the depletion region is high enough to ionize the impurities in that region.[18] Then the term N_A, expressing the average charge concentration in the depletion region and which appears in Eq. (2), is given by the average concentration of dopants. This term varies with temperature when the substrate doping is not uniform.[19] In the neutral substrate where there is no electric field to assist the thermal ionization of impurities, the concentration of ionized impurities can be found using Fermi–Dirac statistics. By equating the concentration of ionized impurities with the hole concentration in the substrate at thermal equilibrium, the Fermi level can be determined as well as the Fermi potential, ϕ_F, at each temperature (T). So the variation of V_{TH} with T is mainly due to the variation of the Fermi potential in the neutral substrate and the slight widening of the depletion region. The equation[15] to be solved numerically is

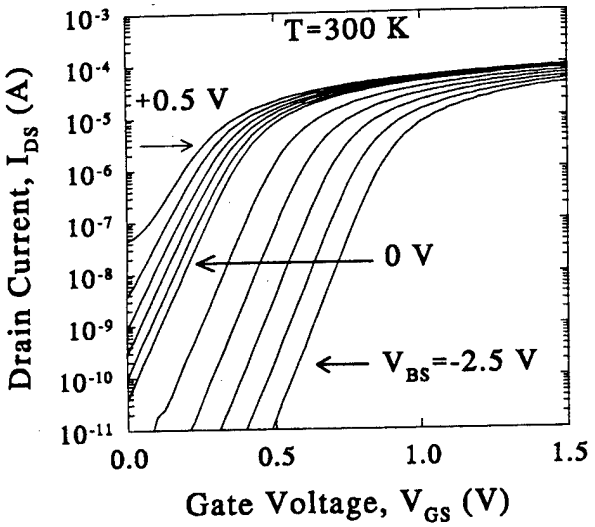


FIG. 1. Experimental drain current vs gate voltage characteristics as a function of substrate bias at room temperature. Steps of 0.5 and 0.1 V were used for $V_{BS}<0$ and $V_{BS}>0$, respectively.

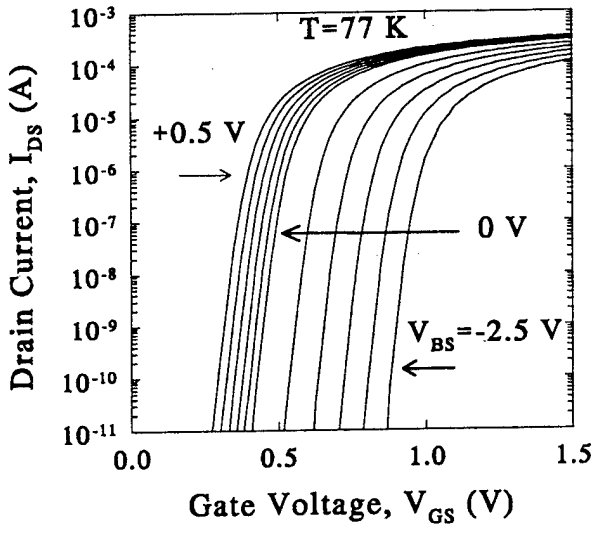


FIG. 2. Experimental drain current vs gate voltage characteristics as a function of substrate bias at 77 K. Steps of 0.5 and 0.1 V were used for $V_{BS}<0$ and $V_{BS}>0$, respectively.

$$\exp\left(\frac{E_v-E_F}{kT}\right)\left[1+4\exp\left(\frac{E_a}{kT}\right)\exp\left(\frac{E_v-E_F}{kT}\right)\right]=\frac{N_A}{N_v}, \quad (8)$$

where N_v is the effective density of states in the valence band and E_a is the acceptor energy level measured from the valence band. E_a is equal to 0.045 eV for boron impurities in silicon. The temperature dependence of N_v and n_i is well known,[20] where for the latter, the effect of the band-gap widening at low temperature has to be included. Once the Fermi level relative to the top of the valence band is known, the Fermi potential can be obtained by using

$$\phi_F=\frac{(E_v-E_F)}{q}+\phi_t \ln\left(\frac{N_v}{n_i}\right). \quad (9)$$

III. EXPERIMENTAL DETAILS

The 2 μm long and 14 μm wide n-MOS transistor was mounted in the cold-head of a closed cycle refrigerator and cooled to 77 K. Measurements were done at temperatures between 300 and 77 K. To obtain the n-MOS drain current–gate voltage characteristics, $I_{DS}-V_{GS}$, the drain voltage, V_{DS}, was kept constant at 50 mV while V_{GS} was swept from 0 to 1.5 V using a 25 mV step; this procedure was repeated for several forward and reverse substrate voltages, where $-2.5\ \text{V}<V_{BS}<0.5\ \text{V}$, and the source was used as the reference terminal; steps of 0.5 and 0.1 V were used for $V_{BS}<0$ and $V_{BS}>0$, respectively. The dc characteristics were measured with the HP 4145B semiconductor parameter analyzer.

The measured n-MOS $I_{DS}-V_{GS}$ characteristics are shown in Fig. 1 at room temperature at different V_{BS} biases. The effect of forward biasing the substrate can be clearly seen; as V_{BS} becomes more positive, the subthreshold characteristic shifts to the left. Here, we see the variation of the subthreshold slope and the appearance of a parasitic current for the more positive V_{BS} biases, and this is definitely altering the subthreshold characteristics. The parasitic current is the consequence of the pre-turn on the parasitic bipolar transistor. This effect disappears completely when the temperature is lowered, as can be seen in Fig. 2, where similar data is presented, but at 77 K. In this case, the subthreshold slope does not show any degradation, and the characteristics exhibits the typical steeper behavior[21] expected at low temperatures.

The threshold voltage was extracted from the experimental data using two different methods. This was because of the largely known ambiguity involved in the definition of the total band bending at the surface, and which is contained in the empirical parameter α in Eq. (5). Both methods are described briefly below.

A. Quasi-constant current method

This method is suitable to obtain V_{TH} with no ambiguity because it corresponds to the case of $\alpha=0$ in Eq. (5),[22,23] which is a common assumption even though the surface potential can be very different from the assumed pinned value of $2\phi_F$.[11] For submicron MOS transistors, where the threshold voltage is several tenths of a volt, this discrepancy can produce serious inaccuracies. Despite the unambiguous determination of V_{TH} by this method, its use presents a couple of disadvantages. It requires more calculations and V_{TH} has to be extracted from the subthreshold characteristics in the weak inversion region and up to the onset of moderate inversion. By making a linear regression to the linear portion of the subthreshold characteristics of Figs. 1 and 2, the slope (m) and the intersection (b) with the $\ln(I_{DS})$ axis (with I_{DS} normalized to 1 A) of this fitted line are obtained, and then V_{TH} can be calculated from the following equation:[22,23]

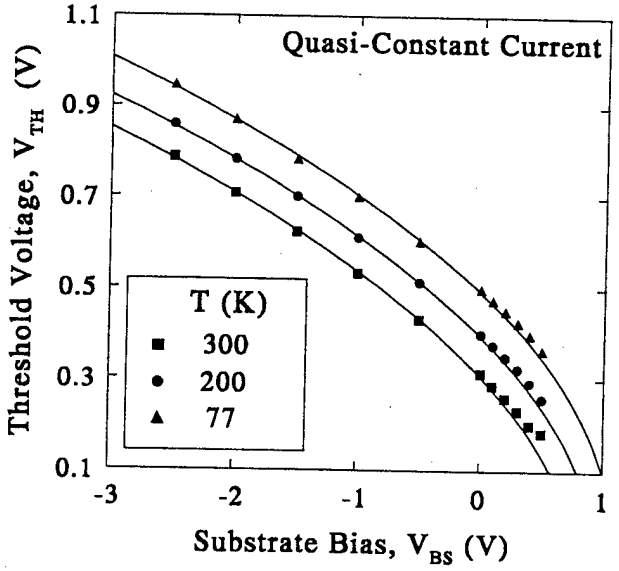


FIG. 3. Quasi-constant current experimental (symbols), fitted ($V_{BS}<0$) and extrapolated ($V_{BS}>0$) (lines) threshold voltage as a function of substrate bias. The results are shown for three different temperatures, from 300 to 77 K.

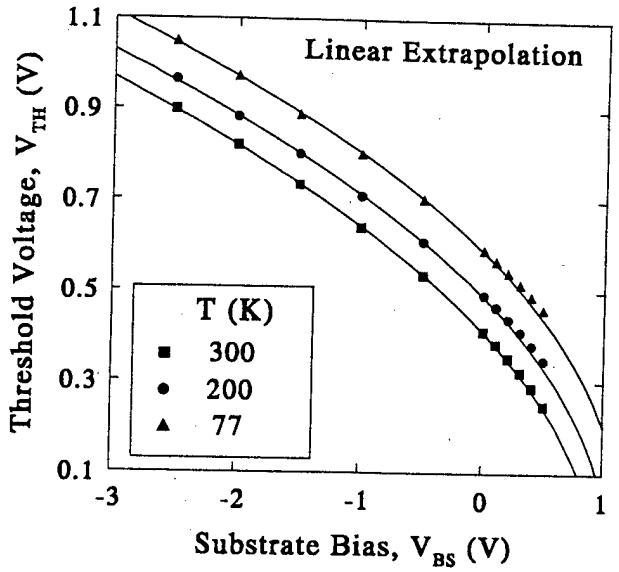


FIG. 4. Linear extrapolation experimental (symbols), fitted ($V_{BS}<0$) and extrapolated ($V_{BS}>0$) (lines) threshold voltage as a function of substrate bias. The results are shown for three different temperatures, from 300 to 77 K.

$$V_{TH}(V_{BS})=\frac{\ln\left\{\beta_{eff}\phi_t^2\left(\frac{1}{\phi_t m}-1\right)\left[1-\exp\left(\frac{-V_{DS}}{\phi_t}\right)\right]\right\}-b}{m}, \tag{10}$$

where β_{eff} is the effective conductance coefficient ($\mu C_{ox}W/L$) obtained by conventional methods from the linear operating region of the MOS transistor. The experimental V_{TH} obtained with this method is presented in Fig. 3 (symbols) as a function of V_{BS}, for three different temperatures.

B. Linear extrapolation method

This method is widely known, and it is based on the linear behavior of the $I_{DS}-V_{GS}$ characteristics when $V_{DS}<0.1$ V (in our case, 50 mV), V_{GS} is higher than V_{TH} but not too high to degrade the channel mobility (in such a case, the linear approximation is violated). The intersection with the V_{GS} axis of the linear extrapolation of the $I_{DS}-V_{GS}$ characteristic, gives the experimental V_{TH}. The experimental V_{TH} obtained by using this method is shown in Fig. 4 (symbols) as a function of V_{BS}, for three different temperatures. Since in this case the device is operating in strong inversion, the surface potential will be higher than in the earlier case, now $\alpha>0$, and the extracted V_{TH} will also be higher at any temperature, as can be observed by comparing Figs. 3 and 4.

IV. OPTIMIZATION RESULTS AND DISCUSSION

In order to investigate the validity of the V_{TH} model of Eq. (1) for a forward biased substrate, this model was fitted to both sets of experimental V_{TH} data only for the $V_{BS}<0$ case, where the validity of Eq. (1) is widely accepted. From the room temperature data set, t_{ox}, N_A, V_{T0}, and α were found by using an optimization routine to solve the system of Eqs. (1)–(9) and the experimental V_{TH}. The temperature dependence of these expressions was included at room temperature [specifically Eqs. (8) and (9)] because even at 300 K there is incomplete ionization of boron impurities in silicon, and assuming the contrary, would have resulted in inconsistencies at the lower temperatures. The simplified model of Eq. (1), without considering specifically V_{FB} [given by Eq. (7)] was used because of the uncertainty of the temperature dependence of ϕ_g, as was already explained, so that this dependence is contained in the temperature variations of V_{T0}.

From the extraction procedure described above, the oxide thickness was 6.3 nm. While the optimum oxide thickness found at 300 K was kept constant for optimization at lower temperatures, the N_A value was allowed to vary as one of the unknowns, because it could vary with temperature as the doping is not uniform. The empirical parameter α was found to be very close to zero for the quasi-constant current method V_{TH}, while its value was very close to 5, though slightly different at each temperature, for the linear extrapolation method. Once the optimum N_A, α, and ϕ_F were obtained, both γ and ϕ_B were calculated at each temperature. The optimization results for the quasi-constant current and the linear extrapolation methods are presented in Tables I and II,

TABLE I. Optimization results for the quasi-constant current method.

T (K)	ϕ_B (V)	V_{T0} (V)	$\gamma(V^{1/2})$
300	0.847	0.308	0.522
250	0.910	0.352	0.521
200	0.969	0.395	0.523
150	1.021	0.436	0.522
125	1.044	0.457	0.525
100	1.065	0.476	0.520
77	1.083	0.492	0.526

TABLE II. Optimization results for the linear extrapolation method.

T (K)	ϕ_B (V)	V_{T0} (V)	$\gamma(V^{1/2})$
300	0.988	0.409	0.558
250	1.025	0.450	0.555
200	1.059	0.488	0.550
150	1.089	0.524	0.542
125	1.102	0.558	0.536
100	1.115	0.574	0.538
77	1.125	0.588	0.542

respectively.

As expected, both ϕ_B and V_{T0} are higher at lower temperatures. γ showed a slight variation with temperature, but there was no definite trend. This is just the result of the optimization procedure and experimental errors. Nevertheless, there is a small ($<7\%$) and systematic difference in γ between both of the methods used. Although this is not well understood yet, it could be due to the difference in the depletion region width as a result of the difference in the surface potential in each method. Therefore, the slight variation of the substrate doping will give different values of the effective doping concentration, whose value is obtained by averaging the doping concentration along the depletion region.

By using the parameters shown in Tables I and II, which were obtained by fitting Eq. (1) to the experimental V_{TH} for the $V_{BS}<0$ biases only, the fitted ($V_{BS}<0$) and extrapolated ($V_{BS}>0$) values of V_{TH} for both methods are shown in Figs. 3 and 4 (continuous lines). The good agreement between the experimental and fitted V_{TH} for $V_{BS}<0$ is a result of the fitting procedure. There is remarkably good agreement between the extrapolated value of V_{TH} and the experimental one for $V_{BS}>0$, over the whole range of temperatures considered in this study.

There is, however, some disagreement between the experimental and fitted room temperature quasi-constant current V_{TH} values for most of the positive V_{BS} biases. This is a consequence of the extraction procedure and not the invalidity of the model. To understand this, we observe the room temperature I_{DS} subthreshold behavior in Fig. 1. As V_{BS} becomes more positive, the subthreshold slope decreases while the current level increases. As was explained above, this is because there is an enhancement of diffusion of mobile carriers due to the reduction of the potential barrier in the S-B junction. Now, the parasitic bipolar is being set in the pre-turn on condition, and the drain, working as the parasitic collector, will conduct a higher current. This phenomenon is dominant for low values of V_{GS}, when the silicon surface is in weak inversion.

As V_{GS} increases, the inversion layer which is being formed is electrically connected to the source, and will operate as a field-induced junction, producing an increment in the leakage current of the substrate. This component also depends on V_{GS}, altering the subthreshold slope. As the temperature is lowered, however, these parasitic contributions become negligible as the injection level is reduced. This can also be appreciated in Figs. 1 and 3, where the lower degradation in the subthreshold characteristics corresponds to a better agreement between the fitted and experimental quasi-constant current V_{TH}. Furthermore, since this parasitic current is negligible in comparison to the drain current in the linear region, it does not alter the extraction procedure of V_{TH} by using linear extrapolation, and the agreement with the model of Eq. (1) is acceptable even at room temperature.

If this low forward biasing were affecting the charge sheet and depletion approximations by the injection of majority carriers, the expected effect would be the reduction of the effective depletion charge;[24] in such a case, the effective γ parameter would be lower at $V_{BS}>0$ than in the corresponding reverse bias case. This would result in an overestimation of V_{TH} by the model. As can be seen in Figs. 3 and 4, there is no such overestimation, demonstrating that the effect of the injected majority carriers is negligible at low forward biases of the substrate.

V. CONCLUSIONS

We have presented the first experimental results, over a wide temperature range, of the effect of forward biasing the source-substrate junction on the threshold voltage of the n-MOS transistor. This work represents a contribution for the realization of low power CMOS ICs where the threshold voltage varies dynamically with the gate voltage. In spite of the mobile carriers injected through the depletion region, it is found that the classic long channel threshold voltage model is still valid for a forward biased substrate. The validity of this statement is extended to the 300–77 K temperature range, for a forward biasing from 0 to 0.5 V, and at two different levels of inversion in the surface: in the onset of moderate inversion and deep in strong inversion. For the former case, nevertheless, a small disagreement between the model and the experimental V_{TH} was found at room temperature. This disagreement was explained not as a violation of the model, but as the result of the presence of parasitic currents that alter the extraction procedure from the subthreshold characteristics. The low temperature measurements showed how the reduction of these parasitic currents results in good agreement between the experimental and fitted values of V_{TH} at the onset of moderate inversion.

ACKNOWLEDGMENTS

This research was supported in part by grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada, Micronet, a federal network center of excellence in microelectronics, and the National Council of Science and Technology (CONACyT) of Mexico. The devices used in this study were obtained from the Laboratoire de Physique des Composants a Semiconducteurs (LPCS), Grenoble, France.

Presented at the 8th Canadian Semiconductor Technical Conference, 12–15 August 1997.

[1]A. P. Chandrakasan and R. W. Brodersen, *Low Power Digital CMOS Design* (Kluwer Academic, Boston, MA, 1995), p. 105.

[2]H. Kotaki, S. Kakimoto, M. Nakano, T. Matsuoka, K. Adachi, K. Sugomoto, T. Fukushima, and Y. Sato, Tech. Dig. Int. Electron Devices Meet. p. 459 (1996).
[3]F. Assaderaghi, D. Sinitsky, S. Parke, J. Bokor, P. K. Ko, and C. Hu, Tech. Dig. Int. Electron Devices Meet. p. 809 (1994).
[4]N. Jankovic, IEEE Trans. Electron Devices **38**, 2527 (1991).
[5]C. Wann, F. Assaderaghi, R. Dennard, C. Hu, G. Shahidi, and Y. Taur, Tech. Dig. Int. Electron Devices Meet. p. 113 (1996).
[6]M. J. Deen, D. S. Malhi, Z. X. Yan, and R. A. Hadaway, 1995 IEEE International Symposium on Circuits and Systems Seattle, WA 29 April–3 May 1995, p. 1968.
[7]Z. X. Yan, M. J. Deen, and D. S. Malhi, IEEE Trans. Electron Devices **44**, 118 (1997).
[8]T. Skotnicki, G. Merckel, and A. Merrachi, Solid-State Electron. **34**, 1297 (1991).
[9]G. Baccarani, M. Wordeman, and R. Dennard, IEEE Trans. Electron Devices **31**, 452 (1984).
[10]M. J. Deen, Proceedings of the 7th International Conference on Microelectronics (ICM '95), Kuala Lumpur, Malaysia, 19–21 December 1995, p. 132.
[11]M. J. Deen and Z. X. Yan, 25th European Solid-State Device Research Conference (ESS-DERC '95), The Hague, Nederland, 25–27 September 1995, p. 643.
[12]T. Yamamoto, T. Mogami, and K. Terada, 1992 Symp. VLSI Technol. Dig., p. 104 (1992).
[13]M. J. Deen and Z. X. Yan, IEEE Trans. Electron Devices **39**, 908 (1992).
[14]M. J. Deen and Z. X. Yan, IEEE Trans. Electron Devices **37**, 1707 (1990).
[15]M. J. Deen, IEE Electron. Lett. **26**, 1493 (1990).
[16]Y. P. Tsividis, *Operation and Modeling of the MOS Transistor* (McGraw-Hill, New York, 1987).
[17]F. H. Gaensslen, V. L. Rideout, and J. J. Walker, IEEE Trans. Electron Devices **24**, 218 (1977).
[18]D. P. Foty, Cryogenics **30**, 1056 (1990).
[19]J. Croon, S. Biesemans, S. Kubicek, E. Simoen, K. De Meyer, and C. Claeys, Proc. Electrochem. Soc. **PV97-2**, 187 (1997).
[20]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley-Interscience, New York, 1983).
[21]M. J. Deen, Microprocessors Microsystems **13**, 245 (1989).
[22]M. J. Deen and Z. X. Yan, Solid-State Electron. **33**, 503 (1990).
[23]Z. X. Yan and M. J. Deen, IEE Proc.-G: Circuits, Devices Syst. **138**, 351 (1991).
[24]S. Mohammadi and C. R. Selvakumar, IEEE Trans. Electron Devices **43**, 185 (1996).

Reduced damage reactive ion etching process for fabrication of InGaAsP/InGaAs multiple quantum well ridge waveguide lasers

B. C. Qiu, B. S. Ooi, A. C. Bryce, S. E. Hicks, C. D. W. Wilkinson, R. M. De La Rue, and J. H. Marsh[a)]

Department of Electronics & Electrical Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

(Received 7 November 1997; accepted 17 April 1998)

The damage introduced into an InGaAs/InGaAsP quantum well structure during CH_4/H_2 reactive ion etching (RIE) processes was measured, for plasma powers from 20 to 100 W, using low temperature photoluminescence. The damage depth profile is estimated to be around 12–70 nm after annealing at 500 °C for 60 s using a rapid thermal annealer. A reduced damage RIE process has been developed to fabricate InGaAs/InGaAsP multiquantum well ridge waveguide lasers. The performance of these lasers has been compared to that of lasers fabricated from the same epilayer using wet etching to form the ridge. The resultant threshold currents were essentially indistinguishable, being 44.5 and 43 mA, respectively, for dry and wet etched lasers with 500 μm long laser cavities. © *1998 American Vacuum Society.* [S0734-211X(98)02804-2]

I. INTRODUCTION

InP and related compound semiconductors are increasingly important for applications in optoelectronics and microelectronics due to their unique electrical and optical properties. Ridge waveguides in optoelectronic devices can be fabricated either by wet chemical etching or by dry etching. Wet chemical etching is believed to be a simple, useful and low damage process, but an important problem with wet etching is that the etch rate is isotropic (the same in all directions) or crystallographic (depending on lattice orientation), so adequate dimensional and profile control can be difficult to achieve. Furthermore, if the waveguide involved is curved or circular, as in, e.g., ring lasers,[1] semiconductor amplifiers with angled facets, and a demultiplexer for optical time division multiplexed (OTDM) systems,[2,3] the crystallographic orientation dependence may rule out the use of wet etching to form the waveguide. On the other hand, dry chemical etching has been shown to be an effective technique due to its highly anisotropic nature. Dry etching is a low pressure process that is used to remove material from a surface. Most dry etching processes are based on sputtering by energetic ions or particle bombardment, evaporation of volatile compounds created through interactions with a reactive gas species or, most frequently, a combination of the two. Dry etching processes that involve particle bombardment usually yield vertical or angled sidewalls regardless of crystallographic orientation. However, the damage which is almost inevitably imparted to the material being etched leads to degradation of the optical and electrical properties of the material.[4–7] It is therefore of great importance to understand and assess the damage, and investigate possible low damage processes for fabrication of optoelectronic devices.

Currently, reactive ion etching (RIE) based on CH_4/H_2 is the most widely used technique for dry etching In containing semiconductors, because it can give better anisotropy[8,9] than RIE using Cl and Br containing compounds.[10] However the damage introduced during the RIE CH_4/H_2 process causes detrimental effects on the performance of optoelectronics devices such as ridge waveguide lasers to the extent that, in some cases, the devices do not work. It is also known that electrical damage due to hydrogen passivation of acceptors occurs during the RIE CH_4/H_2 process, although such damage can be removed by annealing the etched samples[5,6] at a temperature of 350 °C or greater. Therefore it is of great significance to develop a low damage process for fabrication of optoelectronic devices.

In this article, RIE damage was assessed by measuring the damage depth profile using a specially designed multiquantum well (MQW) probe structure and low temperature (5 K) photoluminescence (PL). Post-dry etching annealing was also carried out to investigate its impact on the removal of dry etch damage. Ultimately, 5 μm wide ridge waveguide InGaAs/InGaAsP MQW lasers were fabricated using the RIE dry etching and postannealing procedure developed and were compared to lasers fabricated by wet chemical etching.

II. EXPERIMENTS

A. Material structure

The material structure (see Fig. 1) used in this experiment was grown by metallorganic vapor phase epitaxy (MOVPE) on an n^+InP substrate. First a 100 nm InP buffer layer was grown, followed by 5 $In_{0.53}Ga_{0.47}As$ quantum wells with lattice matched InGaAsP barriers with a band gap of $\lambda_g = 1.26$ μm. The widths of the quantum wells from top to bottom were 2, 4, 6, 8, and 12 nm, respectively. The top four wells were separated by 20 nm InGaAsP barriers, while the 12 nm well was placed 410 nm below the surface to provide the reference signal, this depth being well below any damage induced by the dry etching. Above the top well a 20 nm InGaAsP barrier and a 10 nm InP cap layer were grown.

a)Electronic mail: j.marsh@elec.gla.ac.uk

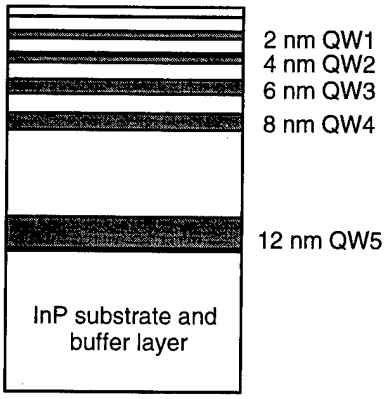


FIG. 1. Structure of the material used to probe the dry etch damage profile. The quantum wells are InGaAs and the barriers are InGaAsP ($\lambda g = 1.26$ μm). The whole structure is lattice matched to InP and has an InP cap.

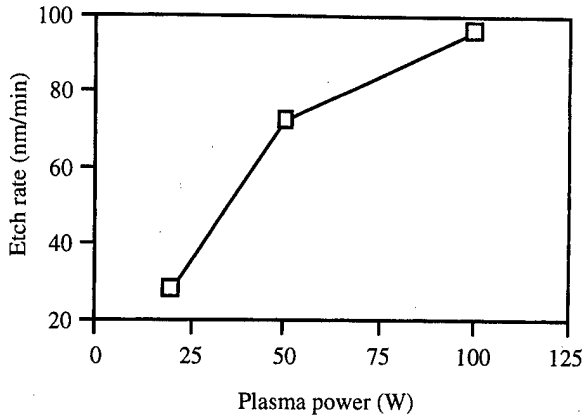


FIG. 2. Etch rate as a function of plasma power.

B. RIE processing

The dry etching was performed in a conventional 13.56 MHz parallel-plate RIE system (Electrotech SRS Plasmafab 340 RIE), which has a 3.3:1 ratio of anode area to cathode area. The samples to be etched were placed on a titanium oxide coated cathode which had a diameter of 17 cm. During etching, the pressure in the chamber was 14 mTorr. The cathode (driven electrode) temperature was regulated by circulating cooling fluid which was maintained at a constant temperature of 30 °C. The plasma powers used were 20, 50, and 100 W.

The etch rates were measured on samples which had patterned SiO_2 masks, the structure of these samples being similar to that of the material used previously.[11] These samples were etched for periods from 15 to 30 min at the different plasma powers. The MQW probe structure was cleaved into samples of size 2×2 mm^2. Dry etching was carried out for 12 s for each plasma power, and the PL was then measured at 5 K.

III. RESULTS AND DISCUSSION

A. Etch rates and surface morphology

The parameters of the RIE processes are detailed in Table I along with the measured etch rates. The etch depths were measured using a depth profile meter. Figure 2 shows the etch rate as a function of the plasma power; it can be seen that the etch rate increases with power. The morphology of the etched samples was also examined using a scanning electron microscope (SEM), and are shown in Fig. 3. Clearly, the higher power etch produced a rougher surface. Such a rough surface would increase the waveguide losses due to scattering and is therefore detrimental to laser operation.

B. RIE damage assessment

Figure 4 shows the low temperature PL corresponding to different dry etch processes, and for an as-grown sample. It should be noted that the PL signals have been normalized with respect to that from the deepest quantum well (QW5), which is believed to be deeper than the range of dry etching damage. Normalization eliminates the requirement for identical sample alignment during the PL measurement. As can be seen from Fig. 4, the wells which have been damaged during the RIE process exhibit broadening of the PL peaks, the peaks from the different wells are no longer well resolved and the intensities are reduced substantially. Only the deepest well shows no broadening, confirming that it is below the damaged region of the samples. QW1 and QW3, from which the PL signals can still be estimated in terms of peak intensity and full width at half maximum (FWHM), have been chosen to characterize the damage. Figure 5(a) shows the ratio of PL peak intensity of QW1 of the etched samples to that of the as-grown sample, as a function of rf power. When 100 W RIE power was used, no signal could be detected from QW1 after etching, even though only about 19 nm of top layer was removed by etching, leaving more than 10 nm InGaAsP above the first well. Figure 5(b) shows the FWHM of the PL spectrum of QW3 of the etched sample. The PL peak becomes increasingly broader with increasing plasma power. The reduction in PL intensity and the increased

TABLE I. Parameters of the RIE CH_4/H_2 processes.

	CH_4 flow rate (sccm)	H_2 flow rate (sccm)	rf power (W)	DC bias (−V)	Pressure (mTorr)	Etch rate (nm/min)
Process 1	7.2	52	20	430	14	28
Process 2	7.2	52	50	556	14	72
Process 3	7.2	52	100	760	14	96

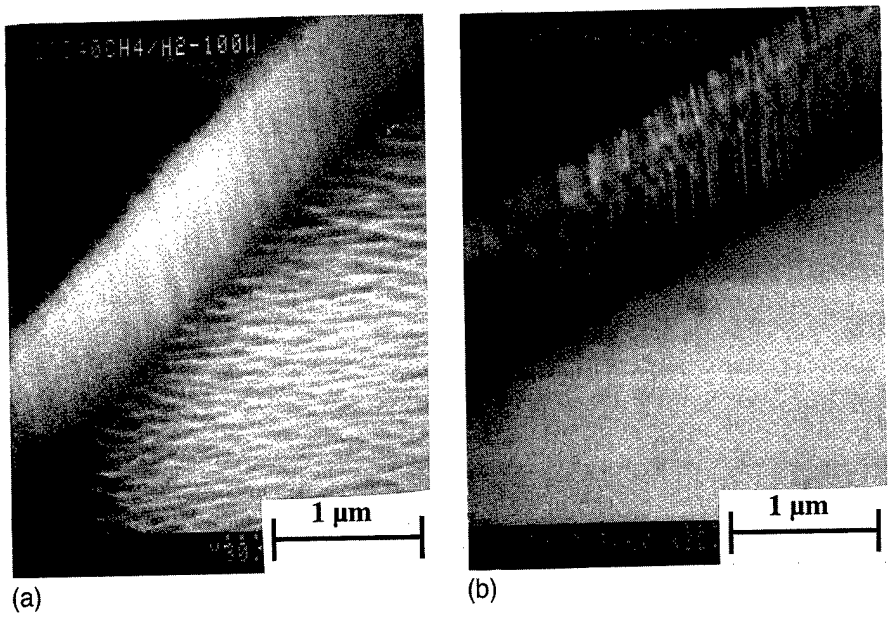


FIG. 3. SEM micrographs showing the surface morphology after etching with plasma powers of 100 (left) and 20 W (right).

broadening as the plasma power increases indicate that more damage has been incorporated into the samples etched with the higher power plasma.

From Fig. 4, it can be seen that significant PL recovery is obtained after annealing. In most cases the peak FWHM is reduced to that of the as-grown value and the peak intensity is increased. This effect may be attributed to the removal of damage, both point defects and hydrogen passivation. However, the PL intensity of the shallow wells is not recovered in all cases. For 20 W of rf power, the PL is almost the same as that of the as-grown sample. When 50 W etch power is used, the top well is not fully recovered and in the case of 100 W

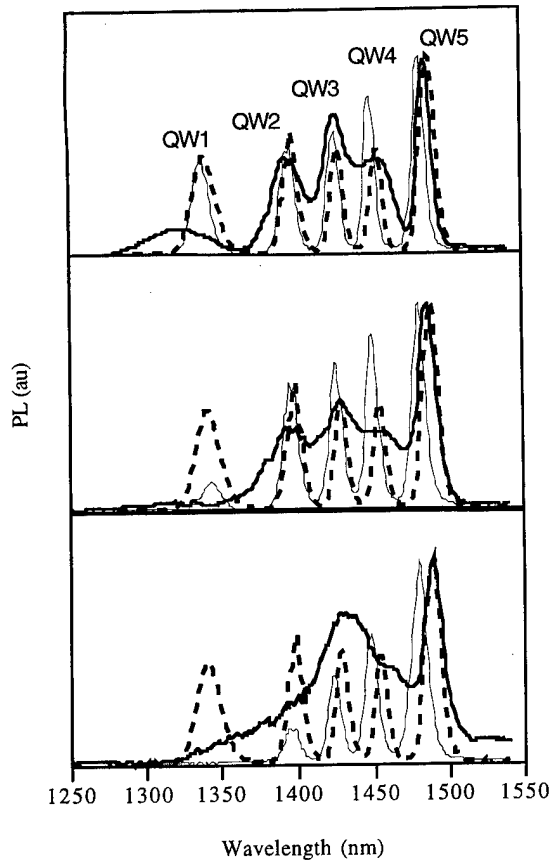


FIG. 4. 5 K PL spectra from the MQW sample. Plasma powers are 20, 50, and 100 W, respectively, from top to bottom. Dashed lines represent the PL from as-grown samples, thick lines from etched samples before annealing and thin lines from etched samples after annealing.

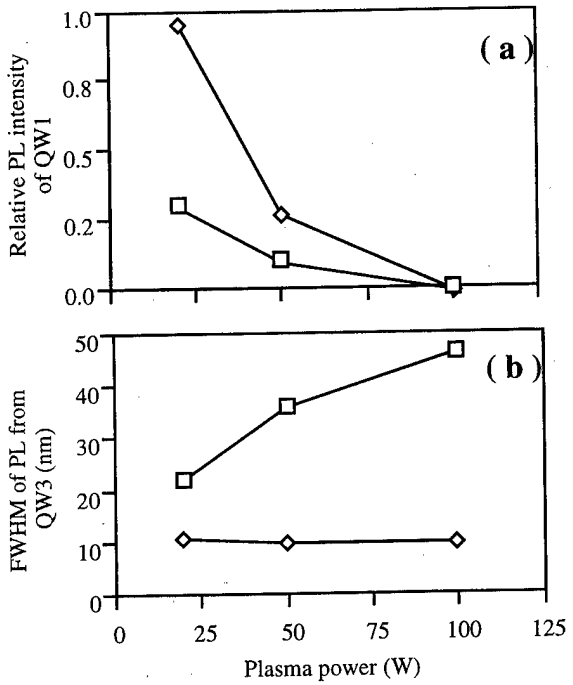


FIG. 5. (a) Ratio of PL intensity of QW1 from the etched sample to that from the as-grown sample as a function of plasma power, and (b) FWHM of PL spectra from QW3 as a function of plasma power. □ before annealing, ◇ after annealing of the etched samples.

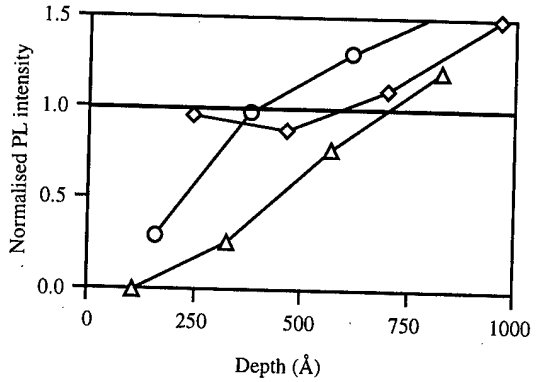


FIG. 6. Normalized PL intensity of the QWs as a function of depth after etching and annealing for plasma powers of 20 W (◇), 50 W (○), and 100 W (△).

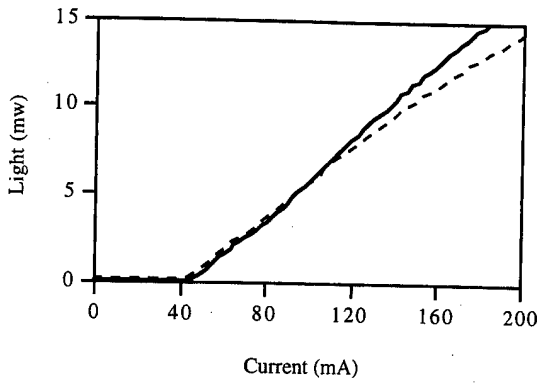


FIG. 8. Comparison of light-current characteristics of dry etched (solid line) and wet etched (dashed line) lasers.

etch power there is no PL signal from the top well and the PL signal from QW2 is only partially recovered, suggesting that extended defects may have been formed in this region. It should also be noted that the peaks are shifted to wavelengths shorter than those of the as-grown samples, this being due to quantum well intermixing caused by the diffusion of point defects through the wells. The blue shift increases with plasma power indicating that greater damage is introduced into the sample as the plasma power is increased.

By normalizing the PL intensities of the quantum wells of the etched samples to those of the corresponding wells of the as-grown sample, the dependence of the relative intensity of PL emission as a function of the depth can be plotted (Fig. 6). From this dependence, it is possible to make an estimate of the damage depth. However, due to the limited number of wells in the material structure used to probe the RIE damage, and the spread of the electron and hole wave functions, this estimate has an error of the order of 12 nm. It should be noted that QW1 of the sample etched with 20 W of power, is assumed to be recovered completely after annealing, as its normalized PL intensity is very close to 1, hence the damage range for this plasma power should be less than the distance between the surface of the etched sample and the QW1. From Fig. 7, it can be seen that the estimated damage depths after annealing are around 12, 27, and 70 nm, corresponding to plasma power levels of 20, 50, and 100 W, respectively.

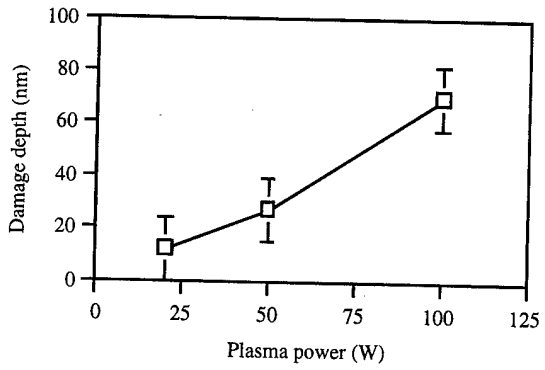


FIG. 7. Estimated damage depth as a function of plasma power.

C. Comparison between the RIE dry etched lasers and wet etched lasers

Note that 5 μm wide ridge waveguide lasers were fabricated by both wet chemical etching and RIE dry etching to assess the RIE process developed in a device application. The laser material, which contained 5 InGaAs/InGaAsP QWs, was grown by MOVPE and was similar to that used by McKee *et al.*[12] Process 2 in Table I was used to form the ridge waveguide, this process being chosen since it produces a relatively high etch rate with lower damage. The sample was etched for around 17 min to give a ridge height of 1.2 μm. After etching, the sample was annealed at 500 °C for 60 s, then routine laser fabrication procedure was followed. The light-current characteristics of groups of the lasers etched by RIE and by wet etching were measured. The average threshold currents are 48.4 and 45.2 mA for the dry etched and wet etched lasers respectively, i.e., essentially indistinguishable given that the ridge widths are not identical. Figure 8 shows the light-current characteristics of the lasers which have the lowest threshold currents, namely 44.5 and 43 mA for the dry etched and wet etched lasers, respectively. It should also be noted that the slope efficiencies of the lasers are very similar. The very close performance of the lasers suggests that the dry etch damage is small enough to have no significant effects on the laser qualities. This correlates with the results obtained when etching the QW probe sample, despite the much longer dry etch time used for the laser (17 min), as compared to that used for the probe material (12 s).

IV. CONCLUSIONS

By using the low temperature PL technique together with a specially designed material structure, the damage introduced by RIE CH_4/H_2 process was estimated to be in the range of 12–70 nm after annealing, for rf plasma powers in the range 20–100 W. The characteristics of InGaAs/InGaAsP multiquantum well lasers fabricated by the RIE process show that the damage does not significantly affect the laser qualities, provided a relatively low power process and post-dry etch annealing are used.

1822

ACKNOWLEDGMENTS

B. C. Qiu was supported by the Sino-British Friendship Scholarship Scheme (SBFSS). The work was supported by the UK Ministry of Defense and the Engineering and Physical Sciences Research Council. The supply of epitaxial wafers by C. C. Button is gratefully acknowledged.

[1]P. B. Hansen, G. Raybon, M.-D. Chien, B. I. Miller, M. G. Young, J.-M. Verdiell, and C. A. Burrus, IEEE Photonics Technol. Lett. **4**, 411 (1992).
[2]C. E. Zah, J. S. Osinski, C. Caneau, S. G. Menocal, L. A. Reith, J. Salzman, F. K. Shokoohi, and T. P. Lee, Electron. Lett. **23**, 990 (1987).
[3]E. Jahn, N. Agrawal, H.-J. Ehrke, R. Ludwig, W. Piepe, and H. G. Weber, Electron. Lett. **32**, 216 (1996).
[4]S. J. Pearton, W. S. Hobson, U. K. Chakrabarti, G. E. Derkits, and A. P. Kinsella, J. Vac. Sci. Technol. B **8**, 1274 (1990).
[5]M. Moehrle, Appl. Phys. Lett. **56**, 542 (1990).
[6]T. R. Hayes, W. C. Dautremont-Smith, H. S. Luftman, and J. W. Lee, Appl. Phys. Lett. **55**, 56 (1989).
[7]D. G. Yu, B. P. Keller, A. L. Holmes, Jr., and E. L. Hu, J. Vac. Sci. Technol. B **13**, 2381 (1995).
[8]E. Andideh, I. Adesida, T. Brock, C. Caneau, and V. Kermidas, J. Vac. Sci. Technol. B **7**, 1841 (1989).
[9]T. R. Hayes, M. A. Dreisbach, P. M. Thomas, W. C. Dautremont-Smith, and L. A. Heimbrook, J. Vac. Sci. Technol. B **7**, 1130 (1989).
[10]K. Takimoto, K. Ohnaka, and J. Shibata, Appl. Phys. Lett. **54**, 1947 (1989).
[11]O. M. Steffensen, D. Brikedal, J. Hanberg, O. Albrektsen, and S. W. Pang, J. Appl. Phys. **78**, 1528 (1995).
[12]A. McKee, C. J. McLean, A. C. Bryce, R. M. De La Rue, and J. H. Marsh, Appl. Phys. Lett. **65**, 2263 (1994).

X-ray photoelectron spectroscopy damage characterization of reactively ion etched InP in CH_4–H_2 plasmas

Y. Feurprier, Ch. Cardinaud,[a] and G. Turban
Laboratoire des Plasmas et des Couches Minces-IMN, UMR 6502 CNRS-Université de Nantes, BP 32229, 2 rue de la Houssinière, 44322 Nantes cédex 3, France

(Received 31 September 1997; accepted 17 April 1998)

The plasma–surface interaction during CH_4–H_2 reactive ion etching processing of InP is described in detail by means of plasma diagnostics (optical emission spectroscopy and mass spectrometry) and x-ray photoelectron spectroscopy (XPS) surface analysis. The influence of the input power is carried out for different CH_4–H_2 mixtures in terms of InP etch rate, etch product and CH_3 radical detection and surface damage characterization. In particular detailed XPS results allow the study of the changes in the stoichiometry and amorphization of the surface with the input power. In addition, for a given power, the quality of the etched surface improves by increasing the fraction of methane in the gas mixture. As an example, the best surface stoichiometry ($InP_{0.86}$) is obtained for a pure methane plasma running at a high power (300 W). In general, it is shown that the lower the P depletion, the lower the amorphization, which is indicative of a general improvement of the etched surface quality. Based on the XPS results, a three-layer model is proposed for the representation of the surface in the course of etching. The damaged layer situated over the bulk InP is composed of a superficial P-depleted layer and of a stoichiometric amorphized InP layer. Using the curve-fitting of the P $2p$ spectra, the thickness of the different layers is estimated. As an example, a damaged layer as low as 37 Å thick is obtained for pure methane plasma at 15 mTorr and a power of 300 W, whereas our standard conditions (10% CH_4–H_2, 50 mTorr, and 80 W) give a damaged layer of 90 Å. The experimental observations give evidence of the need for both ion bombardment and active neutral species to obtain etching. The improvement of the etch process is then explained by an improved In removal rate which is actually the limiting step in the etching mechanism of InP.
© *1998 American Vacuum Society.* [S0734-211X(98)02204-5]

I. INTRODUCTION

Niggebrügge[1] demonstrated the efficiency of methane–hydrogen plasmas to etch III-V materials (GaAs, InP, and InGaAsP). This hydrocarbon-based chemistry then appeared as an alternative to chlorine and chlorinated gases for the etching of In-based III-V semiconductors. In brief, the advantages of this plasma chemistry are in the noncorrosive, nontoxic nature of the gases and in the smooth morphologies of the etched surface obtained at room temperature. Due to the growing interest in materials of the InP system, investigations of the etching conditions of InP and related ternary alloys is well documented for the fabrication of micro- and optoelectronic devices.[2–7] In general, the different discharge parameters are optimized to allow high InP etch rates while keeping good surface morphology and limiting the introduction of damages due to the plasma process.[8,9] But few papers really described experimental results on the CH_4–H_2 reactive ion etching (RIE) of InP based on a plasma–surface interaction approach to give more insight into the understanding of the etching mechanism. Actually a comprehensive study of the InP etching mechanism should rely on a precise analysis of both the gas phase (neutral active species and etch products) and of the surface after etching. In most papers, plasma phase characterization and precise surface analysis data on the induced damages by the RIE process are still lacking.

The objective of this article is to detail the influence of the input power in terms of etch rate (ER), active neutrals, etch products, and chemical damage. In addition experiments are also conducted for different CH_4–H_2 mixture compositions. High resolution x-ray photoelectron spectroscopy (XPS) analyses allow us to determine precisely the surface stoichiometry and to estimate the degree of amorphization of the surface after etching. An evaluation of the damaged layer thickness due to the RIE process is then given. We especially focus on a precise characterization of damages induced by the plasma environment including stoichiometry changes, residual carbon contamination, and crystalline defects by means of XPS.

II. EXPERIMENT

A. Etching conditions

The RIE surface analysis facility has been described elsewhere.[10] It consists of an ultrahigh vacuum surface analysis chamber connected to a plasma reactor. The etching process is operated in a parallel plate diode reactor designed to study the plasma–surface interaction. A schematic drawing of the etch chamber with the different plasma diagnostics available for the analysis of the plasma phase has been reported previously.[11] Briefly, the lower electrode is capacitively coupled to a radio frequency (rf) (13.56 MHz) generator through a tunable matching network. The upper electrode

[a]Electronic mail: Christophe.Cardinaud@cnrs-imn.fr

is grounded and fitted with the mass spectrometer. Electronic grade CH_4 and H_2 gas sources are employed, and etch runs are performed on semi-insulating (100) oriented InP samples with an etching time of 10 min to be in a stable etching regime.[12]

We have varied the applied power from 10 to 300 W for three different mixture compositions, 10% CH_4–H_2, 50% CH_4–H_2 (50 mTorr), and pure methane plasma (15 mTorr). To operate under pure methane plasma, a lower pressure of 15 mTorr is actually required to etch InP, since for 50 mTorr, deposition of amorphous carbon (a-C:H) occurs. A more extensive investigation on the influence of plasma parameters is then provided from our standard plasma conditions (10% CH_4–H_2, 80 W, 50 mTorr).

B. Plasma diagnostics

The light of the discharge is taken through a lateral window that allows the sampling of the region of the plasma in the vicinity of the InP wafer. A Jobin–Yvon HR640 monochromator and an optical multichannel analyzer (EG&G 1455) is used to investigate the emission of the plasma in the 200–800 nm range. Optical emission spectroscopy (OES) is used to monitor the indium etching through the In* emission line located at 451.1 nm. Detailed mass spectrometric (MS) characterization of the gas phase during CH_4–H_2 RIE of InP allows the detection of phosphine as the major etch product of P and of organoindium positive ions.[13] Mass spectrometric sampling of the plasma phase is performed with a Balzers PPM 421 mass spectrometer which is mounted within the grounded electrode. Its extraction hood is thus immersed in the plasma. A 100 μm in diameter orifice allows ionic and neutral species to effuse into the mass spectrometer, yet keeping the operating pressure lower than a few 10^{-5} mbar.

For ionization of neutrals, the energy of the electron beam can be varied from 8 to 120 eV allowing the detection of atoms and radicals by means of the ionization threshold technique. The ionization potential of Ar referenced at 15.7 eV is used to calibrate the electron energy scale. The neutral detection for the etch product identification is performed with an electron energy fixed at 25 eV to reduce the fragmentation of molecules, and the emission current is set at 0.1 mA. In this article, mass spectrometry is used to monitor In and P etch products as a function of the input power for 10% CH_4 RIE of InP. The detection and the determination of the absolute density of the methyl radicals is also performed for 10% and 50% CH_4 containing mixtures.

C. Surface analysis

XPS analyses are carried out *ex situ* at the CNRS-Université de Nantes surface analysis facility (Leybold AG LHS12) in order to benefit from the high resolution of the monochromatized Al $K\alpha$ x-ray source, and have the most confidence in peak curve-fitting. However, some experiments have been carried out with the quasi *in situ* equipment to check that the *ex situ* analysis procedure did not spoil the chemical information. This allows one to obtain precise information on the chemical modifications and on the degree of amorphization of the surface. Using a pass energy of 25 eV, the spectral resolution is 0.34 eV. The spectrometer energy scale is calibrated to the Au $4f7/2$ core level line at 84.00 eV. For the spin-orbit doublets, the binding energy positions and the full width at half maximum (FWHM) given are those of the component of greater angular momentum. The FWHM relative to the P–In bond component of the P $2p$ spectra is used to monitor the ion induced damages of the RIE process. FWHM reference values of 0.57 eV (P $2p$) and 0.63 eV (In $4d$) have been measured for a monocrystalline InP sample. Background subtraction is done by using Shirley's method.[14] Whenever useful, to quantify the relative importance of the various contributions, a curve-fitting is undertaken following Marquardt's algorithm[15] and using Gaussian–Lorentzian curves. Prior to etching the decomposition of the P $2p$ and In $4d$ distributions into oxide and bulk components gives a value of 2.35 for the ratio of the bulk components. An identical result has been directly measured by the analysis of an oxide-free metalorganic chemical vapor deposition (MOCVD) epitaxially grown InP. In the following, we will refer to this ratio as the determination of the mean stoichiometry of the surface in the layer probed by XPS.

When appropriate, curve-fitting results are used to estimate the thickness (z) of any overlayer (M) present on the InP substrate (S) by using the following general expression:

$$I_S = I_S^{\infty} \cdot \exp\left(- \frac{z}{\lambda_M \cdot \cos(\theta)} \right), \tag{1}$$

where λ_M is the inelastic mean free path of photoelectrons in the overlayer M, θ is the detection angle for the photoelectrons with respect to the normal to the surface, and I_S^{∞} is the intensity of the core level peak observed when the substrate is not covered with an overlayer. In our case, I_S^{∞} has been determined on our XPS system using the contamination-free MOCVD epitaxially grown InP sample.

III. RESULTS AND DISCUSSION

A. Etch rate measurements and etch products

InP etch rates are determined by weight loss measurements on a quarter of a 2 in. in diameter InP wafer. Considering the etching homogeneous, data are converted into thickness loss and consequently into a mean etch rate. Figure 1 displays both InP ER and the self-bias voltage as a function of the applied power under our standard plasma conditions. As expected for a parallel plate diode discharge the self-bias voltage follows a square root function of the power delivered to the plasma. For 10 W no measurable ER is found although the surface composition changes as further observed by XPS. The InP ER continuously increases with the rf power from 50 Å/min at 30 W up to 900 Å/min for a power of 300 W due to the higher decomposition and ionization of the plasma. But the increase of the InP ER is not equivalent over the whole range of power. For high power conditions ($\geqslant$70 W) a linear dependence of the InP ER is obtained with the rf applied power. In contrast, for the low power region

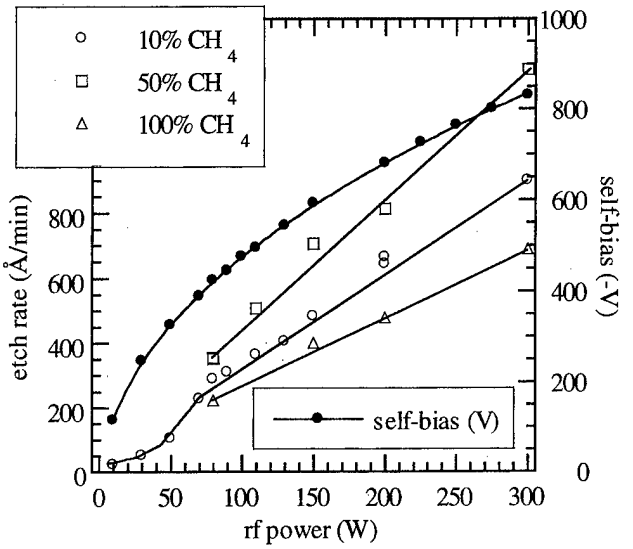


FIG. 1. InP etch rate dependence with the input power using three different CH_4–H_2 mixtures. Plasmas containing 10% CH_4 and 50% CH_4 are performed at a pressure of 50 mTorr while 15 mTorr is used for pure methane plasma. The evolution of the self-bias voltage is reported for the 10% CH_4–H_2 mixture composition and is fitted with a square root function characteristic of a parallel plate diode discharge.

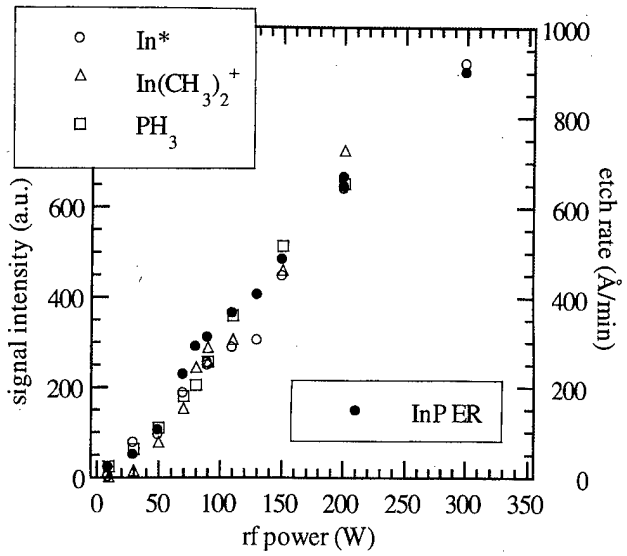


FIG. 2. Intensity of the In and P products present in the plasma phase monitored by emission spectroscopy (In* line) and mass spectrometry [PH_3 and $In(CH_3)_2^+$] as a function of power and compared to the InP ER evolution.

(<70 W) lower ER are observed than those expected from the high power results. This observation is further discussed with the XPS analysis data.

Dry etching of InP in CH_4-based plasmas is also conducted with richer methane fractions in the gas mixture (50% and 100%). Under our standard pressure condition, the etching/*a*-C:H deposition transition is observed for mixtures containing 80%–85% CH_4. A lower pressure of 15 mTorr is necessary to recover etching conditions of InP under pure methane plasma. Figure 1 shows that increasing the mixture composition in methane leads to higher InP ER. As an example, for 300 W it rises from 900 to 1250 Å/min when increasing the methane percentage from 10% to 50%. A similar evolution is observed when operating at 15 mTorr and 80 W, condition for which ER of 180 and 220 Å/min are obtained for 50% and 100% CH_4, respectively. In first approximation, these observations showing the increased InP ER with the power and the CH_4 relative fraction are simply explained by the much higher decomposition of the gas and the higher energy flux of ions bombarding the sample.

B. Plasma phase analysis

1. Etch product detection

It is shown elsewhere[13] that MS and OES on the plasma phase during the RIE process allow the detection of signals closely related to the In and P etch products. MS indicates that phosphine is the major neutral P compound present in the plasma phase and the main reaction product of P with H atoms generated in the discharge. As mentioned by Melville[16] and in a previous work,[13] indium compounds are only detected as positive ions in the form of In^+ and $In(CH_3)_2^+$ ions at $m/e=115$ and 145 amu, respectively. OES displays the existence of a In* line at 451.1 nm. Figure 2 compares the evolution of the ER with the intensity of PH_3, $In(CH_3)_2^+$, and In* signals as a function of the rf applied power. Since the evolution for In^+ and $In(CH_3)_2^+$ ions are identical, only that of $In(CH_3)_2^+$ is reported for clarity. One can see the very good agreement between the various plasma diagnostics. The good correlation of these experimental data together with the etch rate gives confidence in the choice of these signals to represent the etching of InP. In particular, it strongly supports the fact that the In^+ and $In(CH_3)_2^+$ ions likely arise from the ionization in the plasma phase of the organoindium volatile etch products and not from secondary ion–molecule reactions between the In reaction product and ions of the discharge.

2. Etching agent detection

Methyl radicals are regarded as the main active species for the elimination of In by forming a volatile organoindium compound. No direct evidence of the nature of the indium etch product has ever been reported in the literature. However we published experimental results[11] that indirectly support the reaction of CH_3 with In to form organoindium volatile reaction products. CH_3 radicals can be detected by mass spectrometry using the threshold ionization technique. Described in detail elsewhere,[11] the principle of this method is based on the difference of a few electron volts between the dissociative ionization threshold of CH_4 into CH_3^+ (E_2 = 14.3 eV) and the direct ionization threshold of CH_3 radical ($E_1=9.8$ eV). Figure 3 shows that the increase in CH_3 concentration is almost linear with the power even though a slight change in the slope is denoted around 80 W. One must be aware that the CH_3 density one obtains results from the constant balance between source and loss terms. CH_3 radical production is expected to increase because of a higher de-

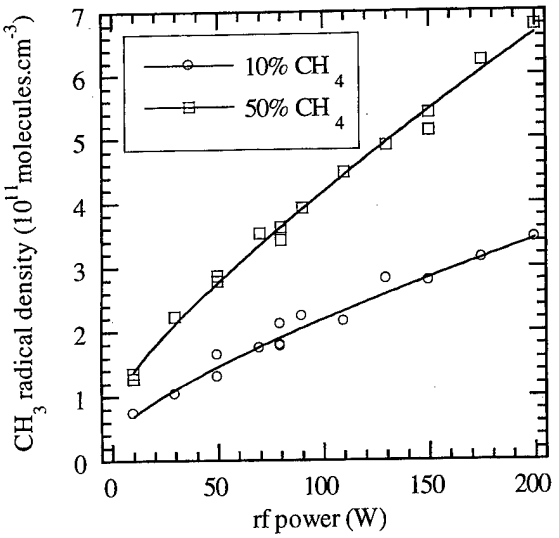


FIG. 3. Evolution of the CH_3 radical density as a function of power for 10% and 50% CH_4 containing mixtures. The absolute density is determined by threshold ionization mass spectrometry.

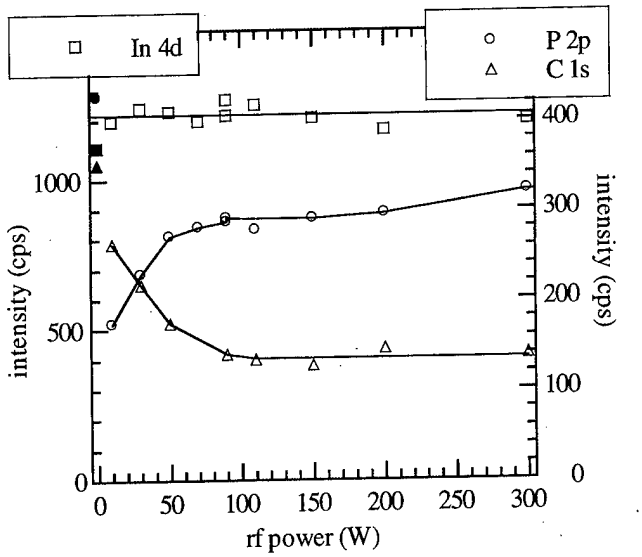


FIG. 4. In $4d$, P $2p$, and C $1s$ XPS intensities vs power to monitor the surface composition changes under our standard plasma conditions (10% CH_4–H_2 and 50 mTorr). Full patterns at 0 W represent the surface before etching.

composition of CH_4. On the other hand, loss due to a-C:H deposition on the chamber walls and to etching of InP is also seen to increase. However, as the overall CH_3 density increases, it can be concluded that loss processes, and InP etching in particular, are not strictly limited by the CH_3 concentration. For any given power, 50% CH_4 in the gas mixture gives a higher CH_3 density in the plasma phase. This is in agreement with previous measurements[11] showing a continuous increase of the CH_3 density with the percentage of CH_4.

The H atom density cannot be estimated by emission spectroscopy since under our experimental conditions the H atomic lines of the Balmer system mainly originate from the dissociative excitation of molecular H_2 rather than the direct excitation of H atoms in their fundamental state. Actinometric measurements are consequently not applicable for an estimation of the concentration of H atoms produced in the CH_4–H_2 discharge. In H_2 rf discharge, Kae-Nune[17] detected and estimated the H atom density by mass spectrometry using the threshold ionization technique. In the experimental conditions we used, the H atom density could not be estimated in a reliable and reproducible manner in CH_4–H_2 mixtures by using the same technique. But due to the presence of CH_4 and H_2 in the mixture the H atom concentration of the plasma phase is thought to be high enough so that the volatilization of P is not limited by the H atom supply to the semiconductor surface. This statement is well correlated with the highly P-depleted surface whatever the plasma conditions as evidenced hereafter by XPS.

To conclude this section, the plasma diagnostics (MS and OES) give complementary results to monitor In and P plasma species which are representative of the removal rate of the group III and V element. Since CH_3 radicals are regarded as the reactive species for In, the measurement of their density is a key element to gain a better understanding of the InP etching since the indium elimination rate, as the limiting step, governs the InP etching mechanism.

C. XPS surface analysis

1. Damage characterization

Plasma etching quickly removes the native oxide from the surface. But due to a short exposure (about 30 s) to atmosphere during the transferring procedure from the etch chamber to the XPS apparatus, a small amount of oxygen is detected on the surface, typically equivalent to half the amount of the native oxide present on the control sample. However, despite this chemical change, the most important surface information is not altered by this *ex situ* procedure as it gives reliable indications on the surface stoichiometry, residual carbon contamination, and degree of amorphization of the surface.

In a previous paper,[12] it has been shown that exposing the InP substrate to the CH_4–H_2 RIE plasma results in a highly P-depleted and amorphized surface. The analysis of C $1s$ and P $2p$ spectra gave strong evidence of the existence of C–In and (In)P–H species, which are indicative of the presence on the surface of residues or precursors of both In and P reaction products. Moreover the P depletion also results in the appearance on the In $4d$ peak of a shoulder attributed to In–In bonds with a second neighbor effect due to C or P (In–In–C or In–In–P species).

Figure 4 shows the C $1s$, P $2p$, and In $4d$ intensities monitored as a function of power for the 10% CH_4–H_2 mixture. The corresponding intensities for the control sample are given at 0 W to allow a direct comparison with those of the etched surface. These results indicate that the In $4d$ intensity remains constant over the whole power range studied, demonstrating no major change of the In atom density. On the contrary, one clearly notices the variation of the P $2p$ intensity. As expected, the major modification in the surface chemical composition is the depletion of phosphorus. A de-

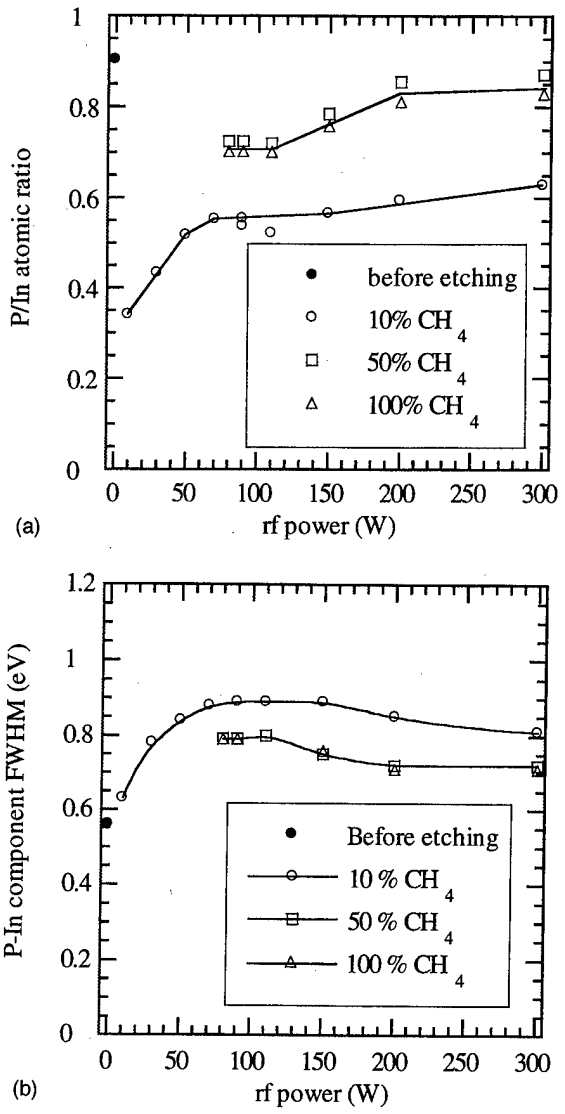


FIG. 5. Surface stoichiometry (a) and P–In component FWHM, (b) vs the power for the different CH_4–H_2 mixture compositions.

crease in the amount of carbon present on the surface is noticed between 10 and 70 W, then reaching a stable value onwards. Considering the residual carbon contamination, Fig. 4 indicates a clean CH_4–H_2 RIE process since the amount of carbon present on the surface after etching is always lower than that on the control sample.

Comparing Figs. 1 and 4, one sees that for powers between 10 and 70 W, the lower the power, the lower the etch rate and the greater the P depletion. It suggests that low power results in a less efficient elimination of In, whereas P is still etched in a normal manner. A preliminary interpretation is that as the power gets lower in the 10–70 W range, the ionic assistance to the etching of In decreases strongly as compared to the neutral contribution.

The mean stoichiometry of the XPS probed depth is estimated by calculating the In $4d$/P $2p$ ratio and comparing it to the reference value of 2.35 for stoichiometric InP. Results are reported in Fig. 5(a). As discussed above a high P depletion is obtained for 10 W ($InP_{0.33}$). This value improves strongly when increasing the power to 70 W ($InP_{0.55}$) and remains stable until 150 W. Increasing the power over 150 W again leads to a lower P depletion (Fig. 4) resulting in a stoichiometry of $InP_{0.63}$ at 300 W [Fig. 5(a)]. Since one of the objectives is to minimize the damages induced by the etch process, further detailed work is focused on high power discharge conditions for richer methane mixtures.

Experiments were then conducted from 80 to 300 W for 50% CH_4–H_2 and pure methane plasma at a pressure of 50 and 15 mTorr, respectively. Similarly to 10% CH_4–H_2 results, increasing the power delivered to the discharge leads to a decrease of the P depletion as shown in Fig. 5(a). In addition, for any power, a better surface stoichiometry is obtained when increasing the percentage of CH_4 in the gas mixture, as the increase of the CH_3 density in the gas phase[11] results in a better elimination of In. As an example, for 80 W, the surface stoichiometry increases from $InP_{0.55}$ to $InP_{0.70}$ for 10% CH_4–H_2 and 50% CH_4–H_2, respectively. Data obtained for rich methane gas mixtures (over 50% CH_4–H_2) are very similar to those at 50% CH_4–H_2. A slightly better stoichiometry is observed in pure methane plasma whatever the input power used. Therefore the best surface quality is obtained in pure methane plasma using a power of 300 W (power density of 1.7 W/cm^2) and results in a surface stoichiometry of $InP_{0.86}$.

The modification of the InP surface after etching is displayed in Fig. 6 which reports the In $4d$ and the P $2p$ spectra for some typical etching conditions. The comparison with the MOCVD InP sample highlights the major chemical changes we observe. The better surface stoichiometry is evidenced by the increasing P $2p$ intensity with power or methane fraction. Accordingly, on the In $4d$ distribution, the decreasing phosphorus depletion results in the decreasing intensity of the shoulder located around 17 eV assigned to In–In–C or In–In–P species. The P $2p$ distribution [Fig. 6(a)] is asymmetric due to the presence of a component situated at 130.3 eV assigned to (In)P–H. The main component located at 129.1 eV is relevant to P–In. A simple decomposition of the P $2p$ spectra with these two contributions then leads to the determination of the FWHM associated with P–In. The decomposition of the In $4d$ spectrum [Fig. 6(b)] is more complicated due to a higher number of contribution: In–P, In–C and some In–In–C or In–In–P species. Moreover the first two are largely overlapping since In–P and In–C contributions have similar binding energies. So only the P $2p$ spectra are used to estimate the degree of damage of the surface after the etch process. Figure 5(b) reports the P $2p$ (P–In) FWHM data obtained for the various plasma conditions. For the 10% CH_4–H_2 mixture, the FWHM increases from 0.63 to 0.89 eV as the power rises from 10 to 70 W, in reason of the increasing ion bombardment. Then it remains stable until 150 W before slowly decreasing with the power and reaching 0.82 eV at 300 W. For 80 W a much lower value of the FWHM is obtained when the percentage of methane is increased to 50% CH_4 clearly demonstrating a decrease in the damage of the surface. Figure 5(b) confirms that for any power increasing the methane content in the gas mixture results in a less

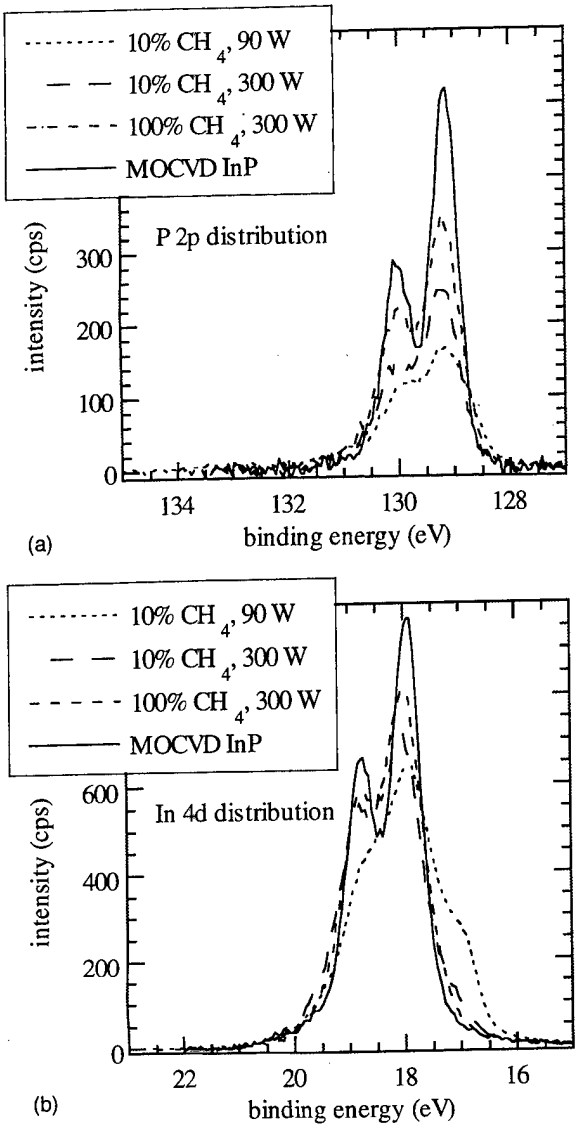


FIG. 6. P $2p$ (a) and In $4d$ (b) XPS spectra at different typical etching conditions showing the improvement of the surface quality with increasing both the power and methane fraction. CH_4–H_2 RIE processed InP samples are compared to the MOCVD InP sample.

damaged surface. It should be recalled that increasing the methane concentration causes the ion flux and the average ion energy to decrease, whereas the CH_3 flux increases. It has been shown previously that this leads to a better surface stoichiometry as well as a less damaged surface.[11] This improvement in the etching of InP has been interpreted as being due to a better balance between the active neutral flux and the ion flux on the surface when increasing the methane fraction. This phenomenon is observed for all applied powers. The effect of the applied power is a little more complex, as increasing the power affects the ion flux, the ion energy, and the CH_3 flux. From 10 to 80 W, the increase of the etch rate, the P/In ratio and the P $2p$ (P–In) FWHM shows that the InP etching mechanism depends strongly on the ion assistance. Between 80 and 150 W the plateau observed in Figs. 5(a) and 5(b) indicates that some equilibrium has been reached between both fluxes. Finally, above 150 W, the slight but significant improvement in the surface stoichiometry as well as in the P–In FWHM suggests that the CH_3 flux/ion flux ratio increases. 50% CH_4–H_2 and 100% CH_4–H_2 RIE processes give the same degree of amorphization [Fig. 5(b)], in agreement with their close P depletion [Fig. 5(a)], suggesting damaged layers of similar thicknesses. The decrease of the surface amorphization with increasing percentage of CH_4 and/or power is demonstrated through a better resolution of the P $2p$ doublet [Fig. 6(a)]. Still under our best RIE conditions (100% CH_4, 15 mTorr, 300 W) the surface is P-depleted and amorphized as seen in Fig. 6.

2. Modeling of the surface in the course of etching

For high power conditions, the simultaneous decrease of the P depletion, and the degree of amorphization of the surface demonstrates the much better surface quality corresponding to a decreasing damaged layer. In fact the interpretation of these data strongly suggests that bulk InP still contributes to the P $2p$ spectra. In other words, the damaged layer thickness is lower than the probed depth by XPS. Recently we proposed a three-layer model to describe the phosphorus depletion and the amorphization of the etched surface.[11] Figure 7(a) recalls this model: from the surface downwards to the bulk it consists of: (i) a superficial layer, amorphous, depleted in P, containing In–In–C (or In–In–P), (In)P–H, In–C, and C–C species, this layer is assumed to present a gradient of concentration in P which we discuss hereafter, (ii) a damaged stoichiometric InP layer, and (iii) the InP bulk substrate. Angular XPS observations[11] have shown a nonuniform location of phosphorus in the analyzed layer. In contrast In is uniformly present and its surface density equals that of the bulk: $N_{In}=5.8\times10^{14}$ cm^{-2}.

To give a representation of the repartition of phosphorus in the superficial layer, an exponential profile has been postulated for the phosphorus atom density:

$$N_P(z)=N_P\cdot\exp\left(-\frac{(z-z')}{d}\right)\quad 0\leqslant z\leqslant z', \tag{2}$$

where $N_P=5.8\times10^{14}$ cm^{-2} for a (100) oriented InP wafer, and z' is the thickness of the superficial P-depleted layer.

Then expressing the intensity of the P $2p$ spectrum as

$$I_P=\int_0^{\infty}N_P(z)\cdot\exp\left(-\frac{z}{\lambda\cdot\cos(\theta)}\right) \tag{3}$$

with $\lambda=28.6$ Å,[17] and considering that $N_P(z)=N_P$ for $z\geqslant z'$, leads to

$$\frac{I_P}{I_P^{\infty}}=\frac{1}{1-\dfrac{\lambda\cdot\cos(\theta)}{d}}\cdot\left[\exp\left(-\frac{z'}{d}\right)-\exp\left(\frac{z'}{\lambda\cdot\cos(\theta)}\right)\right]+\exp\left[\frac{z'}{\lambda\cdot\cos(\theta)}\right]. \tag{4}$$

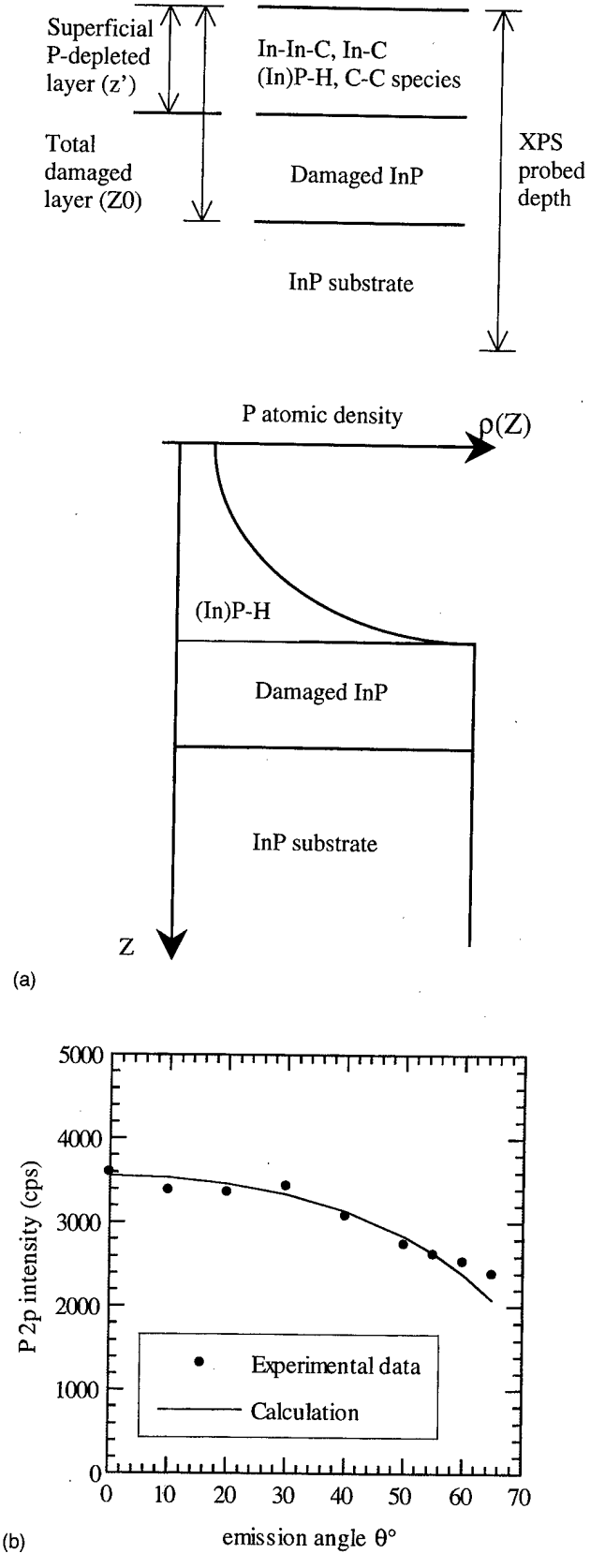


FIG. 7. (a) Representation of the etched surface using a three-layer model: a superficial P-depleted layer in which an exponential P profile is proposed, a damaged stoichiometric InP layer and the InP substrate. (b) Comparison of the experimental and calculated P $2p$ intensities as a function of the emission angle showing the realistic description of the surface in the course of etching by using the three-layer model.

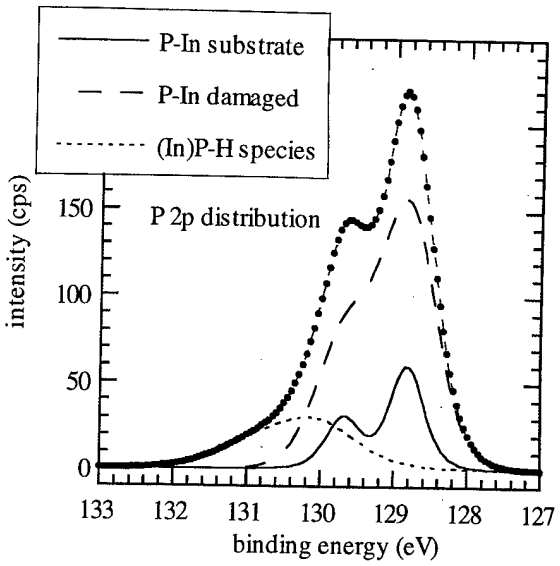


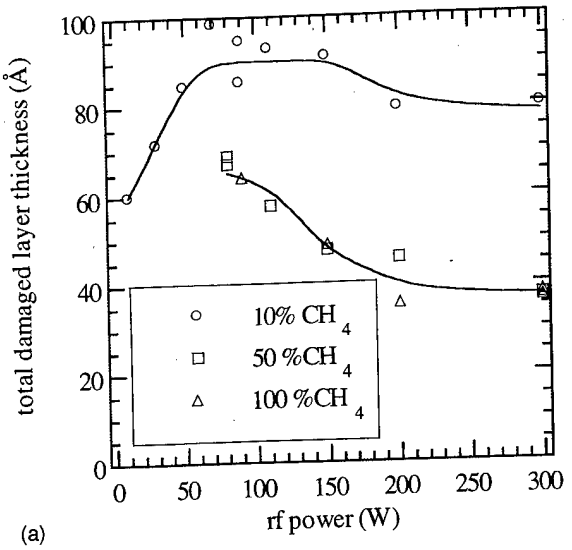
FIG. 8. Decomposition of the P $2p$ spectra using three components including (In)P–H species, P–In bonds originating from the damaged layer, and P–In bonds from the monocrystalline substrate.

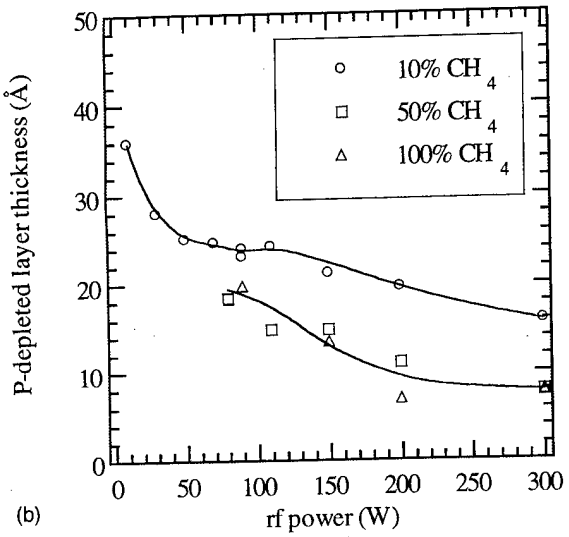
In this relation, the first term is related to the contribution of the superficial P-depleted layer, that is I_P(In)P–H, and the second, to the contribution of the damaged stoichiometric InP layer and the InP substrate.

The decomposition of the P $2p$ spectra into two contributions allows us to deduce the value of the superficial layer thickness z' using Eq. (1), then the exponential factor d is estimated by solving Eq. (4) for $\theta=0$. To be more accurate these determinations have been achieved for several independent etching experiments. This allows us to calculate the theoretical intensity of the P $2p$ distribution as a function of the emission angle θ [Eq. (4)]. A comparison with experimental data is given in Fig. 7(b), and shows a good agreement indicating that our model correctly describes the damaged surface in the course of etching.

3. Layer thickness evaluation

The evaluation of the total damaged layer thickness z_0 is based on the decomposition of the P $2p$ distribution into three components which are presented in Fig. 8. The contribution associated with (In)P–H species located at $+1.2$ eV relative to the main P–In component is unchanged. But the component attributed to P–In is considered as the sum of two contributions: one originates from bulk InP and the second one from the stoichiometric damaged InP material. To operate the curve-fitting of the P $2p$ spectra, the location and the FWHM of the component relative to bulk InP are fixed to the reference values for a monocrystalline sample, 129.1 and 0.57 eV, respectively. The curve-fitting of the P $2p$ spectra then gives the intensities of the different components. A similar operation cannot be performed on the In $4d$ distribution since In–C and In–P components are not easily distinguishable due to their similar binding energy position, and belong to the superficial and the damaged stoichiometric





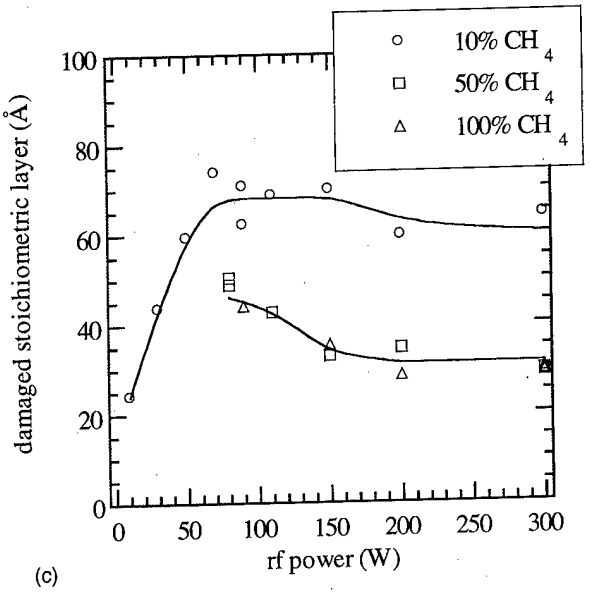


FIG. 9. Thickness evolution of the different layers as a function of power and CH_4–H_2 mixture compositions for the total damaged layer (a), the P-depleted superficial layer (b), and the damaged stoichiometric InP layer (c).

layer, respectively. As a result the decomposition of the In $4d$ spectra is not capable of providing the intensities of the various components.

Using Eq. (1), the total damaged layer thickness z_0 is calculated from the intensity of the bulk InP component deduced from the P $2p$ spectra decomposition. Subtracting z' from z_0 then gives the thickness of the damaged stoichiometric layer. Results are reported in Fig. 9 as a function of the applied power for the different CH_4–H_2 mixtures.

For low power (10–70 W) in the 10% CH_4 plasma condition, the thickness of the total damaged layer increases with power to reach a thickness of about 90 Å. In this range the thickness of the damaged stoichiometric InP layer increases although that of the P-depleted layer decreases. So the lower the power, the more the damage corresponds to a phosphorus depletion. This confirms our interpretation that the In removal depends strongly on the ion assistance. Between 70 and 150 W the total damaged layer thickness remains stable (90 Å), and then slightly decreases down to 80 Å as the power increases to 300 W, in agreement with the decrease observed for both the P depletion and the amorphization of the surface. Results are similar for all mixtures above 50% CH_4. For any power over 80 W, the use of a higher methane proportion leads to a significant decrease of the total damaged layer thickness due to the simultaneous decrease of the P-depleted and damaged stoichiometric InP layer thicknesses. Therefore in terms of chemical damages the lowest thickness is obtained in methane rich mixtures and high power giving values of 37, 8, and 29 Å for the total damaged layer, the P-depleted layer, and for the damaged stoichiometric InP layer, respectively. To fix one's ideas, using our model and the data obtained from XPS, Fig. 10 reports schematic diagrams of the surface etched under typical plasma conditions. It clearly gives an illustration of the influence of the power and of the CH_4 percentage of the mixture on the damages of the surface induced by the plasma process.

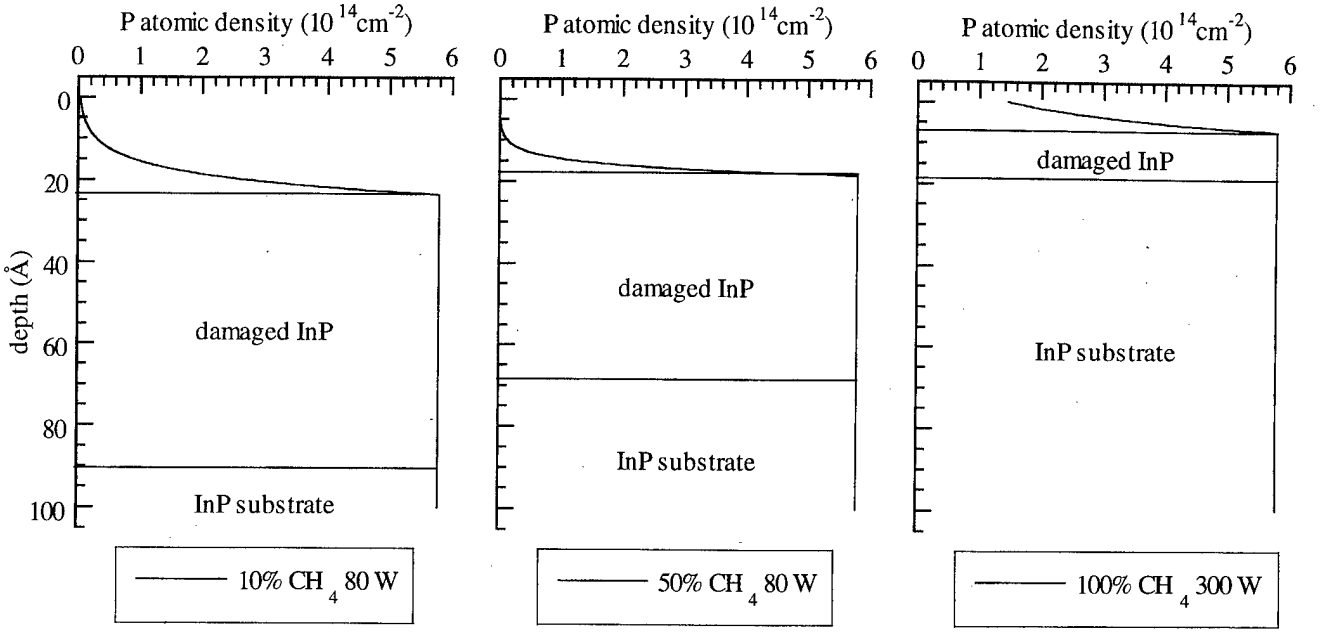


FIG. 10. Schematic diagrams of the InP surface for three typical RIE plasma conditions illustrating the influence of both the power and the CH_4–H_2 mixture composition.

D. Qualitative discussion for an etching model

Let's recall first that H atoms produced in the discharge are the active species responsible for the removal of P from the surface since PH_3 has been evidenced as the major phosphorus etch product, and second that CH_3 radicals are indirectly shown to interact with In and likely form an organoindium volatile product. Under our standard plasma conditions, the surface is found highly P-depleted and damaged. High resolution XPS results show that by increasing both the power and the methane content in the gas mixture, the etched surface quality is greatly improved with lower P depletion and amorphization. Therefore the best result under our RIE conditions is actually obtained in pure methane plasma (15 mTorr) run at a power of 300 W. For any power it is observed that increasing the methane fraction results in the increase of the CH_3 radical density in the plasma phase which in return allows higher ER and significantly improves the surface stoichiometry, although the ion energy flux remains almost constant with the methane content. The other interesting feature comes from the decreasing degradation of the surface with increasing power in the 70–300 W range, and consequently under a higher ion energy flux. And surprisingly, the higher the power, the lower the P depletion and the amorphization of the etched surface. But a much higher decomposition of the gas mixture is obtained when increasing the power, giving a higher CH_3 concentration even though the etch rate and the CH_3 loss increases. Actually the etching of the compound material is controlled by the In removal. The decrease of both the mean surface stoichiometry and amorphization, as the power and the percentage of CH_4 increases, suggest a better balanced interaction between the CH_3 radicals and the ions on the surface. That is likely an increase in the CH_3 flux/ion energy flux ratio, resulting in a more efficient indium etching mechanism. One can say that the etching conditions of In defines the degree of P depletion and amorphization of the surface and as a result the total damaged layer thickness. At low pressure, the etching mechanism is less clear. However, as residual carbon contamination increases and as amorphization decreases, when decreasing the power, it is likely that the CH_3 flux/ion energy flux ratio increases. But, on the contrary to what is observed above 70 W, this increase does not lead to a better surface stoichiometry. It is thus most probable that a "minimum" ion assistance is required to achieve efficient In removal from the surface.

IV. CONCLUSION

Detailed investigations of the RIE process of InP in CH_4–H_2 plasmas were performed as a function of the input power and the gas mixture composition. Plasma phase diagnostics allowed to monitor the etch products and the CH_3 radicals, the active neutral species for In etching. XPS surface analysis gave precise characterization of the damage induced by the plasma exposure, including the P depletion and the degree of amorphization. All these complementary experimental data contribute together for a better understanding of the InP etching. In effect we showed that between 10 and 70 W the ion assistance to the In removal is too weak to allow efficient In etching. As a result, the surface is highly P-depleted but weakly amorphized. On the contrary, increasing the power above 70 W, and the methane fraction, results in a significant improvement of the surface quality as a consequence of the simultaneous decrease of the P depletion and the amorphization of the surface. This is interpreted by a more balanced interaction of the surface with CH_3 radicals and the ion bombardment when using a high power and a

rich methane fraction. The In elimination is the rate limiting step in the etching mechanism of the compound semiconductor and thereby controls the quality of the etched surface.

A model for the representation of the surface in the course of etching is proposed that describes the P depletion and the ion induced crystalline defects. The damaged layer located over the monocrystalline material, which is the result of the plasma–surface interaction, consists of a superficial P-depleted layer and a stoichiometric amorphized InP layer. The thickness of the damaged layer and superficial P-depleted layer are determined for all the plasma conditions studied. A total damaged layer of about 40 Å is obtained under the best plasma conditions.

ACKNOWLEDGMENT

The authors would like to acknowledge France Telecom-CNET (Centre National d'Etude des Télécommunications) for financial support under Contract No. 936B002.

[1]U. Niggebrügge, M. Klug, and G. Garus, Inst. Phys. Conf. Ser. **79**, 367 (1985).
[2]H. Sugimoto, Y. Abe, T. Ohishi, K. Ohtsuka, T. Matsui, H. Yoshiyasu, and Y. Nomura, J. Electrochem. Soc. **139**, 2969 (1992).
[3]S. J. Pearton, in *InP HBT's: Growth, Processing and Applications*, edited by B. Jalali and S. J. Pearton (Artech House, Boston, MA, 1995), p. 47.
[4]C. Y. Park, J. B. Yoo, C. Park, K. S. Hyun, D. K. Oh, Y. H. Lee, C. Lee, and H. M. Park, J. Vac. Sci. Technol. B **13**, 974 (1995).
[5]H. E. G. Arnot, R. W. Glew, G. Schiavini, L. J. Rigby, and A. Piccirillo, Appl. Phys. Lett. **62**, 3189 (1993).
[6]J. E. Schramm, E. L. Hu, J. L. Merz, J. U. Brown, M. A. Melendes, M. A. Thompson, and A. S. Brown, J. Vac. Sci. Technol. B **11**, 2280 (1993).
[7]Y. Kondo, K. Kishi, M. Itoh, H. Oohashi, and M. Yamamoto, 8th International Conference on Indium Phosphide and Related Materials, Schwäbisch Gmünd, Germany, 1996, p. 384.
[8]T. R. Hayes, M. A. Dreisbach, P. M. Thomas, W. C. Dautremont-Smith, and L. A. Heimbrook, J. Vac. Sci. Technol. B **7**, 1130 (1989).
[9]T. R. Hayes, U. K. Chakrabarti, F. A. Baiocchi, A. B. Emerson, H. S. Luftman, and W. C. Dautremont-Smith, J. Appl. Phys. **68**, 785 (1990).
[10]M. C. Peignon, Ch. Cardinaud, and G. Turban, J. Electrochem. Soc. **140**, 505 (1993).
[11]Y. Feurprier, Ch. Cardinaud, and G. Turban, J. Vac. Sci. Technol. B **15**, 1733 (1997).
[12]Y. Feurprier, Ch. Cardinaud, and G. Turban, Plasma Sources Sci. Technol. **6**, 334 (1997).
[13]Y. Feurprier, Ch. Cardinaud, B. Grolleau, and G. Turban, Plasma Sources Sci. Technol. **6**, 561 (1997).
[14]D. A. Shirley, Phys. Rev. B **5**, 4709 (1972).
[15]D. W. Marquardt, J. Soc. Ind. Appl. Math. **11**, 431 (1963).
[16]D. L. Melville, J. G. Simmons, and D. A. Thompson, J. Vac. Sci. Technol. B **11**, 2038 (1993).
[17]P. Kae-Nune, J. Perrin, J. Guillon, and J. Jolly, Plasma Sources Sci. Technol. **4**, 250 (1995).

Germanium etching in high density plasmas for 0.18 μm complementary metal–oxide–semiconductor gate patterning applications

C. Monget[a] and A. Schiltz
France Telecom, CNET/DTM, BP 98, 38243 Meylan Cedex, France

O. Joubert and L. Vallier
France Telecom, CNET/DTM, BP 98, 38243 Meylan Cedex, France and CNRS-LPCM-IMN, 44072 Nantes Cedex 03, France

M. Guillermet
LETI (CEA-Technologies Avancées) DMEL-CEN/G, 38054 Grenoble Cedex 9, France

B. Tormen
SGS-Thomson Microelectronics, 38926 Crolles Cedex, France

(Received 22 January 1998; accepted 3 April 1998)

Oxide masked polysilicon/polygermanium 0.18 μm gates were etched in high density plasma sources. Using gas mixtures of Cl_2 and HBr with O_2 which are commonly used for polysilicon, we observed strong deformation of the poly-Si/poly-Ge gate profiles, whereas perfectly anisotropic etching profiles were obtained for poly-Si gates. A multistep etching recipe was developed allowing anisotropic etching profiles to be obtained while maintaining a good selectivity to the gate oxide when using a Cl_2/N_2 gas mixture. The chemical constituents present on the tops, sidewalls, and bottoms of the etched features were determined by x-ray photoelectron spectroscopy (XPS). XPS analyses have shown that when using a Cl_2/N_2 gas mixture, a thin GeN_x passivation layer is formed on the sidewalls of the poly-Ge features. © *1998 American Vacuum Society.*
[S0734-211X(98)00204-2]

I. INTRODUCTION

Single-crystal silicon–germanium alloys have been extensively studied for device applications. Polycrystalline Ge (referred to as poly-Ge hereafter) is a favorable alternative to the polycrystalline Si (poly-Si) gate electrode for complementary metal–oxide–semiconductor (CMOS) technology. The advantage of using poly-Ge, a so-called midgap material, is that the Fermi level is brought to the edge of the valence band of Ge by P^+ doping [single gate for N and P metal–oxide–semiconductor field-effect transistors (MOSFETs)]. In this way, the gate Fermi level becomes equal to that of the silicon midgap due to the fact that the forbidden gap of Ge is only half of that of Si.[1] The latter suggests that ideal midgap operation is only possible with pure Ge gates. In previous papers,[2,3] we reported the etching of oxide masked poly-Si/poly-$Si_{1-x}Ge_x$ gates in a high density plasma helicon source. In particular, we have optimized a two-step etching recipe using Cl_2/O_2, which provided perfectly anisotropic 0.18 μm features with a minimal consumption of the gate oxide. However, this process was not well adapted for $Si_{1-x}Ge_x$ alloys with high Ge contents: lateral etching was observed for bilayer gates with Ge concentration higher than 75%. X-ray photoelectron spectroscopy (XPS) analyses demonstrated that the passivation layer which forms on the sidewalls of the poly-$Si_{1-x}Ge_x$ alloy is just made of silicon etching products. As a consequence, for high Ge concentration in the alloy, the silicon-based passivation layer is not thick enough to prevent lateral etching of the sidewalls of the gates by the highly reactive species of the discharge.

In this article, the etching of poly-Si/poly-Ge bilayer gates masked with oxide patterns is studied using Cl_2-based chemistries in two different high density plasma etchers. We have also used *quasi in situ* XPS to determine the chemical composition of poly-Ge features after etching.

II. EXPERIMENTAL PROCEDURES

Etching experiments were conducted in two high density plasma etchers. The first one is an helicon source connected to a cluster tool dedicated to the characterization of etching processes (described in more detail in a previous article[4]). It is composed of a load-lock chamber, a transfer chamber, a reactor chamber, and an XPS analysis chamber. The helicon source was made by Lucas Labs®[5,6] and is operated at a rf frequency of 13.56 MHz. The sample can be independently biased using a 600 W maximum power supply at 13.56 MHz. The plasma generation region is equipped with two solenoid magnets which are necessary to initiate the plasma, whereas two other solenoid magnets around the diffusion region maintain a high plasma density on the sample. The tuning of the matching network is computer controlled and performed automatically. In all experiments, the source power was fixed at 2500 W in order to ensure the best etching uniformity and a high plasma density (around 3×10^{11} cm^{-3}).[7] The etching process was monitored in the helicon source using *in situ* ellipsometry.

Etching experiments were also performed in a high density plasma etcher commercialized by LAM® (TCP 9400

[a]Electronic mail: cedric.monget@cnet.francetelecom.fr

SE). The rf power is inductively coupled to the plasma using a flat coil located on top of a 30 cm diameter quartz window (the distance between the coil and the sample is around 10 cm). The sample is clampled on the bottom electrode using an electrostatic chuck. The temperature control is ensured using He backside cooling (pressure: 8 Torr). In all experiments, the temperature of the electrode was maintained at 65 °C. This type of discharge can generate high density plasmas (ion density $>10^{11}$ cm^{-3}) in a pressure range comprised between 2 and 70 mTorr. The etching apparatus is connected to a multidiagnostic SOFIE® platform allowing, in particular, real time monitoring of optical emission signals. Using the first derivative of the optical emission signal obtained by a high-resolution monochromator (theoretical resolution around 1 Å), very accurate and fast optical end point detections can be obtained.

XPS analyses of samples etched in the helicon source were performed *quasi in situ*. After etching, the sample was immediately transferred under vacuum into the XPS analysis chamber. A special mask, already described in detail elsewhere,[6] was designed for XPS analyses. The analyzed areas consisted of regular arrays of trenches, blanket substrate, and unpatterned mask material. As the area of each array is at least 1.5 mm^2, the x-ray beam can thus be focused on one individual array, the resulting signal being the average of many identical lines and spaces. XPS analyses were conducted with the lines aligned perpendicularly or in parallel with the electron energy analyzer. In the first case, the analyzed area is the top and, depending on the space between lines, parts of the sidewalls of the features.[6] When the lines are aligned in parallel with the electron energy analyzer, the analyzed areas are the tops and bottoms of the features. An electron gun emitting low energy electrons (less than 80 eV) was used to charge analyzed surfaces, allowing contributions from insulating surfaces (oxide hard mask) and conducting surfaces (poly-Ge on the bottom and sidewalls of the features) to be differentiated. Chemical composition were derived from the areas of the different XPS spectra. Spectral deconvolution was performed to extract the Ge 3*d*, Si 2*p*, O 1*s*, N 1*s*, Cl 2*p* peak intensities. Individual line shapes were simulated with the convolution of Lorentzian and Gaussian functions. A Shirley function was used to perform background substraction.[8] The elemental Si 2*p* and Cl 2*p* peaks were fitted with a doublet constituted of the two spin-orbit components 2*p*1/2 and 2*p*3/2. Fits were performed forcing an identical full width at half maximum (FHWM) for the two spin-orbit components and an intensity ratio of 0.52 for Si 2*p* and 0.5 for Cl 2*p*. For germanium related analyses, the intensity ratio of the doublet of the elemental Ge 3*d* peak was fixed at 0.66. We verified that a doublet separation of 0.6 eV for elemental Si 2*p* and Ge 3*d*, and 1.6 eV for Cl 2*p* was obtained.[9–11] After XPS analyses, the integrated intensities were divided by the theoretical Scofield cross sections (0.82 for Si 2*p*, 1.42 for Ge 3*d*, 2.29 for Cl 2*p*, 2.93 for O 1*s*, and 1.80 for N 1*s*).[12] As the angle between all of the analyzed areas and the electron energy analyzer is 45°, direct comparisons can be drawn between the coverage of the different elements present on the tops, sidewalls, and bottoms of the features.

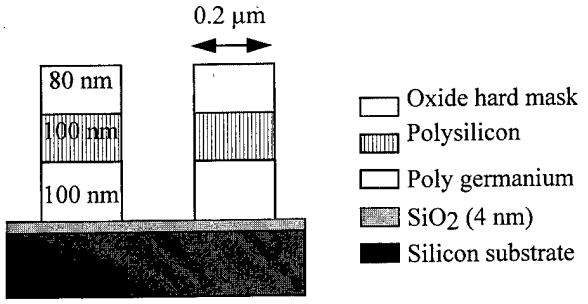


FIG. 1. Schematic view of the poly-Si/poly-Ge multilayer gate stack structure.

Samples were 200 mm diameter wafers covered with different layers. Polysilicon and polygermanium layers were grown by plasma-enhanced chemical vapor deposition (PECVD) using a commercial single wafer cluster tool. The pressure in the reaction chamber was about 80 Torr and the temperature was fixed at 450 °C. The gases used were SiH_4 (for poly-Si growth) and GeH_4 (for poly-Ge growth) with H_2 as a carrier gas. For the etching experiments, a 100 nm poly-Ge layer was grown on top of a 4 nm thin gate oxide. The polysilicon layer grown on top of the poly-Ge was 100 nm thick. Oxide patterns (80 nm thick) were fabricated using conventional lithographic and etching steps (see Fig. 1). For the XPS experiments, a 100 nm thick SiO_2 layer was deposited by PECVD on top of bare silicon. The poly-Ge layer was 500 nm thick, whereas the poly-Si was only 80 nm thick (see Fig. 2). This design was optimized for sidewall analysis: the poly-Ge layer thickness was increased to facilitate sidewall composition analysis, whereas the poly-Si layer simply acts as a protecting layer against oxidation.

III. SUMMARY OF OXIDE MASKED POLY-$Si_{1-x}Ge_x$ ETCHING PROCESS

In a previous work, the patterning of poly-Si/poly-$Si_{0.45}Ge_{0.55}$ bilayer gates has been studied in a low-pressure,

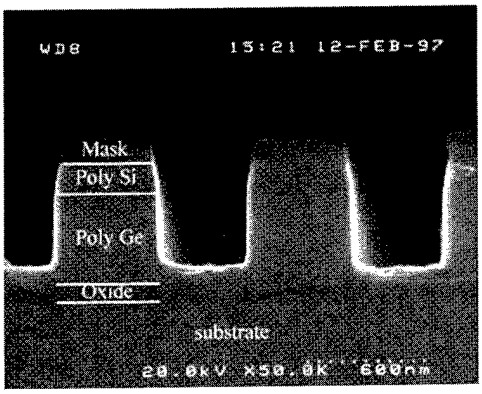


FIG. 2. Poly-Si (80 nm)/poly-Ge (500 nm) structures (0.5 μm line/space) partially etched using a Cl_2/N_2 gas mixture. The remaining thickness of poly-Ge is about 100 nm.

TABLE I. Process conditions used in the helicon source.

Chemistry	Step	Bias (W)	He flow (sscm)	Cl_2 flow (sccm)	O_2-He flow (sccm)	SiH_2Cl_2 flow (sccm)	N_2 flow (sccm)
Cl_2/He	Main etch	150	40	80	···	···	···
Cl_2/He	Lower bias[a]	60	40	80	···	···	···
Cl_2/He	Overetch[b]	60	40	80	···	···	···
Cl_2/SiH_2Cl_2	Main etch	150	···	60	···	20	···
Cl_2/SiH_2Cl_2	Lower bias[a]	60	···	60	···	20	···
$Cl_2SiH_2Cl_2$	Overetch[b]	60	···	60	···	20	···
Cl_2/N_2	Main etch	150	···	73	···	···	7
Cl_2/N_2	Lower bias[a]	60	···	73	···	···	7
Cl_2/N_2	Overetch[b]	60	···	73	40	···	···

[a]Started at the poly-Si/ poly-Ge interface.
[b]Started at the SiGe/oxide interface; duration: 15 s.

high density plasma helicon source.[2] Different gas mixtures used for poly-Si etching (Cl_2, HBr, and O_2 mixtures) have been investigated for poly-$Si_{1-x}Ge_x$ etching. A process using a Cl_2/O_2-He gas mixture was developed allowing anisotropic etching profiles of 0.2 μm bilayer poly-Si/poly-$Si_{0.45}Ge_{0.55}$ gates to be obtained. In order to minimize structural defects formation at the edges of the gate, the bias power was decreased 40 nm before reaching the SiGe/gate oxide interface. Real time ellipsometry was used to control the time at which the bias power was decreased. We have also shown that using the same process, lateral etching was observed again when etching bilayer gates with a Ge concentration in the alloy higher than 75%.

Chemical topography of the features etched was studied using XPS.[3] Analyses have shown that using a Cl_2/O_2 chemistry, the passivation layer formed on the sidewalls of the features during the etching is a silicon–oxide-like film. No germanium–oxide-like etching products were detected on the sidewalls of the features, showing that only silicon etching products are involved in the formation of the passivation layer. As a consequence, the thickness of the passivation layer decreases when the germanium concentration in the alloy increases. When bilayer gates with germanium concentrations in the alloy higher than 75% were etched, the silicon–oxide-like layer formed on the sidewalls of the features was too thin (<1 nm) to prevent lateral etching by reactive species of the plasma.[3]

IV. OPTIMIZATION OF POLYSILICON/ POLYGERMANIUM GATE ETCHING PROCESS

A. Preliminary studies in the helicon source

The helicon source was operated using the maximum rf power of 2500 W to achieve the best etching uniformity (about 10% at 3σ for the investigated operating conditions). The pressure was fixed at 2 mTorr. For each gas mixture investigated (see Table I), a short breakthrough step (5 s) at high bias power (400 W) was used to remove the thin native oxide from the polysilicon surface. The main etching step of the process was then separated into two steps. During the polysilicon gate etch step, the bias power was fixed to 150 W. When reaching the Ge layer, the chuck bias power was then decreased to 60 W. The introduction of the low energy etching step before reaching the gate oxide minimizes the trenching at the edges of the gate[13] and increases the selectivity to the gate oxide. When the poly-Ge/SiO_2 interface was reached, the overetch step was started using the same chuck bias power (60 W).

1. Cl_2/He chemistry

In a previous work, it was demonstrated that, using the two-step etching process described above, 0.25 μm oxide masked poly-Si gates can be anisotropically etched using pure chlorine in the gas phase of the plasma.[14] XPS analyses of oxide and resist masked poly-Si features have shown that a thin oxide-like film (less than 1 nm) was present on the sidewalls of polysilicon.[14] The oxide-like layer was attributed to the erosion of the quartz tube located in the plasma generation region of the helicon source. The oxide-like passivation layer being not thick enough to prevent spontaneous etching, the anisotropy of the etching was possible because of the low spontaneous etching rate of polysilicon by the chlorine atoms of the discharge.

The same chemistry was used to etch oxide masked poly-Si/poly-Ge gates using the process described in Table I. The Cl_2 flow was 80 sccm and the He flow 40 sscm. Lateral etching was observed in the poly-Ge layer (Fig. 3) showing that spontaneous etching reactions between poly-Ge and chlorine atoms are greater than between polysilicon and chlorine atoms. On the other hand, no deformation of the etching profiles was observed after etching the bilayer gate using high chuck bias power conditions (150 W) showing that spontaneous etching reactions between chlorine and poly-Ge mainly occur during the low energy etching step and the overetch step of the process.

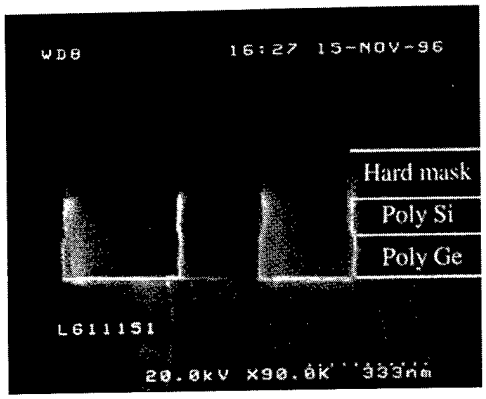


FIG. 3. Cross-sectional view of 0.2 μm poly-Si/poly-Ge gates etched using a Cl_2/He chemistry. Profile deformation due to lateral etching is observed in the poly-Ge layer.

2. Cl_2/SiH_2Cl_2 chemistry

We have seen in previous studies that, contrary to polysilicon, oxygen addition to chlorine does not promote the formation of a sidewall passivation layer on the poly-Ge.[3] The addition of SiH_2Cl_2 to chlorine was investigated in order to add a silicon-containing gas into the feed gas stock of the plasma which could lead to the formation of a silicon–oxide-like layer to passivate the sidewalls of the poly-Ge. The Cl_2/SiH_2Cl_2 gas mixture was used under the process conditions described in Table I. However, even if straight sidewalls were obtained in the polysilicon layer, lateral etching was still observed in the poly-Ge layer (see Fig. 4).

3. Cl_2/N_2 chemistry

As nitrogen may react with germanium to form germanium nitride-based products, N_2 was added to Cl_2 to promote the formation of a passivation layer on the poly-Ge sidewalls. Process conditions are summarized in Table I. After a short breakthrough step, the poly-Si layer was etched under high bias power conditions (150 W) using the Cl_2/N_2 gas mixture. When the poly-Ge layer was reached, the chuck bias power was decreased to 60 W, all other conditions being

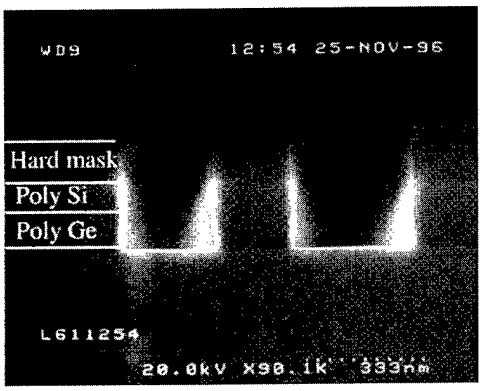


FIG. 4. Cross-sectional view of 0.2 μm poly-Si/poly-Ge gates etched using a Cl_2/SiH_2Cl_2 chemistry. Profile deformation due to lateral etching is observed in the poly-Ge layer.

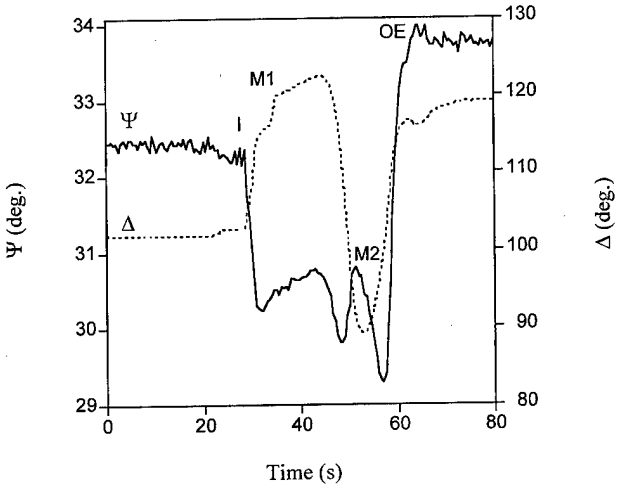


FIG. 5. Real time traces of Ψ (solid line) and Δ (dotted line) recorded at 290 nm during the etch of a poly-Si (80 nm)/poly-Ge (120 nm)/4.5 nm gate oxide structure masked with oxide patterns. Marks I, M1, M2, and OE indicate plasma ignition, transition to the main etching step, transition from poly-Si to poly-Ge and transition to the overetch step, respectively.

unchanged. An ellipsometry curve recorded in real time at a wavelength of 290 nm during the poly-Si/poly-Ge gate stack process is presented in Fig. 5. The first mark (indicated as I) corresponds to the beginning of the process. After a short breakthrough at 400 W, the main etch step starts (mark M1). The transition between the poly-Si and the poly-Ge is reached at mark M2. In order to avoid gate oxide consumption and defects formation in the active areas, main etch step 2 was started at mark M2. The gate oxide is reached at mark OE and the overetch step is started. Figure 6 shows 0.2 μm oxide masked poly-Si/poly-Ge features etched using this process. Despite a slightly sloped oxide hard mask, anisotropic etching profiles were obtained.

The selectivity between poly-Ge and the gate oxide was increased by replacing N_2 by an O_2-He gas mixture during the overetch step. The O_2-He flow was fixed at 40 sccm (8 sccm equivalent O_2). Under the conditions used during the overetch step (source: 2500 W, chuck: 60 W), the selectivity

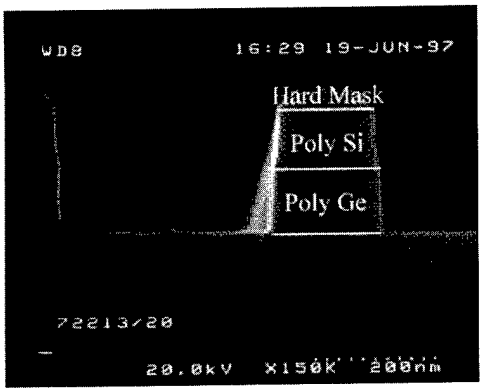


FIG. 6. Scanning electron micrograph of 0.2 μm poly-Si/poly-Ge gates after etching in the helicon source using the final process described in Table I. SEM observation were performed after a short HF rinse (inducing a slight erosion of the oxide hard mask as shown on the picture).

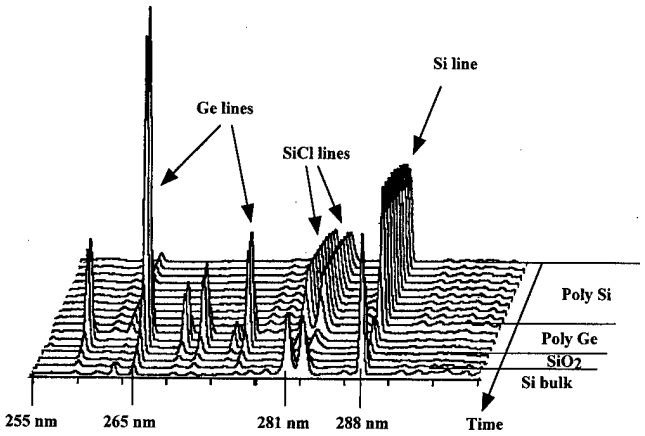


FIG. 7. Optical emission spectra recorded between 250 and 300 nm during the poly-Si/poly-Ge gate process in the LAM TCP 9400 SE. During the etching of the poly-Si layer, emission lines due to SiCl (281 nm) and Si (288 nm) are detected. When the poly-Ge layer is etched, the strongest emission is attributed to Ge species (265 nm).

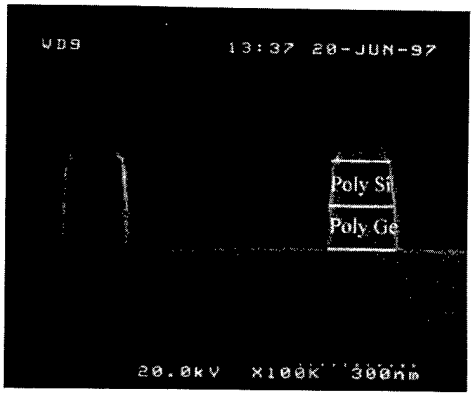


FIG. 8. Scanning electron micrograph of poly-Si/poly-Ge gates after etching in the LAM TCP 9400 SE. SEM observation were performed after a short HF rinse (inducing a slight erosion of the oxide hard mask as shown on the picture).

between poly-Ge and SiO_2 was higher than 100:1, whereas using Cl_2/N_2 the selectivity was less than 40:1. Despite N_2 exclusion during the overetch step, anisotropic etching profiles were obtained for poly-Si/poly-Ge gates down to 0.2 μm features showing that the passivation film which was built up on poly-Ge sidewalls during the main etch step was thick enough to prevent spontaneous etching reactions during the overetch step.

B. Process studies in the LAM TCP etcher

Based on the results obtained in the helicon source, N_2 addition was also used in the LAM TCP to favor the formation of a passivation film on the sidewalls of the poly-Ge features. As in the helicon source, the process designed was a multistep etching recipe. Prior process optimization in the LAM TCP, optical emission spectra were recorded in real time between 255 and 300 nm (see Fig. 7) during the etching of blanket wafers covered with poly-Si (100 nm)/poly-Ge (100 nm)/SiO_2 (4 nm) layers. During the etching of polysilicon, optical emission lines originating from Si (288 nm) and SiCl (281 nm) species are detected. At the end of the poly-Si layer etching, the emission originating from silicon-based species decreases strongly, whereas emission originating from germanium species (265.15 nm) becomes important when the etching of the polygermanium layer starts. At the poly-Ge/SiO_2 transition, the intensity of the 265 nm line decreases strongly (emission at 288 and 281 nm keep very low). When the thin gate oxide (4 nm) has been etched away, emission intensities corresponding to silicon-based species (288 and 281 nm) start growing again, showing that bulk silicon is etched. As a consequence of those experiments, the first derivative of the optical emission line at 288 nm (silicon species) was used as an end point detection for poly-Si etching and the first derivative of the optical emission line at 265 nm (germanium species) was used as an endpoint detection for poly-Ge etching. In a first step, the poly-Si layer was etched using a $HBr/Cl_2/N_2$ gas mixture (70/70/5) at a pressure of 10 mTorr. The rf power injected in the source was 400 W (a typical source power for poly-Si etching) which allows an ion density of around 10^{11} cm^{-3} to be obtained (a much higher rf power is injected in the helicon source to obtain similar densities on the wafer because the wafer is located in the downstream plasma region, whereas in the LAM TCP, the wafer is located in the source region). The chuck power was fixed at 130 W. At the polysilicon end point detection, an overetch time of 10% was applied using the same process conditions. The poly-Ge etch was then performed using Cl_2/N_2 mixture (60/10) at a pressure of 4 mTorr and a chuck bottom power of 50 W (rf power injected in the source was maintained at 400 W). At the end of the poly-Ge layer etch, the overetch step was started using HBr/O_2 (60/2) at a pressure of 30 mTorr. The source power was decreased to 300 W and chuck bias power to 30 W (these conditions are simply those used for the standard poly-Si etch process). Using the series of steps described above, anisotropic etching profiles of 0.2 μm poly-Si (100 nm)/poly-Ge (100 nm) gates were obtained (see Fig. 8).

The process developed in the LAM TCP shows that the passivation layer which builds up on the poly-Ge sidewalls using the Cl_2/N_2 gas mixture is thick enough to prevent lateral etching of poly-Ge during the overetch step. Figure 9 shows 0.2 μm poly-Si/poly-Ge gates running over topography. In this case, the isolation between devices are performed using a 0.25 μm CMOS shallow trench isolation (STI) process. No structural defects (such as pitting of the active areas) are observed at the edges of the gates, showing that the introduction of the low energy etching step and the use of the HBr/O_2 chemistry during the overetch minimizes the gate oxide consumption during the process.

V. CHEMICAL DISTRIBUTION ANALYSES OF OXIDE MASKED POLY-Ge GATES ETCHED USING Cl_2/N_2 CHEMISTRY

In order to analyze the chemical composition of the sidewall passivation layer, samples dedicated to XPS analyses (described in Sec. II) were etched in the helicon source using

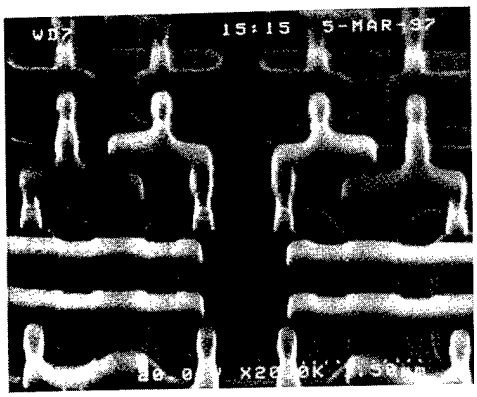


FIG. 9. Poly-Si/poly-Ge gates etched in the LAM TCP and running over topography. Isolation between devices was performed using a shallow trench isolation process.

the Cl_2/N_2 gas mixture. The etching was performed using the high energy main etching step of the process (chuck bias power: 150 W) and stopped in the poly-Ge layer 100 nm before reaching the gate oxide. After etching, samples were transferred under vacuum into the XPS analysis chamber. Figure 2 shows the 0.5 μm L/S gates etched in the helicon source and analyzed by XPS.

A. Chemical analyses of the bottom of the poly-Ge features

Figures 10 and 11 show the Ge 3*d* spectra recorded in a blanket poly-Ge area and in 0.5 μm L/S features with the lines aligned in parallel with the electron energy analyzer (in this configuration, the analyzed areas are the tops and bottoms of the features). In Fig. 10, the two peaks located at 29.3 and 29.9 eV are attributed to the Ge 3*d* doublet of elemental germanium. No $GeCl_x$ species or GeN_x species (located at higher binding energies) were detected on the blanket poly-Ge area. In Fig. 11, the Ge 3*d* spectra recorded with the charge neutralizer turned on exhibits three peaks: the doublet located at 29.4 and 30 eV is attributed to elemental germanium. The peak located at 31.9 eV, which is unshifted when the charge neutralizer is turned on, is originating from the conductive poly-Ge at the bottom of the features and attributed to germanium-nitride-like species (the binding energy corresponds to the binding energy of Ge atoms in

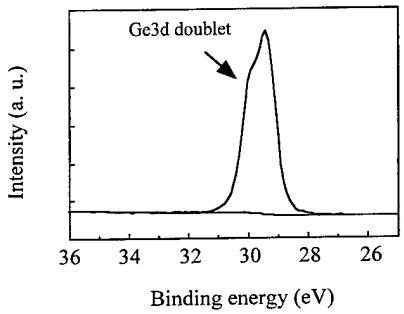


FIG. 10. Ge 3*d* XPS spectra recorded on a blanket poly-Ge area. No GeN_x species are detected on the surface.

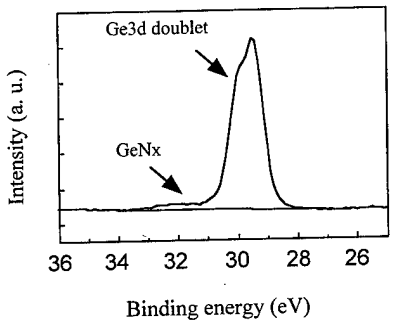


FIG. 11. Ge 3*d* spectra recorded in 0.5 μm L/S lines aligned in parallel with the electron energy analyzer and with the charge neutralizer turned on. GeN_x species are detected on the germanium surface at the bottom of the features.

GeN_4). This result is also in good agreement with Hellman *et al.*: they observe that the Ge 3*d* peak is shifted by 1.9 eV when germanium is exposed to the N_2^+ beam.[15] Zanatta *et al.* found a chemical shift per bound of about 0.5 eV for the Ge–N bond which confirmed that the peak located at 31.9 eV can be attributed to GeN_x ($x=3$) species.[16] When the flood gun was off, another XPS peak was detected at 34.9 eV. This peak was shifted towards lower binding energies and considerably broadened when the flood gun was on and was attributed to germanium-nitride species deposited on top of the oxide hard mask during the etching process.

These results were confirmed by N 1*s* XPS spectra (not shown here) analyses recorded in blanket Ge areas and at the bottom of 0.5 μm L/S features. No N 1*s* signal originating from the blanket poly-Ge area was detected. In 0.5 μm L/S lines features, one N 1*s* peak (located at 399.5 eV with the charge neutralizer turned on and off) was originating from the bottom of the features and confirmed the presence of germanium-nitride-like species at the bottom of the features. Another XPS peak (shifted from 405 to 395 eV when the flood gun was turned on) was indicative of the presence of nitrogen on top of the oxide hard mask (confirming that germanium-nitride-based species were deposited on top of the mask during the etching process).

Figure 12 shows the Cl 2*p* spectra recorded on a blanket poly-Ge area with the charge neutralizer turned on. The

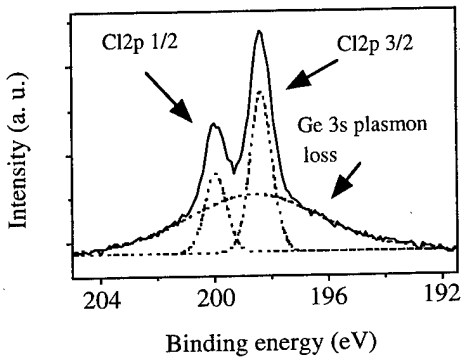


FIG. 12. Cl 2*p* XPS spectra recorded on the blanket poly-Ge. The peaks located at 198.4 and 200 eV are originating from the Cl 2*p* doublet, whereas the broad peak is attributed to Ge 3*s* plasmon loss.

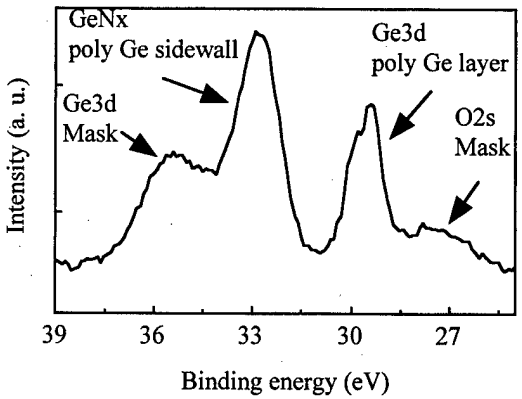


FIG. 13. Ge 3*d* XPS spectra recorded on 0.5 μm L/S features with the lines aligned perpendicularly to the electron energy analyzer and the charge neutralizer turned off.

Cl 2*p* spectra can be divided into three peaks. The Cl 2*p* doublet is located at 198.4 and 200 eV. The broad peak at 198.6 eV is attributed to the Ge 3*s* plasmon loss. XPS spectra (not shown here) recorded in 0.5 μm L/S poly-Ge gates with the charge neutralizer turned on show, as with Ge 3*d*, a clear separation between Cl 2*p* originating from the top of the features (the oxide hard mask) and the bottom of the features (poly-Ge).

Similarly to Ge 3*d*, N 1*s* and Cl 2*p*, O 1*s* and Si 2*p* spectra were recorded in blanket Ge areas and in 0.5 μm L/S poly-Ge features. As above, a perfect differentiation between signals originating from the top of the features (the oxide hard mask) and the bottom of the features (poly-Ge) was possible when the charge neutralizer was turned on.

B. Chemical analyses of the sidewalls of the poly-Ge features

Figure 13 shows the Ge 3*d* spectra recorded in 0.5 μm L/S poly-Ge features with the lines aligned perpendicularly to the electron energy analyzer (in this case, the signal originates from the top and sidewall of the oxide hard mask and the sidewalls of the features). When the charge neutralizer was off (Fig. 13), 5 peaks were detected. One peak, located at 27.7 eV was shifted to lower binding energy when the charge neutralizer was turned on. This contribution was attributed to the O 2*s* XPS peak originating from the oxide hard mask. The doublet located at 29.4 and 30 eV (with the charge neutralizer off and on) was attributed to the elemental germanium of the poly-Ge sidewalls. The XPS peak located at 32.8 eV with the charge neutralizer off and slightly shifted (to 31.9 eV) when the charge neutralizer was turned on, was originating from the conductive Poly-Ge sidewalls and attributed the presence of a germanium-nitride-like layer. The peak originally located at 35.3 eV when the charge neutralizer was off, and shifted to a binding energy of 25 eV when the charge neutralizer was on, was attributed to germanium-nitride-like species deposited on top of the oxide hard mask during the etching process.

Cl 2*p* XPS spectra were also recorded on the same sample (not shown here). The spectrum was fitted into four peaks.

TABLE II. Element concentrations determined on the surfaces of oxide masked poly-Ge features etched using the main etching step of the Cl_2/N_2 process.

Analyzed area	Ge 3*d*	GeN_x or GeO_x	N 1*s*	Cl 2*p*	SiO_x	O 1*s*
Poly-Ge blanket	91.2	...	...	8.8	...	...
Poly-Ge bottom of the features	84.4	4.1	1.5	9.9	...	...
Oxide mask top	...	3.6	2.7	8.6	35.4	49.7
Poly-Ge sidewalls of the features	15.9	21.6	9.2	34.6	8.8	5.1

The doublet located at 201.3 and 199.8 eV with the charge neutralizer off was shifted to 200.5 and 198.9 eV when the flood gun was on. These peaks were attributed to the Cl 2*p* signal originating from the sidewalls of the poly-Ge. The two other peaks (located at 203.7 and 202.6 eV with the charge neutralizer off), were shifted to lower binding energy when the charge neutralizer was turned on. These peaks were therefore attributed to Cl species originating from the oxide hard mask.

Similarly to Ge 3*d* and Cl 2*p*, N 1*s*, O 1*s*, and Si 2*p* spectra were recorded with the lines aligned perpendicularly to the electron energy analyzer. Using the capability of turning the charge neutralizer on and off, a perfect differentiation between Cl 2*p*, N 1*s*, O 1*s*, and Si 2*p* XPS signals originating from the oxide hard mask and the poly-Ge sidewalls was possible.

The quantitative coverage of the different elements derived from the spectra presented above are summarized in Table II. The integrated intensities were divided by the theoretical Scofield cross sections.[12] As the angle between all the analyzed areas and the electron energy analyzer is 45°, direct comparisons were drawn between the coverage of the different elements present on the tops, sidewalls, and bottoms of the features.

VI. DISCUSSION

Chemical topography analyses of oxide masked poly-Si/poly-Ge features etched in a Cl_2/N_2 gas mixtures show the presence of germanium-nitride species on the poly-Ge surfaces at the bottom of the features (4%), on top of the oxide hard mask (3.6%), and on the poly-Ge sidewalls (22%). The results indicate that the Cl_2/N_2 mixture favors the formation of germanium-nitride etching products which are formed on bombarded germanium surfaces during the etch process, sputtered away into the gas phase of the plasma by the high energy ions bombarding the etched surfaces and redeposited on all surfaces exposed to the plasma. This chemistry may be classified as "polymerizing." Here, the term polymerizing simply means that a significant partial pressure of etching products are generated by the etch and redeposited on surfaces exposed to the plasma (the parallel can be made when etching polysilicon gates in $HBr/Cl_2/O_2$, HBr/O_2, Cl_2/O_2

chemistries, where similarly, silicon–oxide-like products are generated by the etch and allow the formation of a passivation layer on the poly-Si sidewalls[17–19]).

The germanium-nitride-like layer formed on surfaces exposed to the plasma is thicker on the unbombarded poly-Ge sidewalls surfaces than on surfaces exposed to the highly energetic ion bombardment (poly-Ge surfaces at the bottom of the features and SiO_2 surfaces on top of the mask). The detection of the elemental germanium originating from the poly-Ge sidewalls when analyses are performed in the perpendicular mode shows that the thickness of the passivation layer is less than the thickness probed by XPS. Taking into account the angle between the electron energy analyzer and the analyzed areas (45°) and the escape depth of Ge 3*d* photoelectrons ($\lambda = 3$ nm),[20] the thickness probed on the sidewalls is approximately 7 nm (which means that the germanium-nitride-like layer is less than 7 nm thick).

VII. CONCLUSION

Poly-si/poly-Ge gates were etched using a multistep etching recipe with various etch chemistries in low-pressure, high density plasma sources. Strong profile deformation of poly-Si/poly-Ge gates were observed when using pure Cl_2, Cl_2/O_2, and Cl_2/SiH_2Cl_2 gas mixtures. In all cases, straight sidewalls were observed in the poly-Si layer whereas lateral etching was observed in the poly Ge layer showing that using those chemistries, no or unsufficient passivation protects the poly-Ge sidewalls.

Anisotropic etching profiles of poly-Si (100 nm)/poly-Ge (100 nm) with a minimal gate oxide consumption were obtained in both high density etchers using Cl_2/N_2 gas mixture. In the processes developed here, the poly-Si layer was etched using plasma operating conditions similar to those used for poly-Si gates, whereas the poly-Ge layer was etched using low chuck bias power conditions. The overetch step was also performed using process conditions (chemistries and plasma operating parameters) similar to those developed for poly-Si gate etching, showing that the germanium-nitride-like passivation layer formed on the poly-Ge sidewalls when using the Cl_2/N_2 gas mixture is thick enough to prevent spontaneous etching reactions between neutral chlorine species and poly-Ge sidewalls.

The chemical topography of the poly-Ge gates after etching in the helicon source was investigated by XPS. XPS analyses show that a GeN_x passivation layer forms on the sidewalls of the poly-Ge features during the etching process. Germanium-nitride species are also detected on top of the oxide hard mask and bottom of poly-Ge features showing that a significant partial pressure of etching products are redeposited on all the surfaces of the features as the etching proceeds.

ACKNOWLEDGMENTS

This work has been carried out within the GRESSI consortium between CEA-LETI and France Telecom-CNET.

[1]T. Skotnicki, P. Bouillon, R. Gwoziecki, A. Halimaoui, C. Mourrain, I. Sagnes, J. L. Regolini, O. Joubert, M. Paoli, and P. Schiavone, Proceedings of the 27th European Solid-state Device Research Conference (ESSDERC), Stuttgart, Germany, 1997.
[2]S. Vallon, C. Monget, O. Joubert, L. Vallier, F. H. Bell, M. Pons, J. L. Regolini, C. Morin, and I. Sagnes, J. Vac. Sci. Technol. A **15**, 1874 (1997).
[3]C. Monget, S. Vallon, F. H. Bell, L. Vallier, and O. Joubert, J. Electrochem. Soc. **144**, 2455 (1997).
[4]F. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 96 (1996).
[5]Lucas Labs Inc., 470C Lakeside Drive, Sunnyvale, CA.
[6]N. Benjamin, B. Chapman, and R. Boswell, SPIE, Santa Clara, CA (unpublished, 1990).
[7]F. H. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 1796 (1996).
[8]D. A. Shirley, Phys. Rev. B **5**, 4709 (1972).
[9]G. Hollinger and F. J. Himpsel, Appl. Phys. Lett. **44**, 93 (1984).
[10]G. Hollinger and F. J. Himpsel, J. Vac. Sci. Technol. A **1**, 640 (1983).
[11]A. R. Zanatta and I. Chambouleyron, Phys. Status Solidi B **193**, 399 (1996).
[12]J. H. Scofield, J. Electron Spectrosc. Relat. Phenom. **8**, 129 (1976).
[13]O. Joubert and F. H. Bell, J. Electrochem. Soc. **144**, 1855 (1997).
[14]F. H. Bell and O. Joubert, J. Vac. Sci. Technol. B **14**, 2493 (1996).
[15]O. C. Hellman, O. Vancauwenberghe, N. Herbots, J. Olson, R. J. Culbertson, and W. J. Croft, Mater. Sci. Eng., B **12**, 53 (1992).
[16]A. R. Zanatta and I. Chambouleyron, Phys. Status Solidi B **193**, 399 (1996).
[17]F. H. Bell and O. Joubert, J. Vac. Sci. Technol. B **15**, 88 (1997).
[18]K. V. Guinn and V. M. Donnelly, J. Appl. Phys. **75**, 2227 (1994).
[19]K. V. Guinn, C. C. Cheng, and V. M. Donnelly, J. Vac. Sci. Technol. B **13**, 214 (1995).
[20]*Practical Surface Analysis by Auger and X-Ray Photoelectron Spectroscopy*, edited by D. Briggs and M. P. Seah (Wiley, New York, 1984).

Estimation of the activation energy for Ar/Cl_2 plasma etching of InP via holes using electron cyclotron resonance

E. W. Sabin[a)]
TRW, Redondo Beach, California 90278

(Received 3 March 1998; accepted 22 May 1998)

The temperature dependence of the via hole etch rate is reported for argon/chlorine plasma etching of indium phosphide in an electron cyclotron resonance etcher. The indium phosphide via hole etch rate was first established for several wafer chuck temperatures. Then the temperature of the wafer was estimated by measuring the heating effect of the plasma and the cooling effect of the helium backside cooling. The wafer temperature was then substituted for the wafer chuck temperature. The etch rate as a function of wafer temperature was found to exist in two regimes. The first regime is where the vapor pressure of the etch by product gas (indium chloride) is below the etcher chamber pressure. In this first regime the activation energy was calculated to be 0.45 ± 0.05 eV. The second regime is where the vapor pressure of the etch by product gas (indium chloride) is above the etcher chamber pressure. In the second regime the activation energy was calculated to be 0.06 ± 0.03 eV. © *1998 American Vacuum Society.* [S0734-211X(98)09604-8]

I. INTRODUCTION

Via hole etching of indium phosphide using chlorine plasmas is shown in the literature to be sensitive to the wafer chuck temperature in the temperature range from about 100 to 170 °C.[1,2] The estimated activation energy based upon the reported wafer chuck temperature was calculated to be 0.19 and 0.22 eV. The literature suggests that this etch rate is strongly dependent upon the volatility of the indium chloride, which is a by-product of the etch process.[1–4] The higher the temperature of the wafer, the higher the volatility of the indium chloride, and the faster the etch rate. To calculate the true activation energy of the etch process it is necessary to know the wafer temperature during the etch. To the best of our knowledge, the actual temperature of the wafer has yet to be reported in the literature because it is difficult to measure during a plasma etch. This inability to measure the wafer temperature has led to claims of being able to etch indium phosphide with a high etch rate but with a very low wafer temperature.[5,6]

This article reports on the temperature sensitivity of indium phosphide etching of via holes using an Ar/Cl_2 plasma. The wafer temperature was determined by measuring the heating of the wafer from the plasma and the cooling of the wafer from the helium backside cooling. This wafer temperature was substituted for the wafer chuck temperature to determine the etch rate as a function of wafer temperature. The etch rate as a function of wafer temperature was found to exist in two regimes. The first regime is where the indium chloride has a vapor pressure below the etch chamber pressure. The second regime is where the indium chloride has a vapor pressure above the etch chamber pressure. The activation energy of the two regimes is also reported.

a)Electronic mail: edwin.sabin@trw.com

II. EXPERIMENTAL SETUP

Indium phosphide wafer chips (0.027 in. thick) were coated with 19 μm of AZ4620 photoresist. The wafer chips were all the same size at approximately 1.9 cm^2. The wafer chips were attached to a 3.25 in. diam sapphire wafer (0.030 in. thick) using KRYTOX® vacuum grease. The thickness of the grease was measured at approximately 0.002 in. A Plasma Therm SLR 770 etch system was used to etch all the wafers chips. The electron cyclotron resonance etch parameters used in this study were: 1 sccm Ar, 19 sccm Cl_2, 50–160 °C wafer chuck temperature, 5 mTorr etch chamber pressure, 5 min of etch time, 1000 W of microwave power (μW), 425 W of rf power (rf) with a corresponding dc bias of $-210+/-10$ V, 180 A for the upper magnet, and 40 A for the lower magnet. After the etch, the wafer chips were cross sectioned and a scanning electron microscope was used to measured the etch depth.

III. RESULTS AND DISCUSSION

A. Etch rate as a function of wafer chuck temperature

Figure 1 shows an Arrhenius plot of the etch rate as a function of inverse wafer chuck temperature. At a wafer chuck temperature above about 90 °C the etch rate appears to be insensitive to temperature. Below about 90 °C the etch rate appears to be a strong function of temperature. If, as reported in the literature, the etch rate is a strong function of the volatility of indium chloride, then it is reasonable to assume that at a wafer chuck temperature lower than 90 °C the vapor pressure of the indium chloride is lower than the etch chamber pressure. Figure 2 shows the vapor pressure of the indium chloride as a function of temperature, calculated from vapor pressure data.[7] Figure 2 shows that below 175 °C the vapor pressure of the indium chloride is below the etch chamber pressure. Figures 1 and 2 indicate that when the wafer chuck temperature is about 90 °C, then the wafer chip temperature is about 175 °C. This method of estimating the

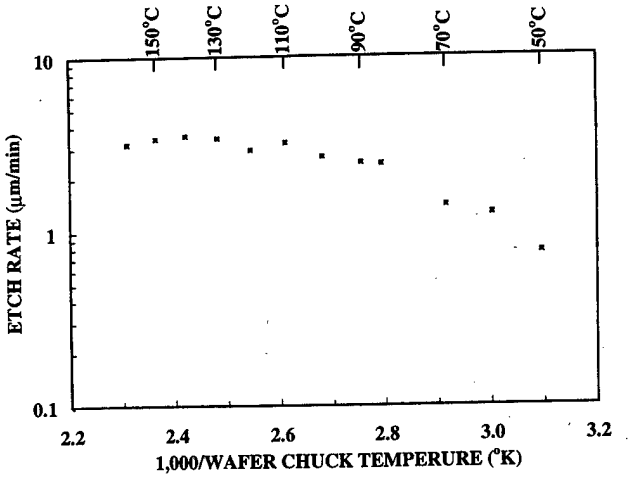


FIG. 1. Etch rate as a function of wafer chuck temperature.

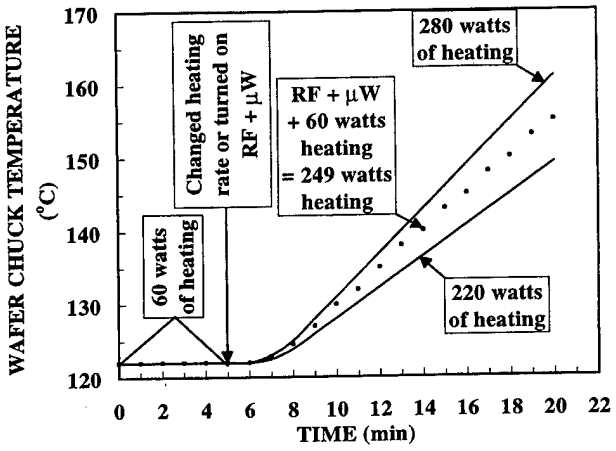


FIG. 3. Increase in wafer chuck temperature due to rf+μW plasma or increase in heater power.

wafer chip temperature is not accurate but does provide a rough estimate of the temperature. A better method of estimating the wafer chip temperature is presented below.

B. Estimating the wafer chip temperature

To determine the wafer chip temperature the following methodology was used. The heating of the wafer chip from the plasma was measured. Then, the cooling of the wafer chip from the helium backside was measured and the net heating/cooling was used to calculate the wafer chip temperature.

The heating of the wafer chip from the plasma was determined by measuring the heating of the wafer chuck from the microwave/rf plasma and comparing this to the heating of the wafer chuck with the heater alone. First, the system was allowed to equilibrate overnight at a heater power level of 60 W. A sapphire wafer was then transferred to the wafer chuck, the gasses turned on, the helium backside cooling turned on, and the microwave/rf (rf+μW) plasma was turned on. The temperature was monitored as the wafer chuck heated up from the plasma. After collecting the data, the sapphire wafer was removed and the system was again allowed to equilibrate at a heater power level of 60 W. A sapphire wafer was again transferred to the wafer chuck and the gasses and helium turned on, but instead of the microwave/rf being turned on, the heater power level was increased to 280 W. Again, the temperature was monitored as the wafer chuck was heated up from the increased heater power. The sapphire wafer was removed and a third time the system was allowed to equilibrate at a heater power level of 60 W. Then, the system was heated up at 220 W. The data from these three heat tests are shown in Fig. 3. The heating of the wafer chuck from the plasma is approximately half way between the heating from 280 W and the heating from 220 W. The heating from the rf+μW plasma plus 60 W of heater power was interpolated to be 249 W. This yields the heating from the rf+μW plasma at 189 W.

To determine the cooling from the helium backside cooling requires knowing the heat transfer coefficient of the helium. The methodology used to determine this is that a wafer on the wafer chuck acts as a shield to the loss of energy from the wafer chuck. If the heater power is held constant and a wafer is placed on the wafer chuck, the temperature of the wafer chuck will increase because the heat loss is reduced. The amount of this temperature increase is dependent upon the radiation heat transfer coefficient. With helium backside

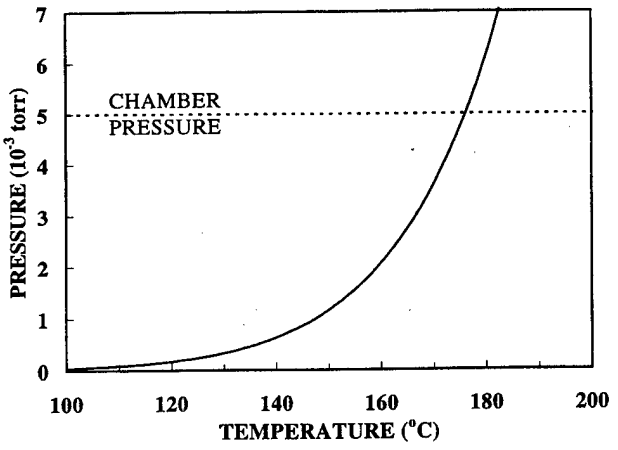


FIG. 2. Vapor pressure of indium chloride as a function of temperature.

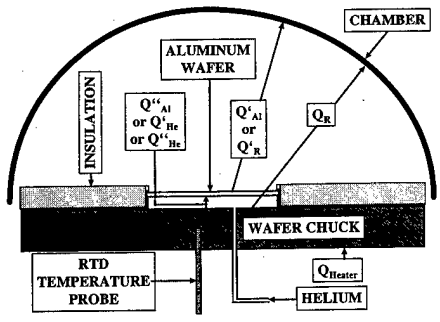


FIG. 4. Schematic drawing of etch system showing energy loss to chamber.

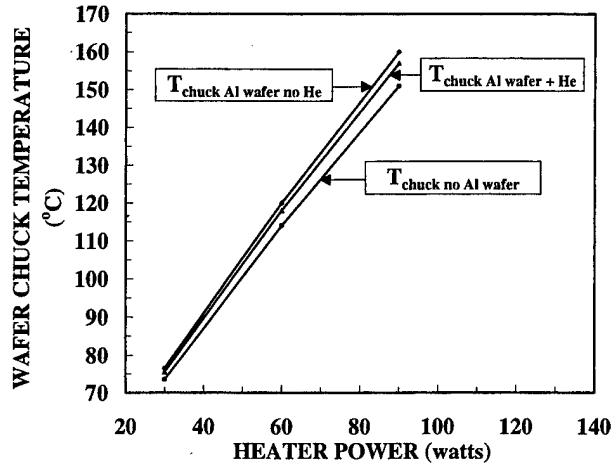


FIG. 5. Equilibrium wafer chuck temperature as a function of heater power.

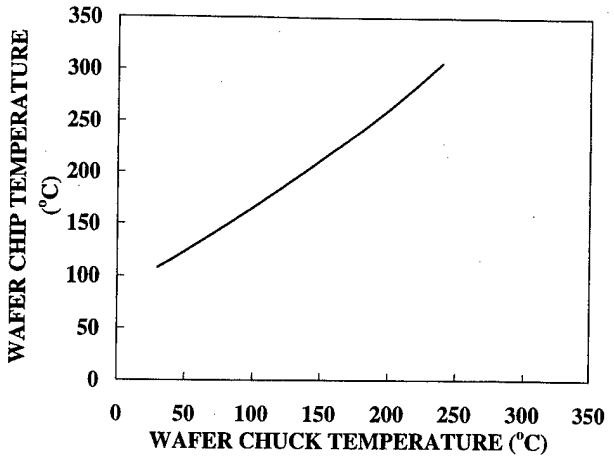


FIG. 7. Wafer chip temperature as a function of wafer chuck temperature during via hole etching.

cooling the amount of shielding is reduced as the helium increases the energy loss. Therefore, the temperature increase with the helium turned on is not as much as with the helium turned off. The amount of this increase is dependent upon the radiation heat transfer coefficient and the helium heat transfer coefficient. Figure 4 shows a schematic drawing of the etch chamber. An aluminum wafer was used for these tests to eliminate the need to determine the emissivity of the wafer chuck and wafer since the wafer chuck was made from aluminum also. First, the temperature of the wafer chuck ($T_{\text{chuck no Al wafer}}$) was determined for several different heater power levels. Then, an aluminum wafer was placed on the wafer chuck and the temperature of the wafer chuck ($T_{\text{chuck Al wafer no He}}$) was determined for the same heater power levels. Then, the helium backside cooling was turned on and the temperature of the wafer chuck ($T_{\text{chuck Al wafer+He}}$) was again determined for the same heater power levels. The data from these tests are shown in Fig. 5. The temperature increase at a constant heater power level was used to calculate the reduction in the energy loss (see the Appendix). This reduction in energy loss was then used to calculate the radiation heat transfer coefficient (see the Appendix). Then, the heat transfer coefficient for helium was calculated (see the Appendix). The helium heat transfer coefficient as a function of temperature is shown in Fig. 6.

The thermal conductivity as a function of temperature for indium phosphide, grease, and sapphire was extracted from the literature.[8–10] The thickness of the indium phosphide, grease, and sapphire wafer was measured with a caliper. The temperature at equilibrium for the wafer chip was then calculated using a plasma heating of 189 W and using the radiation and helium heat transfer coefficients determine above. Figure 7 shows the calculated wafer chip temperature as a function of wafer chuck temperature.

C. Etch rate as a function of wafer chip temperature

The wafer chip temperature from Fig. 7 was substituted for the wafer chuck temperature in Fig. 1 and these data were used to create Fig. 8. Figure 8 shows that the etching process can be separated into two different temperature regimes. For temperatures below about 165 °C the etch rate has an activa-

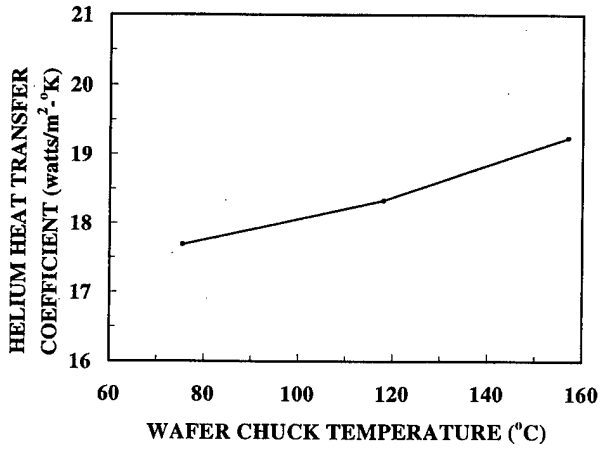


FIG. 6. Helium backside cooling heat transfer coefficient as a function of temperature.

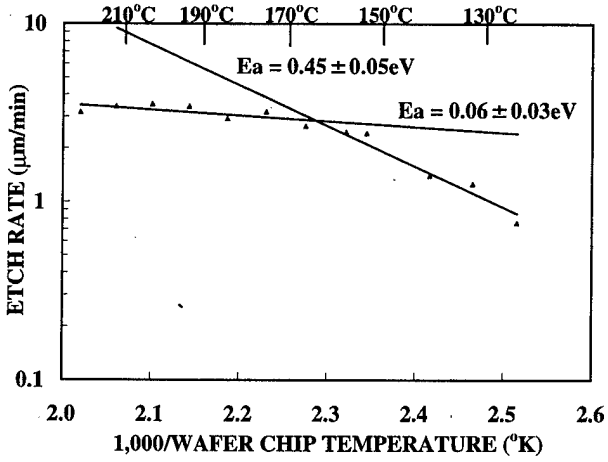


FIG. 8. Arrhenius plot of etch rate as a function of inverse wafer chip temperature.

tion energy of 0.45 ± 0.05 eV. For temperatures above about 165 °C the etch rate has an activation energy of 0.06 ± 0.03 eV. A temperature of 165 °C is just below the temperature where the vapor pressure of the indium chloride is the same as the etch chamber pressure (see Fig. 2). These data indicate that the etching of indium phosphide is dependent upon the vapor pressure of the indium chloride when the wafer chip temperature produces an indium chloride vapor pressure below the etch chamber pressure. For wafer chip temperatures where the vapor pressure of the indium chloride is above the etch chamber pressure the etch rate is very insensitive to the temperature.

IV. CONCLUSIONS

The wafer temperature in an Ar/Cl_2 plasma is much hotter than the wafer chuck temperature. Depending upon the amount of heating the plasma provides to the wafer and the helium backside cooling, the wafer temperature could be as much as 75 °C hotter than the wafer chuck temperature.

The etching of indium phosphide can be separated into two temperature regimes. For wafer temperatures where the vapor pressure of the indium chloride is less then the etch chamber pressure, the etch rate has an activation energy of 0.45 ± 0.05 eV. For wafer temperatures where the vapor pressure of the indium chloride is above the etch chamber pressure, the etch rate has an activation energy of 0.06 ± 0.03 eV. The indication is that the etch rate is dependent upon the volatility of the indium chloride where the vapor pressure is less then the etch chamber pressure.

ACKNOWLEDGMENT

The author wishes to thank his colleagues at TRW for reviewing this article and providing helpful improvements to the writing.

APPENDIX

1. Calculation of the reduction in energy lost by radiation from the wafer chuck

Performing a regression analysis on the data from Fig. 5 produced an equation for the temperature of the wafer chuck as a function of heater power level (Q_{heater}). For the situation with an aluminum wafer on the chuck but *without* helium backside cooling, the equation is

$$T_{\text{chuck Al wafer no He}} = 35\ ^\circ\text{C} + 1.39\ ^\circ\text{C/W} * Q_{\text{Heater}}. \quad \text{(A1)}$$

For the situation with an aluminum wafer on the chuck and *with* helium backside cooling, the equation is

$$T_{\text{chuck Al wafer+He}} = 35\ ^\circ\text{C} + 1.36\ ^\circ\text{C/W} * Q_{\text{heater}}. \quad \text{(A2)}$$

The slope from Eq. (A1) is used to determine the reduction in the energy (Q_{Al}) lost when an aluminum wafer is on the wafer chuck *without* helium backside cooling. The difference in the temperature between the wafer chuck with no aluminum wafer and with an aluminum wafer is dependent upon how much energy reduction occurs. The equation for this reduction at a constant heater power level is

$$Q_{\text{Al}} = (T_{\text{chuck Al wafer no He}} - T_{\text{chuck no Al wafer}})/(1.39\ ^\circ\text{C/W}). \quad \text{(A3)}$$

The slope from Eq. (A2) is used to determine the reduction in the energy (Q_{He}) lost when an aluminum wafer is on the wafer chuck *with* helium backside cooling. The difference in the temperature between the wafer chuck with no aluminum wafer and with an aluminum wafer having helium backside cooling is dependent upon how much energy reduction occurs. The equation for this reduction at a constant heater power level is

$$Q_{\text{He}} = (T_{\text{chuck Al wafer+He}} - T_{\text{chuck no Al wafer}})/(1.36\ ^\circ\text{C/W}). \quad \text{(A4)}$$

2. Calculation of the radiation heat transfer coefficient (s), refer to Fig. 4

When an aluminum wafer is on the wafer chuck the energy that is transferred to the chamber (Q'_{Al}) can be expressed using a simplified equation for blackbody radiation.[11] The radiation heat transfer coefficient (s) has lumped into it the emissivity, the view factor, the Stefan–Boltzman constant, and the area of radiation. The equation is

$$Q'_{\text{Al}} = [(T_{\text{Al wafer}})^4 - (T_{\text{chamber}})^4] * s. \quad \text{(A5)}$$

The equation for the energy that is transferred to the aluminum wafer from the wafer chuck is

$$Q''_{\text{Al}} = [(T_{\text{chuck}})^4 - (T_{\text{Al wafer}})^4] * s. \quad \text{(A6)}$$

Although not all of the energy that is transferred to the aluminum wafer is also transferred to the chamber, for this analysis it is assumed that $Q''_{\text{Al}} = Q'_{\text{Al}}$. Therefore, these two equations are equal. By using an aluminum wafer the emissivity value and view factor are the same for the wafer chuck and the wafer, so the radiation heat transfer coefficient (s) is the same for both equations. Using Eqs. (A5) and (A6), the temperature of the aluminum wafer can be determined:

$$(T_{\text{Al wafer}})^4 = [(T_{\text{chuck}})^4 + (T_{\text{chamber}})^4]/2. \quad \text{(A7)}$$

The total energy transferred to the chamber from the wafer chuck *when there is no aluminum wafer* will be defined as Q_R. To find the ratio between Q_R and Q'_{Al}, the following equation was used:

$$Q_R/Q'_{\text{Al}} = [(T_{\text{chuck}})^4 - (T_{\text{chamber}})^4] * s / [(T_{\text{Al wafer}})^4 - (T_{\text{chamber}})^4] * s. \quad \text{(A8)}$$

The radiation heat transfer coefficient (s) is the same for numerator and denominator since the emissivity and the view factor of the aluminum and the wafer chuck are the same. Using the equation for the wafer temperature Eq. (A7), the ratio was determined to be

$$Q_R/Q'_{\text{Al}} = 2. \quad \text{(A9)}$$

Equation (A9) shows that when an aluminum wafer is placed on the wafer chuck the energy transferred to the chamber is reduced to half of what it was previously. This reduction in the energy loss was previously determined from Eq. (A3) and is Q_{Al}. Therefore,

$$Q'_{Al}=Q_{Al}$$

and

$$Q_R=2*Q_{Al}. \quad \text{(A10)}$$

Knowing the energy transferred to the chamber by radiation from the aluminum wafer, and also knowing the wafer temperature, it is now possible to calculate the radiation heat transfer coefficient (s) using Eqs. (A3), (A5), (A7), and (A10):

$$s=2*Q_{Al}/[(T_{chuck})^4-(T_{chamber})^4]. \quad \text{(A11)}$$

3. Calculation of helium heat transfer coefficient (h), refer to Fig. 4

When a wafer is on the wafer chuck with helium backside cooling, the energy transferred to the wafer due to the helium can be expressed using a simplified heat transfer equation.[12] The heat transfer coefficient (h) has lumped into it the mass flow of the helium and the area of heating. The equation of the energy transferred to the aluminum wafer by the helium is

$$Q'_{He}=h*(T_{chuck}-T_{Al\ wafer}). \quad \text{(A12)}$$

The equation of the energy transferred to the aluminum wafer by radiation during helium backside cooling is

$$Q''_{He}=s*[(T_{chuck})^4-(T_{Al\ wafer})^4]. \quad \text{(A13)}$$

The equation of the energy transferred to the chamber by radiation from the aluminum wafer is

$$Q'_R=s*[(T_{Al\ wafer})^4-(T_{chamber})^4]. \quad \text{(A14)}$$

The energy transferred to the aluminum wafer by radiation and helium backside cooling equals the energy transferred to the chamber by the aluminum wafer:

$$Q'_R=Q'_{He}+Q''_{He}. \quad \text{(A15)}$$

The total heat transferred to the chamber with no aluminum wafer is Q_R, which was previously calculated by Eq. (A10). This quantity is reduced when an aluminum wafer is on the wafer chuck and helium backside cooling is turned on. The amount of this reduction was determined previously in Eq. (A4) and is Q_{He}. Therefore,

$$Q'_R=Q_R-Q_{He}. \quad \text{(A16)}$$

Calculating Q'_R from Eq. (A16), and knowing the radiation heat transfer coefficient (s) calculated from Eq. (A11), it is now possible to calculate the aluminum wafer temperature. By rearranging Eq. (A14) and substituting Eq. (A16) for Q'_R, the following equation for the aluminum wafer temperature is derived:

$$(T_{Al\ wafer})^4=(Q_R-Q_{He})/s+(T_{chamber})^4. \quad \text{(A17)}$$

Now knowing the aluminum wafer temperature, it is possible to calculate the energy transferred to the wafer by radiation using Eq. (A13). Finally, it is possible to calculate the helium heat transfer coefficient (h) using Eqs. (A12), (A13), (A15), and (A16).

$$h=(Q_R-Q_{He}-Q''_{He})/(T_{chuck}-T_{Al\ wafer}). \quad \text{(A18)}$$

[1]R. Shul, A. Howard, C. Vartuli, P. Barnes, and W. Seng, J. Vac. Sci. Technol. A **14**, 1102 (1996).
[2]R. Khare, J. Brown, M. Hu, D. Pierson, M. Melendes, and C. Constantine, J. Vac. Sci. Technol. B **12**, 2947 (1994).
[3]F. Ying, W. H. Juan, and S. W. Pang, J. Vac. Sci. Technol. B **15**, 665 (1997).
[4]C. Youtsey and I. Adesida, J. Vac. Sci. Technol. B **13**, 2360 (1995).
[5]J. Lee, J. Hong, and S. Pearton, Appl. Phys. Lett. **68**, 847 (1996).
[6]J. W. Lee, J. Hong, C. R. Abernathy, E. S. Lambers, S. J. Pearton, W. S. Hobson, and F. Ren, J. Vac. Sci. Technol. B **14**, 2567 (1996).
[7]C. J. Smithells and E. A. Brandes, *Metals Reference Book*, 5th ed. (Butterworths, London, 1976), p. 235, Table 14.
[8]I. Kudman and E. F. Steigmeier, Phys. Rev. **133**, A1665 (1964).
[9]Private communication with Du Pont.
[10]R. C. Weast, *Handbook of Chemistry and Physics*, 53rd ed. (The Chemical Rubber Company, Cleveland, OH, 1972), p. E5.
[11]W. L. McCabe and J. C. Smith, *Unit Operations of Chemical Engineering*, 3rd ed. (McGraw-Hill, New York, 1976), p. 293, Eqs. (11)–(9).
[12]W. L. McCabe and J. C. Smith, *Unit Operations of Chemical Engineering*, 3rd ed. (McGraw-Hill, New York, 1976), p. 383, Eqs. (14)–(25).

Abrupt reduction in poly-Si etch rate in HBr/O_2 plasma

Shinji Kuroda
Radioisotope Research Center, Osaka University, 2-4 Yamadaoka, Suita, Osaka 565, Japan

Hiroaki Iwakuro[a)]
Shindengen Electric Manufacturing Co., Ltd., 10-13 Minami, Hanno, Saitama 357, Japan

(Received 23 October 1997; accepted 29 May 1998)

The effect of oxygen on polycrystalline-Si (poly-Si) and SiO_2 etching in hydrogen bromide (HBr) reactive ion etching plasmas has been studied by measuring etch rates and using x-ray photoelectron spectroscopy (XPS) to study compositional changes in the surface layer. The etch rate of the poly-Si increases dramatically from 30 to 235 nm/min as the O_2 concentration increases from 0% to 25%, whereas the SiO_2 etch rate gradually decreases from 3 to 1 nm/min. Above 30% O_2 in HBr, the poly-Si etch rate abruptly decreases by a factor of 16 compared with that at 25%. From XPS analysis, it is found that the abrupt decrease of the poly-Si etch rate at O_2 concentrations of more than 30% is closely related with the composition and thickness of an $SiBr_xO_y$ layer formed during the HBr/O_2 plasma exposure. The $SiBr_xO_y$ layer has a composition of nearly SiO_2. Br ions cannot permeate the $SiBr_xO_y$ layer formed in plasmas containing 30% O_2 in HBr (or greater), and therefore, the poly-Si etch terminates. © *1998 American Vacuum Society.* [S0734-211X(98)13904-5]

I. INTRODUCTION

Reactive ion etching (RIE), which combines physical sputtering and chemical reaction to achieve anisotropic etching, has become a key dry etching process in recent microfabrication technology. Recently, RIE with hydrogen bromide (HBr) gas has been found to offer high potential in meeting some of the challenges of submicron fabrication requirements. The etching characteristics such as etch selectivity, anisotropy, etch rate, and damage have been reported both for polycrystalline silicon (poly-Si) and crystalline silicon.[1–3]

We have etched a poly-Si film as a mask of an SiO_2 film in pure HBr plasma. The resulting etch selectivity (poly-Si:SiO_2) had a value of 10, which is insufficient for device fabrication. So, we examined the effect of the addition of O_2 to an HBr plasma on the etch rates of poly-Si and SiO_2 films, and we found that the etch rate of the poly-Si increased dramatically up to an O_2 concentration of 25%, and that the poly-Si etch rate abruptly decreased by a factor of about 16 with more O_2 addition. This abrupt reduction in the poly-Si etch rate did not appear in CF_4/O_2 plasma,[4,5] in which a gradual decrease in the etch rate is observed following the addition of O_2 past the maximum etch rate. We considered that this abrupt reduction in the poly-Si etch rate was closely related to the resulting reaction layer on the surface, and we have analyzed the reaction layer with x-ray photoelectron spectroscopy (XPS).

II. EXPERIMENT

The Si substrates used in the present study were 4 in. in diameter, (001)-oriented *n*-type silicon wafers with a resistivity of 0.5 Ω cm. Silicon dioxide films 100 nm thick were prepared by thermal oxidation of the Si substrates at 1373 K in dry oxygen. Poly-Si films 500 nm thick were deposited on the SiO_2 films by a chemical vapor deposition system.

The poly-Si and SiO_2 were dry etched in a magnetron ion etching system equipped with a 3 kW rf generator (13.56 MHz).[6] The reaction chamber was evacuated by a turbomolecular pump to 2.7×10^{-4} Pa. Etching conditions included an rf power of 2.0 kW (1.6 W/cm^2 power density) and gas pressure of 0.67 Pa, which is a typical plasma condition for the etching of poly-Si. Etching time was 1 min. The substrate was held controllably at room temperature. Hydrogen bromide gas, available from Asahi Denka Kogyo, was four-nine purity. In this plasma, the ion energy is not monoenergetic, as the plasma has an energy distribution that correlates with its peak ion energy. The average ion energy is exactly represented by the sum of the self-bias voltage (about 300 V) plus the plasma potential (5–20 V), but is taken to be roughly equal to the self-bias voltage.[6]

In the magnetron ion etching system, poly-Si films were usually etched using the gases SF_6 and CHF_3, so that the reactor was contaminated by the etching gases and the etching species. Accordingly, the reactor was cleaned before the experiments in this study were carried out, but the contamination could not be completely removed from the reactor.

Etch rates of the poly-Si and SiO_2 were determined by measuring the thickness before and after dry etching using interferometry.

Etched surfaces of the poly-Si were characterized by XPS. The samples were always exposed to the atmosphere after dry etching, and were analyzed by XPS within 2 h. X-ray photoelectron spectra were obtained with a Dupont ESCA 760B apparatus using an Mg $K\alpha$ (1253.6 eV) x-ray source. The acceleration voltage and emission current of the x-ray source were 8 kV and 30 mA, respectively. The base pressure of the analysis chamber was 2.7×10^{-6} Pa during accumulation of the spectra.

a)Electronic mail: hiwakuro@post.click.or.jp

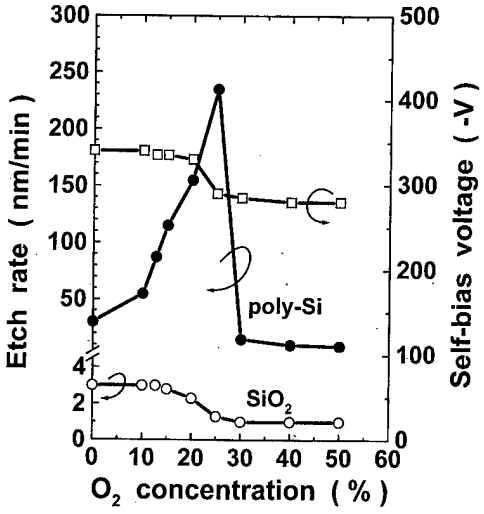


FIG. 1. Etch rates of the poly-Si and SiO_2 and the dc self-bias voltage as a function of O_2 concentration in HBr/O_2 plasma.

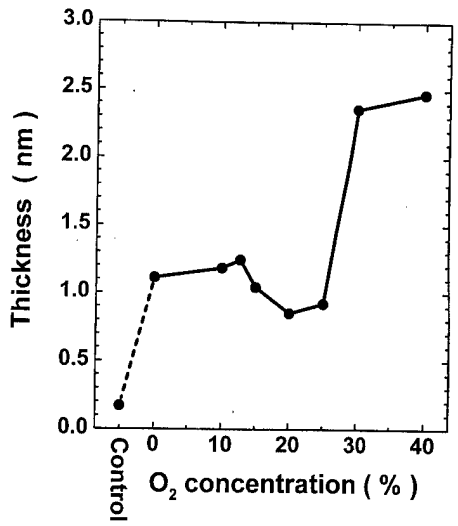


FIG. 2. Etch selectivity of the poly-Si with respect to SiO_2 as a function of O_2 concentration in HBr/O_2 plasma.

III. EXPERIMENTAL RESULTS

Figure 1 shows etch rates of poly-Si and SiO_2 and the dc self-bias voltage as a function of O_2 concentration in HBr/O_2 plasma. The etch rate of the poly-Si has a value of 30 nm/min in pure HBr plasma. The poly-Si etch rate increases dramatically when O_2 is added to HBr. A maximum etch rate of 235 nm/min is achieved with 25% of added O_2. A further increase in O_2 concentration leads to an abrupt decrease by a factor of about 16 in the poly-Si etch rate. On the other hand, the dependence of the SiO_2 etch rate on O_2 addition differs from that of the poly-Si etch rate. The SiO_2 etch rate gradually decreases from 3 to 1 nm/min with increasing O_2 concentration from 0% to 25%, and keeps constant at the value of 1 nm/min at O_2 concentrations above 25%. The decreases in the SiO_2 etch rate and the self-bias voltage show a similar tendency as a function of O_2 concentration.

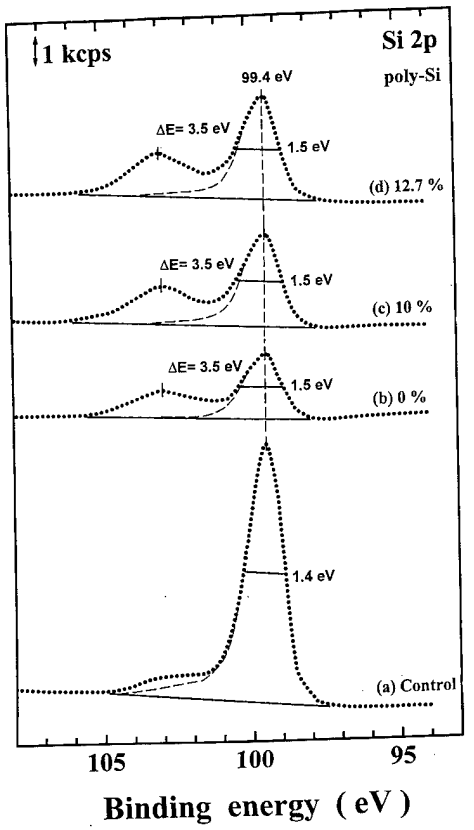
Figure 2 shows the etch selectivity of the poly-Si with respect to the SiO_2 as a function of O_2 concentration. The etch selectivity (poly-Si:SiO_2) rises rapidly with an increase in O_2 concentration, and then reaches the highest value of about 235 at 25% O_2. Additional O_2 leads to a large reduction in the etch selectivity because of the abrupt decrease in the poly-Si etch rate.

Figure 3 shows Si $2p$ spectra obtained from poly-Si samples that have been exposed to HBr/O_2 plasma. The Si $2p$ spectrum obtained from a nonplasma-exposed poly-Si sample is also shown in Fig. 3(a) as a control. The Si $2p$ spectra were separated into two components: an elemental Si $2p$ spectrum with a binding energy of 99.4 eV and a chemically shifted Si $2p$ peak at a binding energy of 102.9 eV. These components correspond to silicon in poly-Si and a reaction layer, respectively, which is formed on the poly-Si surface by plasma exposure. The symbol ΔE refers to the difference between the binding energies of the two peaks.

In Fig. 3(a), the small signal around 103 eV indicates a slight oxidation on the poly-Si surface due to atmospheric exposure. In the pure HBr plasma-exposed sample [Fig. 3(b)], the two peaks at 99.4 eV (from silicon in poly-Si) and 102.9 eV are observed. The peak at 102.9 eV (from the reaction layer) is attributed to an Si–Br–O oxide ($SiBr_xO_y$), not to $SiBr_x$. This is because binding energies of $SiBr_x(x=1,2,3,4)$ have values of 100.05, 100.7, 101.35, and 102.0 eV, respectively,[7] and because signals of Br and O are detected from the sample in the XPS data (Br $3d$ and O $1s$ spectra are not shown). Although no O_2 was added to the HBr, a Si oxide is present on the poly-Si surface. Atmospheric exposure of the HBr plasma-etched poly-Si surface could explain the formation of such a layer. The peak intensity at 99.4 eV in Fig. 3(b) is much lower compared with that of the control [Fig. 3(a)]. The peak intensity at 99.4 eV increases with increasing O_2 concentration from 0% to 25%, and then decreases with more O_2 addition. On the other hand, the intensity of the 102.9 eV peak increases consistently with increasing O_2 concentration, and becomes larger than the 99.4 eV peak at an O_2 concentration of 30%. In addition, the energy difference ΔE increases from 3.6 to 4.1 eV with increasing O_2 concentration, as shown in Figs. 3(e)–3(i) for 15%–40% O_2 in HBr plasma.

Since the Si of SiO_2 is in its +4 oxidation state and showed a chemical shift of about 4.5 eV to the Si $2p$ binding energy of bulk Si, every unit increment of oxidation number (each Si–O bond) roughly contributes a chemical shift of about 1.1 eV.[8] On the other hand, each Si–Br bond contributes a chemical shift of 0.65 eV.[7] This means that in the Si of $SiBr_xO_y$ the chemical shift to the Si $2p$ binding energy of bulk Si increases as the degree of oxidation increases. Accordingly, the increase in the chemical shifts of the ΔE 3.6–4.1 eV is attributed to an increase in the degree of oxidation of the $SiBr_xO_y$.

Figure 4 shows normalized peak intensities of Si $2p$ (the poly-Si and the Si oxide) and Br $3d$ spectra obtained from the poly-Si samples as a function of O_2 concentration in the



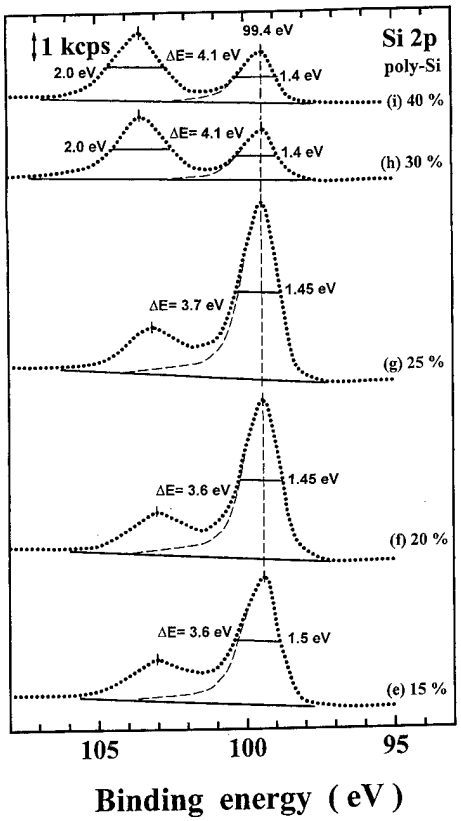


FIG. 3. Si 2*p* spectra obtained from the poly-Si samples that have been exposed to HBr/O_2 plasma.

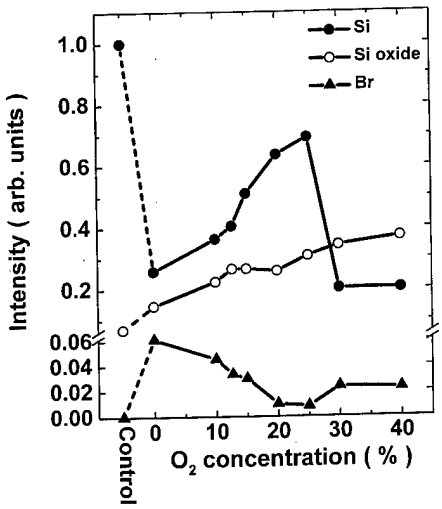


FIG. 4. Normalized peak intensities of the poly-Si samples as a function of O_2 concentration in HBr/O_2 plasma.

HBr/O_2 plasma. The intensity values on the vertical axis were determined as follows: each peak intensity was divided by the cross section for photoelectron production which was obtained from Scofield.[9] The corrected peak intensities were normalized to that of the Si 2*p* spectrum for the control sample.

The intensity of the Si peak from the poly-Si is smallest at 0%, increases with an increase in O_2 concentration, and achieves a maximum value at O_2 concentration of 25%. A further increase in O_2 concentration leads to an abrupt reduction by more than two times in the Si peak intensity from the poly-Si. The behavior of the Si peak intensity from the poly-Si is similar to that of the poly-Si etch rate (Fig. 1). This suggests that the thickness of the $SiBr_xO_y$ formed on the poly-Si is related to the poly-Si etching, since the Si 2*p* photoelectrons of the poly-Si bulk have escaped from the surface through the $SiBr_xO_y$.

In O_2 concentration of 0%–25% the Br peak intensity is largest at 0%, and decreases gradually with increasing O_2 concentration. The Br peak intensity slightly increases at O_2 concentrations greater that 30%. On the other hand, the intensity of the Si oxide increases consistently with increasing O_2 concentration. On comparison of the Si oxide and Br, the the Si oxide turns out to have an intensity much larger than the Br. In particular, in O_2 concentration of 20%–40%, the intensity of the Si oxide is larger by more than 20 than the Br intensity. This indicates that the composition of the $SiBr_xO_y$ is nearly SiO_2.

Figure 5 shows thicknesses of the Si oxides formed on the poly-Si as a function of O_2 concentration in the HBr/O_2 plasma. The thickness was computed from the intensity ratio of the shifted Si 2*p* peak corresponding to Si oxide, to the elemental Si 2*p* peak corresponding to the poly-Si,[10] assuming that the composition of the $SiBr_xO_y$ was SiO_2. Cross sections for photoelectron production were taken from Scofield,[9] and electron mean free paths in Si and SiO_2 (λ_{Si}

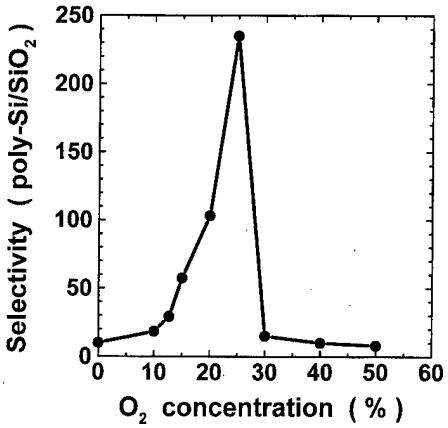


FIG. 5. Thickness of silicon oxide formed on the poly-Si surface as a function of O_2 concentration in HBr/O_2 plasma.

$=1.3$ nm and $\lambda_{Ox}=2.1$ nm) were taken from Hochella and Carim.[10]

The oxide thickness increases slightly with increasing O_2 concentration from 0% to 12.7%, and then decreases to 0.9 nm with an increase in O_2 concentration to 25%. However, a further O_2 addition results in an abrupt increase by a factor of 2.6 in the thickness (2.37 nm at 30%). The O_2 concentration at which the abrupt increase in the thickness occurs corresponds to that of the abrupt decrease in the poly-Si etch rate.

IV. DISCUSSION

When Br ions in pure HBr plasma bombard the poly-Si surface, a reaction layer of $SiBr_x$ is formed.[2,3] The $SiBr_x$ is likely to remain on the poly-Si surface, since the volatility of $SiBr_4$ is very low (the $SiBr_4$ boiling point is 427 K).[11] Nevertheless, the $SiBr_x$ can be desorbed from the poly-Si surface, resulting in etching of the poly-Si, because the surface is subjected to bombardment by ions, electrons, and photons from the plasma. Poly-Si etching in HBr plasma only occurs in the presence of ion bombardment, which indicates that the Si etching is dominated by physical sputtering.[1]

In pure HBr plasma, the etch rate of SiO_2 is very low compared with that of the poly-Si (Fig. 1). The lower SiO_2 etch rate is attributable to the larger binding energy (191.1 kcal/mol) of the Si–O bond relative to that (78.1 kcal/mol) of the Si–Si bond,[12] and because there is no effective chemical mechanism for the removal of oxygen from the SiO_2 surface.

Previous studies of HBr reactive ion etching have suggested that the presence of carbon in the reactor has a strong influence on the SiO_2 etch rate; the SiO_2 etching was enhanced by carbon contamination.[2,3] In the present study, we used a reactor which was contaminated by carbon. Accordingly, the decrease in the SiO_2 etch rate with increasing O_2 concentration is due to the decrease in the amount of carbon in the reactor due to CO or CO_2 formation and to the observed slight reduction in the self-bias voltage.

In HBr/O_2 plasma, the poly-Si etch rate is determined by the concentrations of both Br and O ions in the plasma because they compete for active poly-Si surface sites. For low O_2 concentrations, the poly-Si etch rate is controlled by the arrival rate of Br ions to the surface and responds to an increase in the Br ion concentration in the plasma. The increase in the concentration of Br ions is similar to the situation for F ions in CF_4/O_2 plasma, in which the F concentration increases dramatically at first (up to about 20% O_2 addition).[13]

For high oxygen concentrations, the poly-Si etch rate is limited by oxidation of the poly-Si surface and correlates well with the increase in the thickness of the $SiBr_xO_y$ layer. Figures 3 and 4 show that an increase in the O_2 concentration leads to a change in the composition from primarily Si–Br bonding to Si–O bonding (an increase in the degree of oxidation), and that the composition of the $SiBr_xO_y$ is nearly SiO_2 above a 30% O_2 addition to the HBr/O_2 plasma. The $SiBr_xO_y$ layer with a composition of nearly SiO_2 prevents the poly-Si from being etched because there is no effective chemical mechanism for the removal of oxygen from the SiO_2 surfaces.

If the $SiBr_xO_y$ layer on the poly-Si surface is thin (up to about 1.2 nm, as shown in Fig. 5), and the Br ions from the plasma can reach through the $SiBr_xO_y$ layer to the poly-Si layer, then poly-Si etching can proceed. When the $SiBr_xO_y$ layer is thick enough that the Br ions cannot diffuse to the poly-Si layer, the poly-Si etch rate decreases abruptly. The $SiBr_xO_y$ layer grows to a thickness of 2.37 nm during plasma exposure of 30% O_2 in HBr, as shown in Fig. 5. Accordingly, the abrupt decrease in the poly-Si etch rate at the 30% O_2/HBr plasma concentration is related to the composition and thickness of the $SiBr_xO_y$ formed on the surface during the plasma exposure.

V. CONCLUSION

The effect of oxygen on the poly-Si and SiO_2 etching in HBr reactive ion etching plasmas has been studied. The poly-Si etch rate increases dramatically from 30 to 235 nm/min as the O_2 concentration in HBr increases from 0% to 25%, whereas the SiO_2 etch rate gradually decreases from 3 to 1 nm/min. The poly-Si etch rate abruptly decreases by a factor of 16 at O_2 concentrations greater than 30%.

The abrupt decrease of the poly-Si etch rate at O_2 concentrations greater than 30% is closely connected with the composition and thickness of the $SiBr_xO_y$ layer formed during the HBr/O_2 plasma exposure. The $SiBr_xO_y$ layer, with a composition of nearly SiO_2, and a thickness large enough to prevent Br ion diffusion through it to the underlying poly-Si layer, prevents the continued etching of the poly-Si.

ACKNOWLEDGMENT

The authors would like to thank Professor T. Yamamoto of Osaka University for valuable discussions.

[1]L. Y. Tsou, J. Electrochem. Soc. **136**, 3003 (1989).
[2]M. Nakamura, K. Iizuka, and H. Yano, Jpn. J. Appl. Phys., Part 1 **28**, 2142 (1989).
[3]O. S. Nakagawa, S. Ashok, and J. K. Kruger, J. Appl. Phys. **69**, 2057 (1991).

[4]C. J. Mogab, A. C. Adams, and D. L. Flamm, J. Appl. Phys. **49**, 3796 (1978).
[5]R. d'Agostino, F. Cramarossa, S. D. Benedictis, and G. Ferraro, J. Appl. Phys. **52**, 1259 (1981).
[6]A. A. Bright and S. Kaushik, J. Vac. Sci. Technol. B **7**, 542 (1989).
[7]T. M. Bestwick and G. S. Oehrlein, J. Vac. Sci. Technol. A **8**, 1696 (1990).
[8]S. Iwata and A. Ishizaka, J. Appl. Phys. **79**, 6653 (1996).
[9]L. J. Scofield, J. Electron. **8**, 129 (1976).
[10]M. F. Hochella and A. H. Carim, Surf. Sci. Lett. **197**, L260 (1988).
[11]R. C. Weast, *CRC Handbook of Chemistry and Physics* (Chemical Rubber, West Palm Beach, FL, 1989), p. B-126.
[12]R. C. Weast, *CRC Handbook of Chemistry and Physics* (Chemical Rubber, West Palm Beach, FL, 1989), p. F-175.
[13]C. J. Mogab, A. C. Adams, and D. L. Flamm, J. Appl. Phys. **49**, 3796 (1978).

Structural, optical, and electrical properties of nanocrystalline silicon films deposited by hydrogen plasma sputtering

B. Garrido, A. Pérez-Rodríguez,[a] and J. R. Morante
Departament d'Electrónica, Universitat de Barcelona, Diagonal 645-647, 08028 Barcelona, Spain

A. Achiq, F. Gourbilleau, R. Madelon, and R. Rizk
LERMAT, UPRESA CNRS 6004, ISMRA, 6 Bd. du Maréchal Juin, F-14050 Caen Cedex, France

(Received 6 October 1997; accepted 17 April 1998)

Nanocrystalline silicon films were deposited by radio frequency sputtering in a pure H_2 plasma on glass and monocrystalline ⟨100⟩ silicon at various substrate temperatures, T_s. The detailed structural, optical, and electrical analysis of the films has been performed by transmission electron microscopy, Raman scattering, infrared spectroscopy, x-ray diffraction, optical absorption, photoluminescence and electrical measurements. The data obtained show that, to a significant extent, control of the structure and hence of the optical and electrical properties of the films can be achieved by changing T_s. Increasing T_s from 50 to 250 °C leads to an increase of the average grain size (from a few nm to a few tens of nm) and crystalline fraction (from 37% to 74%) and the optical band gap decreases from 2.40 to 1.95 eV. Hydrogen incorporation, together with T_s, are thought to be at the origin of the resulting microstructure and consequently determine the optical and transport properties. Moreover, hydrogen content was found to be associated with void formation which induces structure relaxation with very low residual stress. Finally, electrical conductivity in the layers increases by more than six orders of magnitude with T_s. The high dark conductivity measured from the sample deposited at the highest T_s ($>10^{-3}$ Ω^{-1} cm^{-1}) and its low activation energy (0.13 eV) are in agreement with the high crystalline fraction of this layer, where tunneling of carriers between the crystallites likely occurs. © *1998 American Vacuum Society.*
[S0734-211X(98)02704-8]

I. INTRODUCTION

Hydrogenated nanocrystalline silicon (*nc*-Si:H) is a structurally inhomogeneous material which consists of nanometer-size crystallites usually embedded in an amorphous tissue, together with a more or less important fraction of microvoids. Its fabrication requires relatively small modifications from the well known techniques for deposition of amorphous silicon (*a*-Si). The hydrogenated microcrystalline material shows much higher conductivities, mobilities, and doping efficiency, and a lower optical absorption than its amorphous counterpart, which gives interest to these films for applications such as thin film photovoltaic cells and transistors for flat panel displays.[1–4] Nanocrystalline silicon (*nc*-Si) films are also of interest for other optoelectronic applications as luminescent devices. When the crystal size decreases below 10 nm and becomes comparable with the diameter of the bulk exciton (4–5 nm), quantum size effects occur leading to a widening of the band gap and a concomitant increase of the probability for optical transitions.[5]

Nevertheless, it is difficult to fabricate nanostructured grains with reproducible sizes in the nanometer scale. The difficulty on achieving control of the average size and distribution of sizes of the crystallites and the crystalline fraction in the films causes severe problems for the reproducibility of the characteristics of most thin film devices.[6] In addition to these parameters, other structural issues are also of great importance in controlling those properties, as the nature of grain boundaries with the amorphous tissue, the fraction of voids, the hydrogen content and its bonding configuration. These features affect significantly the optical and electrical properties of the material.

In principle, two kinds of methods have been proposed for the fabrication of high performance *nc*-Si structures: (i) recrystallization of amorphous films and (ii) direction deposition. Device grade microcrystalline Si (μc-Si) and *nc*-Si films have been obtained by solid phase recrystallization of *a*-Si using methods such as laser, flash lamp and conventional furnace annealing.[7–9] The quality of the crystallized layers is found to depend strongly on the structural configurations, defects, and stresses at the level of grain boundaries, which give exponential band tails of electronic states near the band edges.[9] In general, these methods have difficulties in the accurate control of the size of nanocrystals and the crystalline fraction of the material.

On the other hand, the direct deposition of *nc*-Si:H has been found to occur as soon as hydrogen dominates the other gaseous components in the deposition chamber.[10] This fact has been reported by several authors which used, for instance, H_2-diluted silane in glow discharge technique or radio frequency sputtering with a plasma of nearly pure hydrogen.[10,11] In contrast with other methods, the control of growth, grain size, and crystallinity can be achieved, within certain limits, by sputtering technique using the simple change of the substrate temperature.

In this framework, a detailed analysis of *nc*-Si:H films

[a]Electronic mail: perez-ro@iris1.fae.ub.es

obtained by radio frequency sputtering deposition in a pure H_2 plasma is reported. The structure-properties relationship is investigated through a variety of techniques dealing with structural, optical, and transport properties. The growth of crystallites is described as resulting from complex interactions of silicon hydrides present in the plasma with the growing surface layer through continuous breaking or etching of strained near surface Si–Si bonds. The incorporation of hydrogen and its bonding configurations are analyzed, being these key factors determining the crystallization mechanism which is accompanied by hydrogen desorption and the possible formation of microvoids.

II. EXPERIMENTAL DETAILS

The samples investigated were deposited on glass substrates (Corning 7059) for the structural, optical, and electrical measurements and on single crystal ⟨100⟩ Si substrates for infrared measurements. All the substrates were submitted to a standard chemical cleaning (successively) in trichlorethylene (80 °C, 5 min), aceton (room temperature, 5 min), boiling propanol (5 min), and then rinsed in deionized water and dried in blow nitrogen. Prior to this cleaning, the monocrystalline silicon Si (c-Si) substrates used for infrared measurements were polished on both faces (to avoid scattering on rough surfaces) using diamond paste with decreasing granulometric size down to 1 μm. To remove the polishing-induced damage (cold work), the mechanically polished c-Si substrates were chemically etched at room temperature during 2 min in acidic solution of H_2SO_4 diluted in H_2O_2, in order to form a thin oxide layer at the substrate surfaces. This oxide film was subsequently removed by 10% diluted HF during 10 s. The silicon substrates were then rinsed with propanol and dried with a nitrogen flux. Before deposition, the sputtering chamber was backed out by a cryogenic pump to a pressure lower than 1×10^{-7} Torr. The atmosphere during the deposition was pure hydrogen gas (99.9999%) at a pressure of 0.4 Torr and the radio frequency power density was fixed at 0.6 W/cm^2. The substrate temperature (T_s) was varied between 50 and 250 °C. This temperature was measured by means of a Chromel/Alumel thermocouple wedged, close to the substrates, on the sample holder. The thickness of the deposited films is in the range 0.5–1.0 μm.

The hydrogen content (C_H) and bonding configuration were obtained from the infrared (IR) absorption spectra in which Si–H_n species have the main vibration modes. The IR spectra were measured with a Nicolet 750-II spectrometer. Suitable bare Si substrates were used as references. The optical absorption coefficient, optical band gap, and refractive index were determined from optical absorption measurements from 0.4 to 4.1 eV using a Perkin Elmer UV/VIS spectrometer.

The structural analysis of the samples was performed by Raman scattering, transmission electron microscopy (TEM), and x-ray diffraction (XRD) techniques. Raman scattering measurements were done with a Jobin-Yvon T64000 spectrometer coupled with an Olympus metallographic microscope, in backscattering configuration. Both the excitation and collection of the scattered light were done through the microscope objective. The excitation was made with the 514.5 nm line of an argon ionized laser, for which the penetration depth of scattering light is estimated to be of about 300 nm in single crystal Si. The diameter of the light spot on the sample was slightly submicronic. To avoid heating effects in the spectra, the excitation power on the samples was kept below 0.26 MW/cm^2. All the spectra were systematically compared to those obtained on bulk silicon under the same experimental conditions. The films were observed in cross-section TEM using a JEOL 200 CX operating at 200 kV. For this, samples were mechanically grinded, dimpled down on the two faces, and then thinned with the ion milling. The measurement of the average grain size and texture analysis of the layers were performed by XRD measurements using a high resolution Phillips instrument in the standard θ-2θ configuration. For each film the three most intense diffraction peaks (corresponding to the reflections ⟨111⟩, ⟨220⟩, and ⟨331⟩) were recorded, and the intensity of the different peaks was corrected by the appropriate geometrical angle dependent factors and normalized to those of a randomly oriented Si powder.

Photoluminescence (PL) measurements were carried out between 4.2 K and room temperature with the sample mounted in an Air Products cryostat. The excitation source was the 488 nm line of an Ar laser and the emitted signal was detected by a photomultiplier after dispersion by a 1 m grating monochromator.

Finally, the electrical conductivity of the films was measured in darkness at different temperatures between 77 and 400 K using a 6517 Keithley electrometer. For these measurements, planar In electrodes were obtained by deposition with electron gun evaporation and annealing at 140 °C for 20 min. The Ohmic character of the In contacts was checked at room and liquid nitrogen temperatures for voltages as high as 100 V.

III. RESULTS AND DISCUSSION

A. Morphology and structure of the nanocrystallized layers

1. TEM analysis

TEM observations were performed in all the samples whose images and corresponding electron diffraction patterns provide clear evidence of the formation of Si grains. Figure 1 shows bright field micrographs of two samples which are representative of the films obtained with low and high T_s, respectively. The cross-section image of the sample at T_s=50 °C [Fig. 1(a)] shows the presence of a 300 nm in average thick amorphous buffer layer on the substrate. Since the thickness of this layer is 550 nm, the amorphous buffer is a considerable fraction of the whole thickness and thus might influence the main optical and electrical properties of the film. Such an underlying amorphous layer has been already observed in previous works dealing with direct deposition of μc-Si films.[12,13] This layer appears in all the samples deposited at $T_s \leq 100$ °C, being absent for higher values of T_s, as shown in Fig. 1(b).

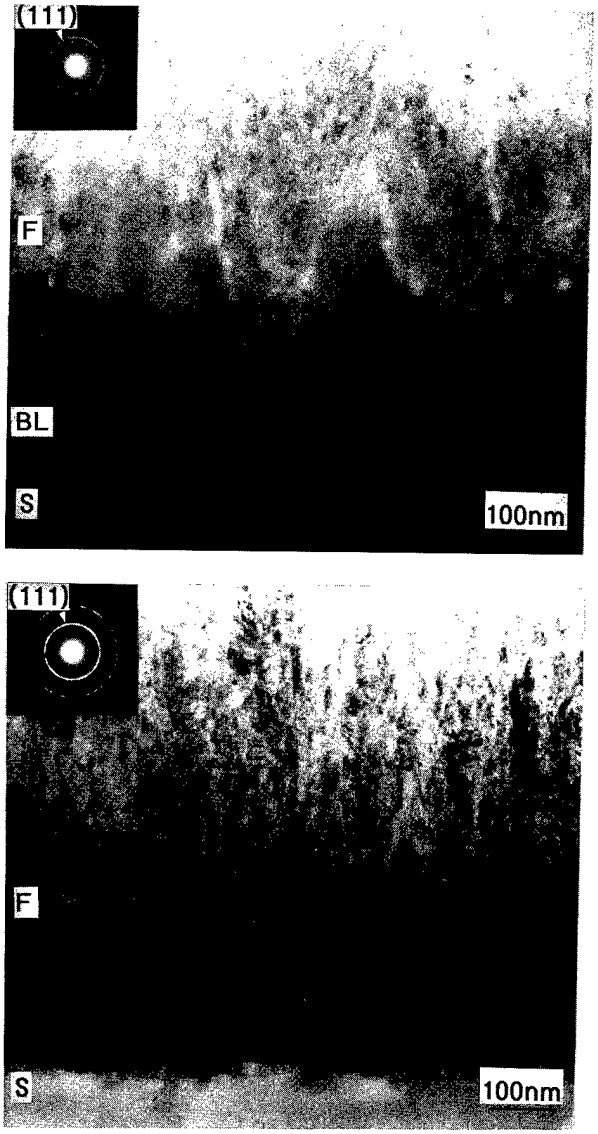


FIG. 1. Bright field transmission electron microscopy micrographs of the samples deposited at substrate temperature of (a) T_s=50 °C and (b) T_s=250 °C.

The crystalline phase in the films shows a columnar structure, with each column restricted in length for low T_s and consisting of aggregates of crystallites with the same orientation. The columnar structure originates from the crystallization induced by the reactive Si–H_2 species because of its low surface diffusivity compared to that of Si–H_3.[13,14] This is suggested by the relatively high concentration of dihydrides in our samples (see IR spectra). In the sample with T_s=50 °C [Fig. 1(a)], most of the columns show a conelike shape.

On the other hand, the interface with the substrate for T_s>100 °C is characterized by contrast fringes which are typical of the presence of a high density of defects like twins. In the region of the film close to the substrate smaller crystallites and shorter columns are observed in comparison with the rest of the layer. This suggests the presence of a disordered interface with a high density of defects and likely a higher amorphous component. Although this is not observed in the TEM images, we have to bear in mind that the TEM contrast is dominated by the crystalline component and, so, TEM usually underestimates the amorphous fraction. This would explain why for the samples grown at the highest T_s (250 °C) many columns are observed to extend through most of the thickness of the layer.

This behavior suggests that crystallites develop near the substrate when the growth proceeds and act as seeds for the growth of columns. Moreover, there is a gradation in the size of the crystallites, which increases when they are more distant from the interface. The lower the growth temperature, the higher the interface region with a gradation of crystallite sizes. For low T_s, cones arise connected to the random event that occurs when a crystallite with enough size to act as a seed is formed. For $T_s \leq 100$ °C, an amorphous layer grows before crystallization occurs. Upon increasing the substrate temperature, the thickness of the amorphous layer recedes, small precipitates start to form, and the amorphous content in this region decreases. This is due to the increasing diffusivity and reactivity of silicon hydrides, whose interactions with the surrounding bonds result in breaking of the weak and strained bonds, leading to the relaxation towards the crystalline state and desorption of hydrogen.[13,14] As will be shown later, this agrees with the amount of hydrogen and density of microvoids in the layers.

2. Raman scattering

The Raman spectra measured in the layers show the contribution of both amorphous and crystalline Si. The Raman spectrum of amorphous Si is characterized by four broad bands, which can be fitted with Gaussian curves. These bands are located at about 480 cm^{-1} (transverse optic), 380 cm^{-1} (longitudinal optic), 310 cm^{-1} (longitudinal acoustic), and 150 cm^{-1} (transverse acoustic). For crystalline Si, the Raman spectrum only shows a first order line with Lorentzian shape, centered at 521 cm^{-1} and with a full width at half maximum of about 3 cm^{-1}. This corresponds to a triply degenerated zone center vibrational mode.

Figure 2 shows the Raman spectra as measured from the top surface of the layers for layers obtained at different temperatures. Even without further analysis, the direct inspection of these spectra shows an increase of the crystalline fraction with T_s. This is indicated by the increasing contribution of the 521 cm^{-1} line in relation to the 480 cm^{-1} amorphous band. We have to remark that these spectra give information corresponding to the surface region of the layers. As already indicated, for the used excitation wavelength (514 nm), the penetration depth of scattered light is of about 300 nm for single crystal Si. However, for *nc*-Si films with a significant amorphous fraction a significant decrease of penetration depth is expected. According to this, Raman scattering measurements have also been performed from the back side, focusing the light spot on the film-glass interface. In Fig. 3 are plotted the spectra of the layers deposited at T_s=50 °C from both top and back sides. From this figure it is clear that the

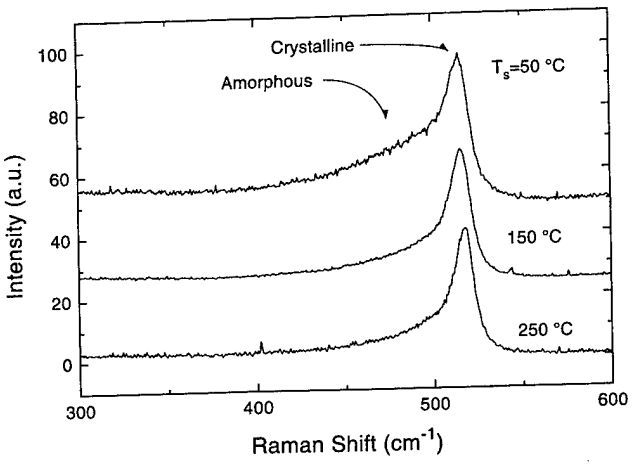


FIG. 2. Raman spectra from the top surface of the layers for layers obtained at different temperatures, T_s.

contribution to the crystalline band is very small for the back side spectrum, in agreement with the thick amorphous buffer layer observed by TEM [Fig. 1(a)]. The sample deposited at 100 °C showed a higher amorphous contribution from the back side as well. However, for the layers obtained with T_s>100 °C, the spectra measured from both sides are the same, in agreement with the absence of the buffer amorphous layer as observed by TEM.

The amorphous contribution to the spectra has been quantified by fitting the corresponding bands with Gaussian curves. Then, the crystalline contribution was obtained by substracting the amorphous one from the experimental spectra. The obtained crystalline spectra show an asymmetric shape, with a broadening towards the low frequency side in relation to the reference spectrum from single crystal Si. This shape is characteristic of phonon confinement effects induced by disorder in the lattice. These effects have been quantified by a correlation length model.[15–17] The geometry of confinement is also of importance and the best fitting to our results is obtained when spherical confinement is used.

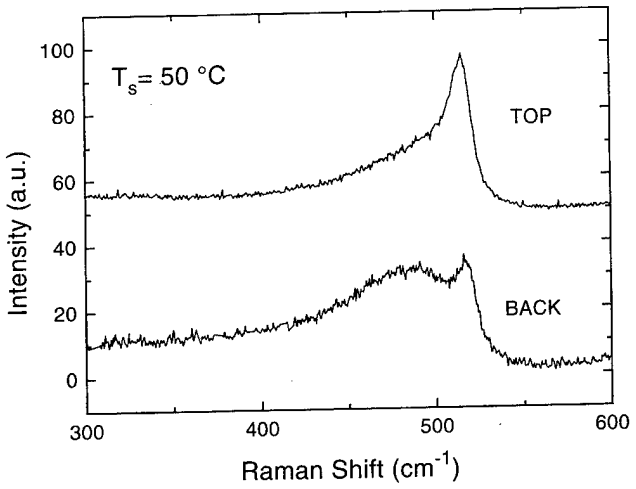


FIG. 3. Raman spectra of the layer deposited at substrate temperature T_s=50 °C from both top and back sides.

The correlation length (L) obtained from the model is assigned to the average distance between defects which break the translational symmetry and then confine the phonons, as dislocations, twins and grain boundaries. Hence, a high density of extended defects in the crystal grains will lead to correlation lengths smaller than the grain size. On the other hand, for crystal grains with low amount of defects, the correlation length would give an estimate of the grain size.

This model has allowed us to obtain an accurate fitting of the crystalline contribution in the spectra. This also points out the absence of stress effects in the spectra. So, in a general way the stress in the crystals can also be estimated by assuming a model in which both the correlation length and the stress are uniformly distributed in the scattering volume. In such a case the stress does not affect the shape of the Raman spectrum and leads to an additional shift of the spectra.[17–19] The absence of such a shift in our spectra indicates that the level of residual stress in the grains is below a value of about 10 MPa, taking into account our experimental accuracy.[17]

The crystalline fraction of the layers (X_c) has also been calculated from the Raman spectra, from the integrated intensities of the crystalline and amorphous bands, I_c and I_a, respectively. The crystalline fraction was determined from

$$X_c=\frac{I_c}{I_c+yI_a}, \tag{1}$$

where y is the ratio of the Raman efficiencies for crystalline and amorphous silicon. Tsu *et al.*[20] have reported that the value of y strongly depends on the grain size and becomes larger when grain size diminishes down to a few tens of nanometers. Bustarret *et al.*[15] have studied the relationship between the Raman intensity ratio of crystalline and amorphous silicon as a function of the crystalline size. Assuming phonon confinement in spherical crystals, these authors have fitted their results with the following function:

$$y(\delta)=0.1+\exp(-\delta/250), \tag{2}$$

where δ is the domain size of the crystallites. This expression is valid for δ>30 Å, a value which is also the minimum for the thermodynamical stability of a silicon nanocrystal. For this calculation, we have used in each case the value estimated for the correlation length L.

The data estimated from the Raman spectra (correlation length L and crystalline fraction X_c) for the different samples are shown in Fig. 4, versus the deposition temperature T_s. As it is shown, both parameters tend to increase with T_s. L increases from 50 Å for T_s=50 °C to 62 Å for T_s=250 °C. On the other hand, the absence of stress effects in the Raman spectra is likely due to the presence in the layers of a significant amount of microvoids, which allows relaxation of the residual stress. The formation of microvoids in the layers is also attested by optical measurements, as will be shown later. From previous studies with samples performed on layers recrystallized by annealing we have observed that a more compact tissue is strongly correlated with a stressed nanocrystal pattern.[11]

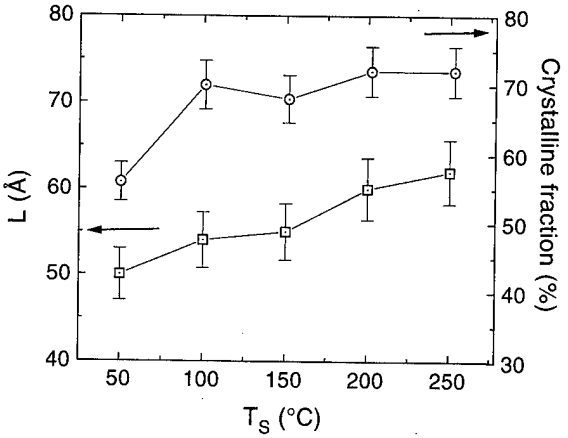


FIG. 4. Evolution of correlation length (L) and crystalline fraction (X_c) vs substrate temperature, T_s.

3. Hydrogen content and bonding configuration (infrared)

The concentration of bonded hydrogen in atomic percentage, C_H, has been estimated from the area under the Si–H wagging mode around 640 cm^{-1}, by taking the oscillator strength value recently determined by Langford *et al.*,[21] $A_w = 2.1\times10^{19}$ cm^{-2}. The results show that C_H decreases monotonically from 9.1% for T_s=50 °C down to 1.9% for T_s=250 °C, in concordance with the increasing crystallite size and crystalline fraction. The evolution with substrate temperature is shown in Fig. 5. Although a thin native oxide layer and hydroxides are likely formed on the surface of all deposits due to ambient exposure, they were undetected by the transmission infrared measurements. This might be related to the very small thickness of the native oxide layer (2–3 nm) compared to the thickness of the layers (microns).

The IR absorption bands due to the stretching vibrational modes of the Si–H_n species are located in the region 2000–2100 cm^{-1}. The experimental spectra are plotted in Fig. 6(a) and reveal stretching vibrations around 2090 cm^{-1} and a substantial reduction of the hydrogen content with increasing T_s. It is important to notice the quasiabsence of the vibrational mode at 2000 cm^{-1}, normally attributed to isolated

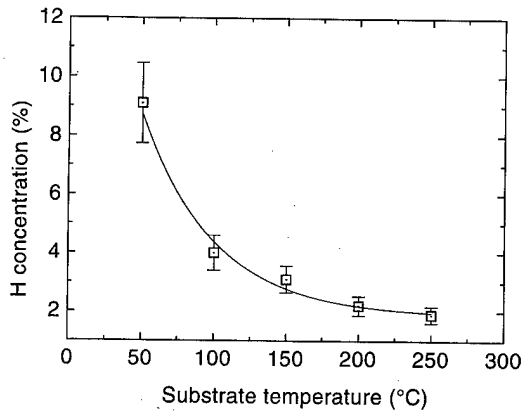


FIG. 5. Concentration of bonded hydrogen, C_H, as a function of substrate temperature, T_s.

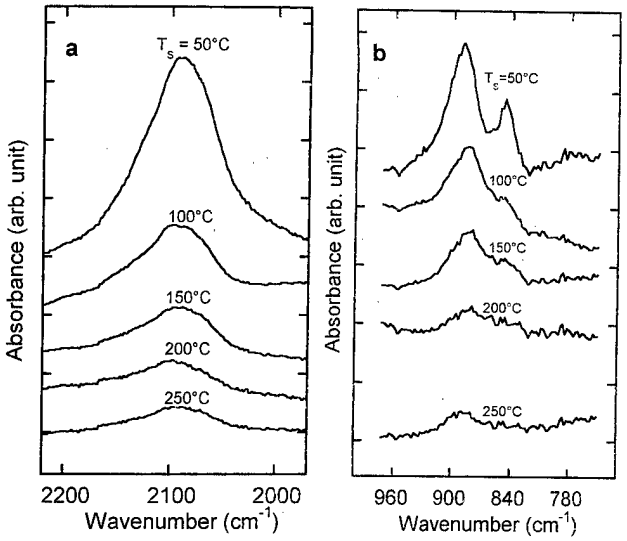


FIG. 6. Experimental infrared spectra of the samples with substrate temperature, T_s, as a parameter, (a) stretching bands and (b) bending bands.

Si–H bonds in a-Si:H. This suggests that hydrogen atoms in the amorphous tissue bonded to Si atoms are mainly in a dihydride configurations Si–H_2. This is supported by the detection of the bending mode at 890 cm^{-1} and to a lesser extent the mode at 840 cm^{-1} which is usually attributed to $(Si–H_2)_n$ chains [Fig. 6(b)]. This interpretation is also supported by the fact that both the contribution of the modes at 890 cm^{-1} in the spectra and the amorphous fraction in the films decrease as substrate temperatures increases. Nevertheless, Si–H single bonds at the crystalline ⟨111⟩ surface also absorb in the 2085–2090 cm^{-1} region, while dihydride species at the crystalline surface are shifted up to 2115 cm^{-1}. According to this the absorption bands plotted in Fig. 6(a) can also have a significant contribution from Si–H bonds in ⟨111⟩ surfaces.

As already mentioned and reported by many authors,[1,4–11] the crystallization of the layers is due to the diffusivity and reactivity of silicon hydrides, especially SiH_2, which obviously increase with the substrate temperature. Hydrogen is swept out from the recrystallized volume and stabilizes the Si nanocrystal by passivating its surface. Our results can be explained with these models, assuming that the hydrogen plasma creates a subsurface region of the nanocrystal with a lower connectivity between Si atoms and where the various interactions with the hydrogen species would induce, as mentioned above, the relaxation of the amorphous network to the crystalline configuration. Moreover, by increasing the temperature of deposition some competition between the growing and etching of the silicon surface atoms is also expected, especially for the strained Si–Si bonds. This process would be accompanied by desorption of molecular hydrogen and/or silicon hydrides (SiH_4,Si_2H_6).

The dihydride configuration is characteristic of the passivation of Si ⟨100⟩ planar surfaces (two silicon dangling bonds) while the monohydride configuration is characteristic of the more dense ⟨111⟩ Si planes. So, it has been reported that after a HF dip, planar ⟨111⟩ silicon surfaces are com-

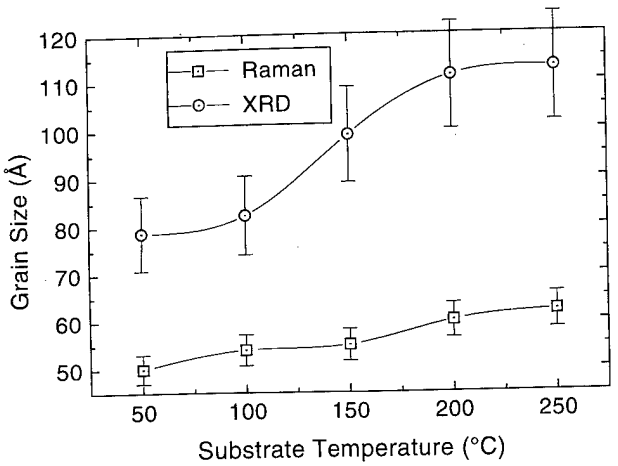


FIG. 7. Crystal size obtained from x-ray diffraction and correlation length (L) measured from Raman, vs substrate temperature, T_s.

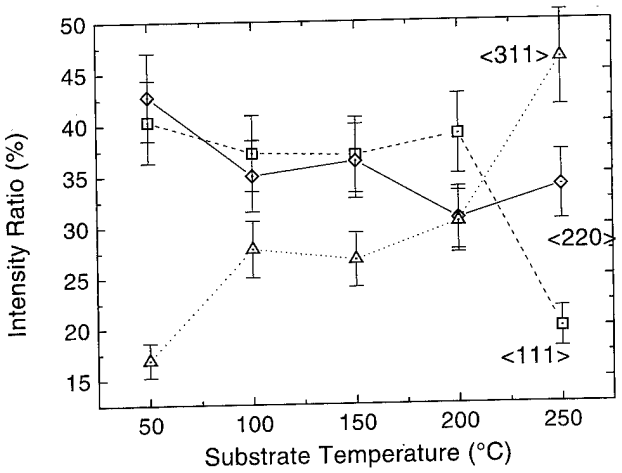


FIG. 8. Normalized intensities of the ⟨111⟩, ⟨220⟩, and ⟨331⟩ x-ray reflections vs substrate temperature, T_s.

pletely passivated by SiH.[22,23] On the contrary, SiH_2 species passivate mainly ⟨100⟩ planar surfaces. Infrared spectra indicate the presence of SiH_2 lying in the amorphous tissue and SiH passivating crystallite surfaces. This suggests the nanocrystals to be limited by ⟨111⟩ planes with terraces and steps that might be occupied by SiH and SiH_3 species, respectively. The contribution of other hydride species (mainly SiH_3) leads to a widening of the stretching band centered at 2090 cm^{-1}. In principle, for smaller crystals an increase in the number of steps and defects is to be expected. This agrees with the higher contribution in the high frequency side of the IR spectra from the samples with smaller crystals (obtained at low T_s values), as can be seen in Fig. 6(a). Anyway, further experiments are needed to clarify the role of the hydrogen as a passivating agent of the crystalline surface.

4. XRD analysis

The average grain size and texture in the layers have been estimated by x-ray diffraction measurements. The average grain size can be obtained from the width of the diffraction peaks, being given by the Scherrer formula: $\delta = k\lambda/(W \cos\theta)$, where k is approximately 1, λ is the wavelength of the x rays, W is the width of the band after correcting the instrumental contribution and θ is the Bragg angle of the diffracting planes. In Fig. 7 are plotted the crystal sizes obtained by XRD together with the correlation length estimated from the Raman spectra. All these data have a common upward trend with the deposition temperature T_s; however, the absolute values do not coincide. The data from the Raman spectra underestimate the crystal size by a factor of 2 as L corresponds to the average distance between defects in the nanocrystal.

The formation of textures and preferentially oriented growth of μc-Si and nc-Si is connected with the mechanism of crystal nucleation and has been reported by several authors.[24,25] Despite considerable interest, the detailed mechanisms for preferential growth are still unclear. Due to the anisotropy of the growth rate, the crystals that have the larger growth velocities in the direction normal to the surface will give a preferential orientation in the material. The intensities of the XRD peaks relative to a random oriented silicon powder give information related to the weight of certain crystallographic orientations. We have illustrated in Fig. 8 the normalized intensities for the ⟨111⟩, ⟨220⟩, and ⟨311⟩ reflections, considering that higher angle reflections have negligible contribution in the XRD spectra. For $T_s \leqslant 200$ °C, the lower angle reflection ⟨111⟩ has a similar contribution of around 40% to the total integrated intensity. In this range of T_s, ⟨311⟩ reflection is the smallest and tends to increase gradually at the expense of the ⟨220⟩ direction. For the highest T_s (250 °C) an unexpected drop of the ⟨111⟩ contribution is observed, while the ⟨311⟩ reflection steeply grows towards a final contribution of around 50%.

B. Optical and electrical characteristics

The film thickness, the absorption coefficient (α) and the refractive index (n) of the layers have been determined from appropriate calculations by using the transmission spectra. This approach has been first proposed by Manifacier *et al.*,[26] greatly improved afterwards by Swanepoel[27] and briefly summarized in our previous study.[28] The optical energy band gap, E_0, is commonly determined by extrapolating the linear plot of $(\alpha E)^{1/2}$ versus $E - E_0$, where E is the photon energy, to the $\alpha = 0$ value. Figure 9 shows the variation of E_0 versus T_s for the samples analyzed. E_0 decreases from a value as high as 2.40 eV for $T_s = 50$ °C down to approximately 1.95 eV for $T_s = 250$ °C.

The behavior of the optical band gap can be related to the structure and the role of hydrogen. A widening of the band gap has been previously reported for layers which were recrystallized by thermal annealing of sputtered samples.[9] A reduction of the localized defect states at the grain boundaries and/or passivation of dangling bonds with hydrogen lead to a consistent variation in the density of band-tail states which generate transitions with energy lower than the gap. To study these effects, The Urbach energy (E_u) has been determined in the samples.[9] This parameter corresponds to

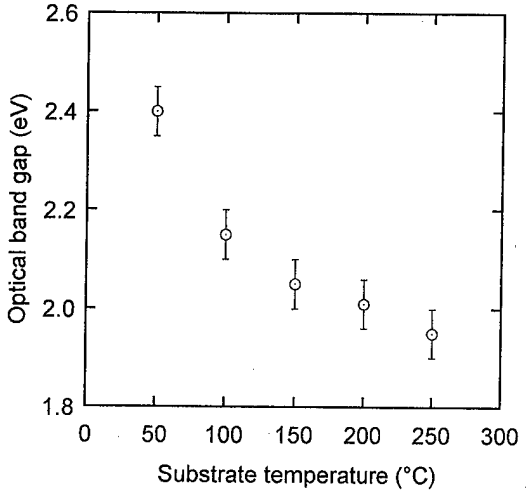


FIG. 9. Values of the optical band gap as a function of the substrate temperature, T_s.

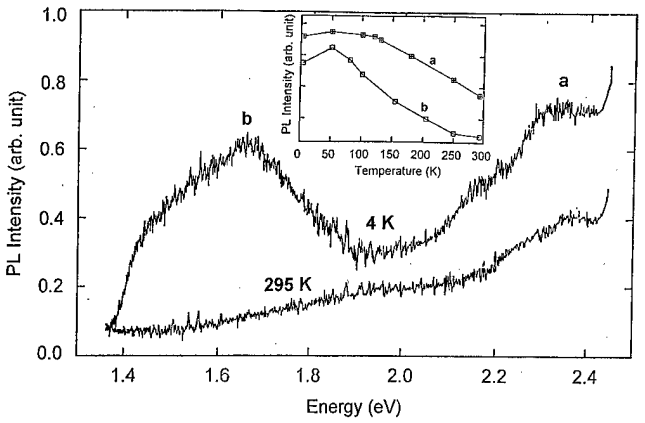


FIG. 10. Visible photoluminescence at 4.2 K and room temperature of the sample deposited at a substrate temperature of 50 °C. The inset shows the intensity of the peaks as a function of the measurement temperature.

the exponential-like behavior of the absorption versus the photon energy near the optical absorption edge. This band tail is originated from defect states that are induced by structural disorder.[29] E_u was found to increase in a gradual way from 120 meV for T_s=50 °C to 260 meV for T_s=250 °C. Compared to its value in a standard a-Si:H (50 meV approximately), the high values obtained reflect the microscope disorder of Si nanograins, likely related to the distribution of crystallite sizes and shapes, in addition to the disorder originated from the microstructure within the grain and the roughness of its surface.[30] Thus, the increasingly important values of E_u for high T_s are consistent with the increase of the amount of grains and/or crystalline fraction, which are expected to induce more disorder.

Besides, other effects cannot be excluded. Remarkably, for samples deposited at low temperatures (T_s=50 and 100 °C), the optical band gap lies between 2.40 and 2.15 eV, and the correlation length is around 5–6 nm according to Raman measurements. To clarify this, photoluminescence measurements have been performed at different temperatures from 4.2 K to room temperature. In Fig. 10 are plotted the spectra measured at 4.2 K and room temperature (RT) from the layer grown at T_s=50 °C. The spectra show two emission bands, located in the red (1.63 eV) and blue (2.30 eV) regions of the visible spectrum, whose intensities vary with the measurement temperature as indicated by the inset. Only the blue one survives with the half of its intensity at RT, whereas the red one shifts slightly to lower energy (1.59 eV) and is detectable between 4.2 and 120 K for the film deposited at T_s=100 °C. For the other layers, no PL is detectable in the visible range. The appearance of PL emission in the visible range might support the existence of quantum confinement effects. However, quantum confinement effects can hardly explain these PL features as they require grain sizes of 2–3 nm.[5] The high energy PL peak persists only on that sample deposited at 50 °C which is the most hydrogenated and which contains $(SiH_2)_n$ (n>2) chains (polysilane). This is evidenced by the 840 cm^{-1} infrared band detected for this sample. Moreover, hydrogen passivates dangling bonds at the surface of the nanocrystals, thus eliminating possible pathways for nonradiative recombinations. According to this, one cannot exclude emission as being originated in the surface of the nanocrystals. So, interface states passivated by hydrogen and/or polysilane in the amorphous tissue as well as surface oxygen passivation could also be involved in the measured photoluminescence.

On the other hand, the refractive index has been estimated from optical transmission measurements, and shows a monotonic increase from n=2.5 (T_s=50 °C) up to n=2.9 (T_s=200 °C). As both amorphous and crystalline Si have comparable values of refractive index, the fraction of voids in the layers can be estimated by considering the material as a composite with voids (n=1) embedded in silicon and thus applying suitable models for averaging the dielectric function (Bruddemann model). We have represented in Fig. 11 the fraction of voids determined by this method together with the hydrogen concentration. As shown in the figure, there is a strong correlation between both magnitudes. The fraction of voids is maximum for the sample at T_s=50 °C, for which the hydrogen content is also maximum (C_H=9.1%) and the

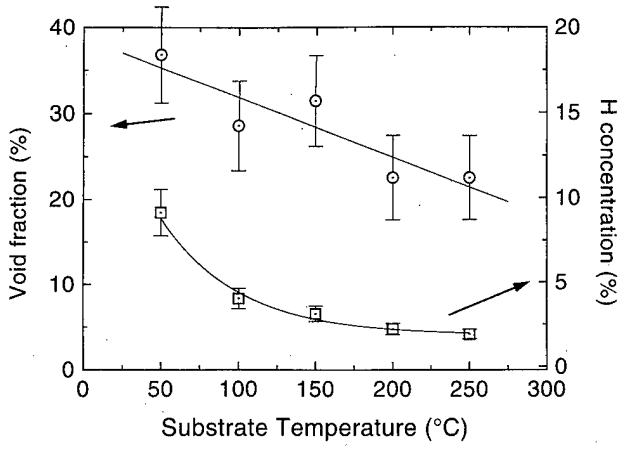


FIG. 11. Fraction of voids and hydrogen concentration (C_H) vs substrate temperature (T_s).

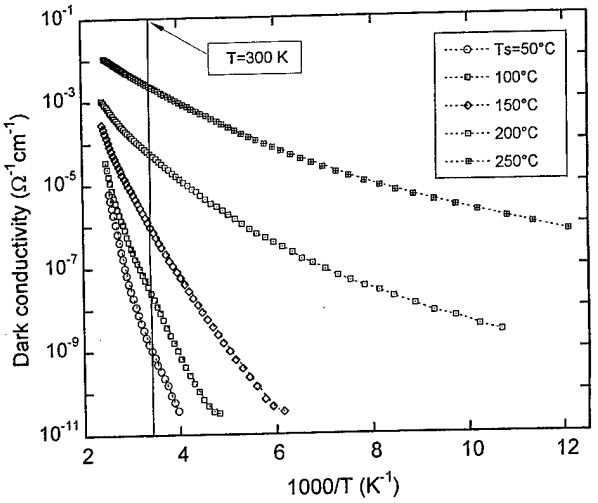


FIG. 12. Dark conductivity measurement for different substrate temperature (T_s), as a function of the measurement temperature (T).

correlation length is minimum (L=50 Å). Moreover, the overall quantity of voids for all the samples as extremely high ($\geqslant$22%) suggesting that a porouslike tissue is developed for these directly deposited samples. A similar porouslike tissue has been claimed by others authors to explain the initial growth stage and the structural development of μc-Si deposited by plasma enhanced chemical vapor deposition on glass.[31,32] The growth of this tissue is related to the fact that the high porosity for all T_s leads to the relaxation of the nanocrystals to the most favorable atom coordinates, leading to residual stress levels below the sensitivity of the Raman spectra.

Finally, further information concerning the growth of crystallites is provided by the dark conductivity measurements. Figure 12 shows the dark conductivity as a function of the measurement temperature (1000/T) and with substrate temperature (T_s) as a parameter. The activation energy decreases from 1.25 eV for T_s=50 °C to a value as low as 0.13 eV for T_s=250 °C, being this characteristic of a composite structure consisting of an amorphous tissue with a high crystalline fraction.[33] We have to notice the dramatic increase of conductivity from T_s=50 °C to T_s=250 °C of more than six orders of magnitude, reaching a value close to values characteristic of monocrystalline silicon. In addition, the reduction of the density of grain boundaries as grain size increases would induce the lowering of the activation energy. This can also be explained by the high crystallinity (>70%) achieved for the samples deposited at the higher temperatures. For these layers, the distance between crystals is of few atomic spacings, and tunneling of carriers between crystals becomes possible, which would give a low value of the activation energy for conductivity.

IV. CONCLUSIONS

The detailed structural, optical and electrical characterization of nc-Si:H layers obtained by sputtering deposition in a pure hydrogen plasma shows the ability to control the grain size, crystalline fraction, optical band gap and electrical conductivity in the layers by the temperature of deposition T_s. Increasing T_s from 50 to 250 °C leads to an increase of the average grain size (from few nm to a few tens of nm) and crystalline fraction (from 37% to 74%), and the optical band gap decreases from 2.40 to 1.95 eV. Besides, electrical conductivity shows a dramatic increase of more than six orders of magnitude. The presence of a significant content of voids in the layers allows relaxation of stress to values below those detectable by Raman measurements. On the other hand, the role of hydrogen species and particularly that of SiH_2 in the crystallization mechanism has been evidenced, being these effects greatly enhanced by an increase of the substrate temperature.

ACKNOWLEDGMENTS

This work has been partially funded by the "Acciones Integradas" cooperation program,, Ref. HF96-0128, and the Spanish CICYT under Project Ref. MAT96-1010-C02-01.

[1]C. Wang and G. Lucovsky, *Proceedings of the 21st IEEE PVSC* (IEEE, New York, 1990), p. 1614.
[2]S. D. Brotherton, D. J. McCulloch, J. B. Clegg, and J. P. Gowers, IEEE Trans. Electron Devices **40**, 407 (1993).
[3]T. Kamins, *Polycrystalline Silicon for Integrated Circuit Applications* (Kluwer, Norwall, MA, 1988), p. 160.
[4]N. Yamagushi, J. J. Hajjar, and R. Reif, IEEE Trans. Electron Devices **38**, 55 (1991).
[5]S. Veprek, *Polycrystalline Semiconductors IV*, edited by S. Pizzini, H. P. Strunk, and J. H. Werner, Vol. 51/52 (Scitec, Winterthur, Switzerland, 1996), p. 225.
[6]A. Itoh, T. Ifuku, M. Otobe, and S. Oda, Mater. Res. Soc. Symp. Proc. **452**, 749 (1997).
[7]A. Borghesi, M. E. Giardini, M. Marazzi, A. Sassella, and G. de Santi, Appl. Phys. Lett. **70**, 892 (1997).
[8]K. Tanaka, Mater. Res. Soc. Symp. Proc. **453**, 3 (1997).
[9]A. Achiq, R. Rizk, R. Madelon, F. Gourbilleau, and V. Voivenel, J. Phys. D: Condens. Matter (to be published).
[10]S. Veprek, Mater. Res. Soc. Symp. Proc. **164**, 39 (1990).
[11]B. Garrido, A. Achiq, J. Maćia, J. R. Morante, A. Pérez-Rodríguez, P. Ruterana, and R. Rizk, Mater. Res. Soc. Symp. Proc. **452**, 785 (1997).
[12]T. Kaneko, M. Wakagi, K. Onisawa, and T. Minemara, Appl. Phys. Lett. **64**, 1865 (1994).
[13]Y. Sun and T. Miyasato, Jpn. J. Appl. Phys., Part 2 **34**, L1248 (1995).
[14]T. Akasaka and I. Shimizu, Appl. Phys. Lett. **66**, 3441 (1995).
[15]E. Bustarret, M. A. Hachicha, and M. Brunel, Appl. Phys. Lett. **52**, 1676 (1988).
[16]Y. He, C. Yin, G. Chen, L. Wan, X. Liu, and G. Y. Hu, J. Appl. Phys. **75**, 797 (1994).
[17]J. Macía, E. Martin, A. Pérez-Rodríguez, J. Jiménez, J. R. Morante, B. Aspar, and J. Margail, J. Appl. Phys. (to be published).
[18]M. Yang, D. Huang, P. Hao, F. Zhang, X. Hou, and X. Wang, J. Appl. Phys. **75**, 651 (1994).
[19]H. Xia, Y. L. He, L. C. Wang, W. Zhang, X. N. Liu, X. K. Zhang, and D. Feng, J. Appl. Phys. **78**, 6705 (1995).
[20]R. Tsu, J. González-Hernández, S. S. Chao, S. C. Lee, and K. Tanaka, Appl. Phys. Lett. **40**, 534 (1982).
[21]A. A. Langford, M. L. Fleet, A. P. Nelson, W. A. Landford, and N. Maley, Phys. Rev. B **45**, 13 367 (1992).
[22]G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, Appl. Phys. Lett. **56**, 656 (1990).
[23]P. Dumas, Y. J. Chabal, and G. S. Higashi, Phys. Rev. Lett. **65**, 1124 (1990).
[24]A. Matsuda, K. Kumagai, and K. Tanaka, Jpn. J. Appl. Phys., Part 2 **22**, L34 (1982).

[25]Y. Sun, T. Miyasato, and J. K. Wigmore, Appl. Phys. Lett. **70**, 508 (1997).
[26]J. C. Manifacier, J. Gasiot, and J. P. Fillard, J. Phys. E **9**, 1002 (1976).
[27]R. Swanepoel, J. Phys. E **16**, 1214 (1983).
[28]R. Rizk, A. Achiq, R. Madelon, F. Gourbilleau, and F. Cruege, *Polycrystalline Semiconductors IV*, edited by S. Pizzini, H. P. Strunk, and J. H. Werner, Vol. 51/52 (Scitec, Winterthur, Switzerland, 1996), p. 243.
[29]G. D. Kody, T. Medje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. **47**, 1480 (1981).
[30]Y. Kanemitsu, Phys. Rev. B **53**, 13 515 (1996).
[31]G. N. Parsons, J. J. Boland, and J. C. Tsang, Jpn. J. Appl. Phys., Part 1 **31**, 1943 (1992).
[32]I. Simizu, J. Non-Cryst. Solids **114**, 145 (1989).
[33]D. Das, Phys. Rev. B **51**, 10 729 (1995).

Deep, vertical etching of flame hydrolysis deposited hi-silica glass films for optoelectronic and bioelectronic applications

A. J. McLaughlin,[a] J. R. Bonar, M. G. Jubber, P. V. S. Marques,[b] S. E. Hicks, C. D. W. Wilkinson, and J. S. Aitchison
Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

(Received 29 September 1997; accepted 10 April 1998)

We report on the use of CHF_3, C_2F_6, and SF_6 as etch gases for deep reactive ion etch processing of germano-boro-silicate glass films prepared by flame hydrolysis deposition (FHD). The glass film under study had a composition of 83 wt % SiO_2, 12 wt % GeO_2, and 5 wt % B_2O_3. The scope of the study was to identify the benefits and drawbacks of each gas for fabrication of deep structures (>10 μm) and to develop an etch process in each gas system. The etch rate, etch profile, and surface roughness of the FHD glass films were investigated for each gas. Etch rates and surface roughness were measured using a surface profiler and etch profiles were assessed using a scanning electron microscope. It was found that SF_6 had the highest FHD glass etch rate and nichrome mask selectivity (>100:1) however, it had the lowest photoresist mask selectivity (<1:2) and highest lateral erosion. CHF_3 had the lowest FHD glass etch rate but high selectivity over both nichrome (>80:1) and photoresist (>10:1) and the etch profile was found to be smooth and vertical. C_2F_6 had a similar etch profile to that of CHF_3, but the selectivity over both mask materials was lower than in CHF_3. Fused silica was used as a reference material where it was found the percentage drop in etch rate in C_2F_6, SF_6, and CHF_3 was -12%, -15%, and -37%, respectively. From the results presented here CHF_3 proved to be the most versatile etch process as either photoresist or nichrome masks could be used to attain depths of 20 μm, or more. © *1998 American Vacuum Society.* [S0734-211X(98)00704-5]

I. INTRODUCTION

Flame hydrolysis deposited (FHD) silica glass films are widely used in planar lightwave circuits (PLCs) due to their low loss and inherent compatibility with optical fibers.[1] Potential applications of these devices range from optical communications to automotive control systems. In these applications waveguide dimensions range from 4 to 50 μm. Recently there has been growing interest from industry and academia in the use of FHD films in bioelectronic applications such as fluid handling and optical biosensing where structures with dimensions of >20 μm are commonplace.[2] One potential limitation to the production of these devices is the dry etch processing.

Deep etching of silica glass films has been reported previously, for example using C_2F_6 in electron cyclotron resonance reactive ion beam etching (ECR-RIBE),[3–5] and CHF_3/Ar in hollow cathode reactors.[6,7] Numerous publications also exist on the mechanisms of SiO_2 for semiconductor applications, however to our knowledge, there are no reports which examine the problems and advantages of using different fluorine based etch gases for deep reactive ion etching (RIE) of FHD silica glass films.

The main objective of this study was to develop optimized CHF_3, C_2F_6, and SF_6 RIE processes which have high FHD glass etch rates, high mask selectivities, and low sidewall damage. In order to develop the process we examined how selection of reactive ion etch parameters, for example: etchant gas, pressure, and forward rf power, influenced the RIE characteristics of germano-boro-silicate glass films prepared by FHD and the potential masking materials. We outline the advantages and disadvantages of using CHF_3, C_2F_6, and SF_6 as etchant gases in deep reactive ion etching of flame hydrolysis prepared glass films.

II. EXPERIMENT

High silica films were formed by flame hydrolysis deposition.[1,8] In this process metal halides, such as $SiCl_4$, $GeCl_4$, and BCl_3, are hydrolyzed in an oxyhydrogen flame to form a low density oxide "soot." The hydrolysis reactions can be described by

$$SiCl_{4(v)} + 2H_2O_{(v)} \rightarrow SiO_{2(s)} + 4HCl_{(v)},$$

$$GeCl_{4(v)} + 2H_2O_{(v)} \rightarrow GeO_{2(s)} + 4HCl_{(v)},$$

$$2BCl_{3(v)} + 3H_2O_{(v)} \rightarrow B_2O_{3(s)} + 6HCl_{(v)}.$$

The resulting oxide "soot" is deposited on the surface of a thermally oxidized silicon wafer. The "soot" was subsequently annealed at 1300 °C to form a germano-boro-silicate glass (GBSG) film. The FHD film composition was estimated by Microprobe analysis to be 83 wt % SiO_2, 12 wt % GeO_2, and 5 wt % B_2O_3. The thickness measured by a prism coupling technique was found to be ~8 μm. The wafer was subsequently diced up to provide enough samples for etch process characterization.

[a]Electronic mail: A.McLaughlin@elec.gla.ac.uk
[b]Also at Centro de Física do Porto, Rua da Campo Alegre, 687 4150 Porto, Portugal.

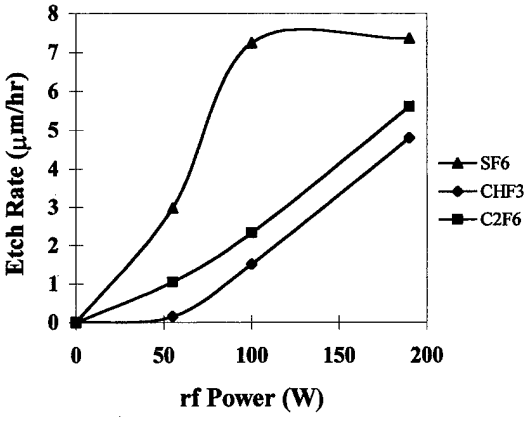


FIG. 1. Etch rate of flame hydrolysis deposited glass as a function of forward rf power in CHF_3, C_2F_6, and SF_6 (pressure 50 mTorr, flow rate 20 sccm). Thin film deposition was evident above 100 W in SF_6 and below 50 W in CHF_3.

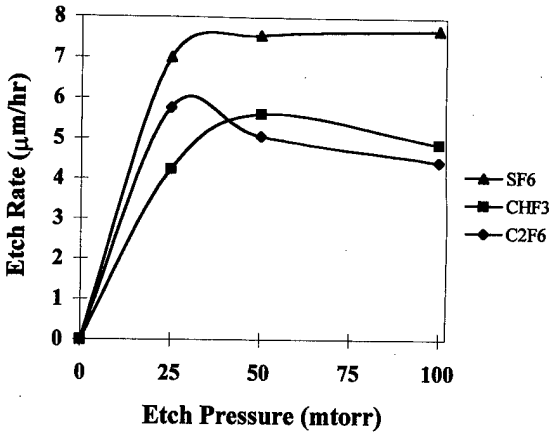


FIG. 2. Etch rate of flame hydrolysis deposited glass as a function of etch pressure in CHF_3, C_2F_6, and SF_6 (rf power 190 W and flow rate 20 sccm).

A Plassys electron beam evaporator was used to coat samples with nichrome (60% Ni: 40% Cr), which was subsequently patterned using a Shipley S1818 photoresist, and postbaked prior to the wet etch formation of the nichrome dry etch mask in a solution of ammonium ceric nitrate, glacial acetic acid, and RO water. Samples were also prepared which used photoresist only as the dry etch mask.

In this work we used an Oxford Plasmatechnology RIE80 reactive ion etching machine. The machine was cleaned prior to each etch run in an oxygen plasma and the sample was placed on the driven electrode. Etch characterization was restricted to pressure and rf power variation as the maximum C_2F_6 and SF_6 flow rates were machine limited to less than that of CHF_3 thus restricting the range over which the influence of flow rate could be studied. The etch tests were performed with a constant flow rate of 20 sccm. The rf power was varied in the range 30–200 W, and the pressure was varied in the range 25–100 mTorr for all gases.

A DekTak surface profiler was used to measure the mask thickness prior to etching. The etch depth was measured before and after mask removal, giving the FHD glass and mask etch rate as well as the glass to mask selectivity. The surface profiler was also used to determine the roughness of the etched surfaces.

A Hitachi S-800 scanning electron microscope was used to examine the etch profiles of the samples. Finally, fused silica samples were prepared and etched using the same processes for comparison of etch rates.

III. RESULTS AND DISCUSSION

A. Process development

1. CHF_3 development

The FHD glass etch rate with CHF_3 increased with increasing rf power, as shown in Fig. 1, above a threshold of 50 W, below which deposition of a thin film was evident. The photoresist mask etch rate decreased with increasing rf power, whereas the etch rate of the nichrome mask increased with increasing rf power, as did the dc bias. The FHD glass etch rate was also found to increase with pressure before saturating above 50 mTorr, as shown in Fig. 2. Above 50 mTorr a thin film was visible on the etched surface of the sample. The photoresist mask etch rate decreased with increasing pressure whereas the nichrome etch rate remained constant. The dc bias did not vary with pressure. The optimal process was obtained when the rf power, pressure, flow rate, and dc bias were set to 190 W, 50 mTorr, 20 sccm and −500 V, respectively. The mask selectivities and glass etch rate using this process are illustrated in Table I.

2. C_2F_6 development

In C_2F_6 the etch rate of the FHD glass was found to increase with rf power (Fig. 1), with no observed threshold for etching. The nichrome etch rate also increased as did the dc bias, whereas the photoresist mask etch rate was observed to decrease. From Fig. 2 we can see that the FHD glass etch rate increased with pressure before saturating above 25 mTorr. The photoresist mask etch rate was found to decrease

TABLE I. Performance of developed CHF_3, C_2F_6, and SF_6 etch processes for deep reactive ion etching of flame hydrolysis deposited silica glass films.

Gases	FHD glass etch rate (μm/h)	Nichrome (nm/h)	Photoresist (μm/h)	Sidewall profile of FHD glass	Fused silica etch rate (μm/h)
CHF_3	5.6	60	0.5	vertical	3.5
C_2F_6	5.8	150	2	vertical	5.1
SF_6	7.7	77	3.9	undercut	6.5

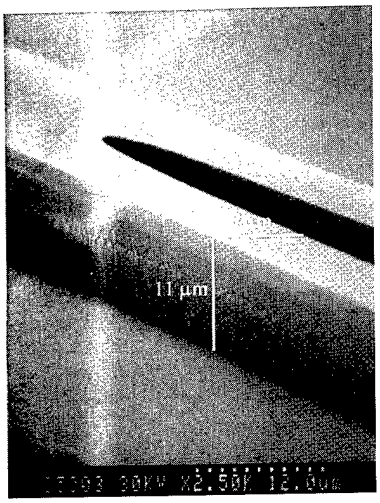


FIG. 3. FHD glass etch profile using CHF_3 RIE and a nichrome mask. Of note is the vertical profile and low lateral erosion of the sidewalls. The etched depth is 11 μm and ridge width 6 μm.

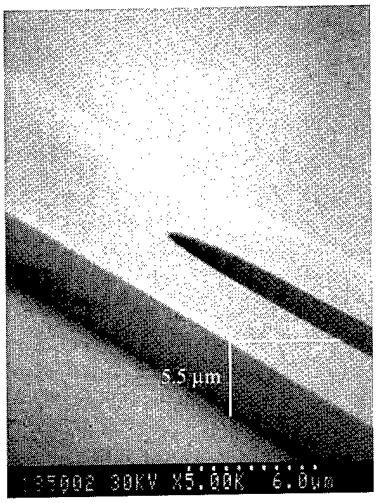


FIG. 4. FHD glass etch profile using C_2F_6 RIE and a nichrome mask. The etched depth was 5.5 μm.

whereas the nichrome mask etch rate remained constant, again with no change in dc bias. The etch process was optimized where the rf power, pressure, flow rate, and dc bias were set to 190 W, 25 mTorr, 20 sccm, and -500 V, respectively. The mask selectivities and glass etch rate using this process are illustrated in Table I.

3. SF_6 development

Using SF_6 the etch rate increased rapidly with increasing rf power, before saturating above 100 W as shown in Fig. 1. The etch rate of the nichrome mask also increased with rf power, as did the dc bias, however a high photoresist mask erosion rate was evident which increased with rf power. The etch rate of the FHD glass also increased with pressure as shown in Fig. 2, however the etch rate of the nichrome mask decreased, as did the dc bias. Again the photoresist mask erosion was high and increased with pressure. A thin film was evident on the sample at pressures >50 mTorr and at rf powers >190 W. The FHD glass etch rate and nichrome selectivity were optimized when rf power, pressure, flow rate, and dc bias were set to 190 W, 100 mTorr, 20 sccm, and 130 V. The mask selectivities and FHD glass etch rate using this process is presented in Table I.

B. Assessment of developed processes

1. Sidewall profile

Samples were prepared with nichrome masks and etched using the optimized processes already described. From the micrograph of Fig. 3 it can be observed that the CHF_3 process results in a vertical etch profile with no lateral erosion. A similar sidewall profile can also be observed under C_2F_6 etching (Fig. 4). This lack of lateral erosion indicates the presence of polymer protection of the sidewalls, typical of etching in fluorocarbon etch gases.[7] Sidewall erosion was evident using SF_6 as shown in Fig. 5. The lateral erosion can be attributed to two factors: the lateral erosion was minimized under conditions where the dc bias was high, and the obvious lack of fluorocarbon protection of the sidewalls. The amplitude of sidewall corrugation was estimated by scanning electron microscope investigation to be less than 150 nm using all etch processes, indicating that the sidewall roughness is mask, rather than process limited.

2. Comparison of etch rates with fused silica

Fused silica samples were prepared with nichrome masks and etched in the optimized processes to standardize each process as a means of comparison of the FHD glass etch properties. In C_2F_6 and SF_6 RIE the fused silica and FHD glass etch rates were found to drop by only 12% and 13%, respectively, whereas in CHF_3 the etch rate was found to fall by 37%.

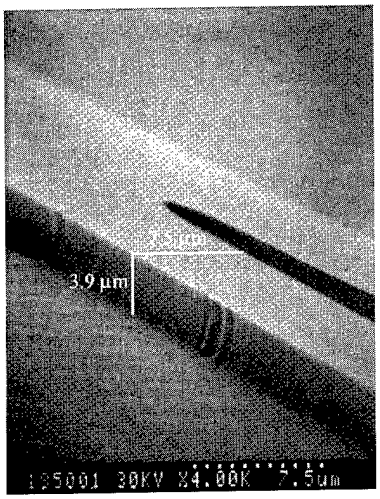


FIG. 5. FHD glass etch profile using SF_6 RIE and a nichrome mask. Of note is the defect in the mask which illustrates the lateral erosion of the sidewall. The etched depth was 3.9 μm.

3. Surface roughness

Any process where a surface is bombarded by reactive ions and neutrals can cause damage in the form of surface roughness. The surface roughness of the samples were measured using the DekTak surface profiler and in all cases the average surface roughness, R_{av}, was found to be less than 10 nm, which was less than the vertical resolution of the machine.

4. Comparison between the developed processes for deep etching

It is clear from the results presented here that nichrome masks are preferable to photoresist for deep RIE of FHD glass because of the enhanced mask:glass etch rate selectivity, especially in the case of SF_6 where the nichrome etch rate is very low. The nichrome etch rate is bias, and therefore ion energy, dependent as the bias under SF_6 etching is only -130 V compared to -500 V under fluorocarbon etching conditions.

From the results presented here we may deduce that fluorocarbon deposits play an important role in enhancing the mask:glass etch rate selectivity by protecting the mask surface from erosion. We suggest that this is the reason the photoresist mask etch rates were lower in C_2F_6 and CHF_3 RIE than SF_6 RIE. Previous studies have shown that selectivity can be improved through the inclusion of hydrogen in the fluorocarbon process, for example in CHF_3[9] or through H_2 addition to CF_4.[10] This is confirmed by the fact that the mask:glass selectivity is higher with CHF_3 compared to C_2F_6 processing. A fluorocarbon film protects the mask surface as well as the etched sidewalls reducing (or eliminating) the degree of lateral erosion. However, we have observed that too much fluorocarbon deposition leads to a grossly contaminated chamber resulting in a decrease in etch rate with time. Therefore a compromise between mask protection and high etch rates must be found dependent upon the application.

From the work presented here it is clear that CHF_3 RIE is the most suitable process for deep and vertical etching of FHD glass as it exhibits a high mask:glass etch rate selectivity and results in vertical sidewalls. SF_6 RIE, although having a high etch rate and attractive mask:glass selectivities, suffers from a high degree of undercutting. In optoelectronic applications, this would result in excessive optical losses and a large polarization sensitivity. Although C_2F_6 RIE of FHD glass results in vertical profiles, it suffers from a low mask:glass etch rate selectivity which limits the depth attainable.

IV. CONCLUSIONS

In summary, the implications for deep etching (>10 μm) of flame hydrolysis deposited silica glass using CHF_3, C_2F_6, and SF_6 as etchant gases have been presented. From the experimental data obtained, CHF_3 presents the most versatile etch process, in terms of FHD glass etch rate and high selectivity over both photoresist and nichrome. However higher etch rates are obtainable with the other etchant gases at the expense of selectivity, and in the case of SF_6, sidewall quality. Furthermore, it has been illustrated that SF_6 has a high selectivity over nichrome, which could be exploited by using a CF_x generating gas for increased sidewall protection. The experimental data also illustrate that doped silica glasses prepared by FHD have higher etch rates compared to pure fused silica. Work is ongoing to optimize the RIE of doped silica glasses through the use of mixtures of etch gases.

ACKNOWLEDGMENTS

The authors acknowledge the assistance of the technicians in the dry etch group for their invaluable assistance to this work. A. J. McLaughlin is supported by Lucas Industries, Advanced Vehicle Systems Development, and by EPSRC under a CASE award. M. G. Jubber acknowledges the support of the EPSRC under a ROPA Award. P. V. S. Marques acknowledges the Portuguese Research Council, JNICT, for financial support through the Program PRAXIS XXI.

[1]M. Kawachi, Opt. Quantum Electron. **22**, 391 (1990).
[2]G. L. Duveneck, E. Verpoorte, P. Oroszlan, M. Pawlak, C. Erbacher, A. Spielmann, D.Neuschäfer, and M. Ehrat, Anal. Methods Inst. Special Issue μTAS'96, 158 (1996).
[3]A. K. Dutta, Jpn. J. Appl. Phys., Part 1 **34**, 365 (1995).
[4]A. K. Dutta, Jpn. J. Appl. Phys., Part 1 **34**, 1663 (1995).
[5]A. K. Dutta, J. Vac. Sci. Technol. B **13**, 1456 (1995).
[6]M. V. Bazylenko and M. Gross, Appl. Phys. Lett. **69**, 2178 (1996).
[7]M. V. Bazylenko and M. Gross, J. Vac. Sci. Technol. A **14**, 2994 (1996).
[8]M. Kawachi, M. Yasu, and T. Edahiro, Electron. Lett. **19**, 583 (1983).
[9]G. S. Oehrlein and H. L. Williams, J. Appl. Phys. **62**, 662 (1987).
[10]H. W. Lehmann and R. Widmar, J. Vac. Sci. Technol. **15**, 319 (1978).

Sidewall slope control of chemically assisted ion-beam etched structures in InP-based materials

J. Daleiden,[a] K. Czotscher, C. Hoffmann, R. Kiefer, S. Klussmann, and S. Müller
Fraunhofer-Institut für Angewandte Festkörperphysik, Tullastrasse 72, D-79108 Freiburg, Germany

A. Nutsch
Walter Schottky Institut, Am Coulombwall, D-85748 Garching, Germany

W. Pletschen and S. Weisser
Fraunhofer-Institut für Angewandte Festkörperphysik, Tullastrasse 72, D-79108 Freiburg, Germany

G. Tränkle
Ferdinand Braun Institut für Höchstfrequenztechnik, Rudower Chausse 5, D-12489 Berlin, Germany

J. Braunstein and G. Weimann
Fraunhofer-Institut für Angewandte Festkörperphysik, Tullastrasse 72, D-79108 Freiburg, Germany

(Received 29 September 1997; accepted 10 April 1998)

Chemically assisted ion beam etching (CAIBE) of InP-based materials has been newly developed with BCl_3/Ar in comparison to Cl_2/Ar and IBr_3/Ar. Using halogen gases and an argon ion beam at 400 V a very good surface morphology was obtained at a low substrate temperature of −5 °C in any case; with BCl_3/Ar a surface roughness of 0.2 nm was observed. The etch rates were in the range of 40–75 nm/min depending on the reactive gas. Mixing the reactive gases or tilting the substrate with respect to the impinging ion beam allowed us an excellent control of the etched sidewall slope. By mixing BCl_3 and IBr_3 we were able to tune the sidewall slopes between 15° (measured to the surface normal) for pure BCl_3 and 38° for the pure IBr_3, respectively. Tilting the substrate allowed us to adjust the slope angle between 0° and 60°. In addition we have analyzed the etched surfaces by energy dispersive x-ray measurements. The low temperature processes yielded stoichiometric InP surfaces, etching at higher substrate temperatures results nonstoichiometric surfaces. These low temperature halogen CAIBE processes were successfully applied for the fabrication of gratings, ridge waveguides and facets for long wavelength (InGa)(AsP) and (AlInGa)(AsP) laser diodes. In this article we present ridge waveguide (InGa)(AsP)/InP lasers (1.55 μm) with direct CW modulation bandwidths of 9.5 GHz. © *1998 American Vacuum Society.*
[S0734-211X(98)00604-0]

I. INTRODUCTION

For the fabrication of optoelectronic devices anisotropic etch processes are needed which yield smooth sidewalls as well as flat bottom surfaces. Dry etching techniques in general offer high anisotropy, a good control on the etch profiles and an adjustable etch selectivity. Several dry etching techniques are used for the processing of (InGa)(AsP) and (AlInGa)(AsP) layers which are important for long wavelength telecommunication systems. Methane–hydrogen processes often suffer from polymer formation,[1–3] while processes with chlorine usually have to be run at rather high substrate temperatures[4] (>150 °C) or high ion energy[5] (>1 keV) because of the low volatility of the $InCl_x$ reaction products. With these high process temperatures a standard photoresist (PR) masking is not possible. Recently we have shown, that chemically assisted ion beam etching (CAIBE) of InP-based materials is possible using halogen gases at low substrate temperatures.[6] In this article we report on a newly developed low temperature CAIBE process with BCl_3/Ar for the etching of InP-based materials and compare it to CAIBE with Cl_2 and IBr_3.

II. EXPERIMENT

The samples were etched in a Technics Plasma Ribetch 160 ECR UHV LL system,[7] which is equipped with an electron cyclotron resonance ion source having a diameter of 160 mm. The load locked chamber is pumped by a cryopump and has a base pressure of 2×10^{-8} mbar. The reactive gases are introduced through four centrosymmetric tubes directed towards the substrate. Ar was used as sputter gas at a bias of 400 V. For good thermal contact to the cooled substrate holder the samples were mounted using a 1-mm-thick silicone pad. The temperature of the substrate holder was controlled by a thermocouple, while the surface temperature of the sample was monitored by a pyrometer. The samples were masked by the photoresist AZ 1518 which had to be hardened in an oven at 120 °C for 30 min.

The surface roughness was analyzed by atomic force microscopy (AFM) using a Si tip with a total cone angle of 20° and a radius of curvature <10 nm. In order to examine the etched sidewalls the samples were tilted by 10°.

[a]Formally with the Fraunhofer-Institut für Angewandte Festkörperphysik, D-79108 Freiburg, Germany. Current address: Department of Electronics, Royal Institute of Technology, Electrum 229, S-164 40 Kista, Sweden; electronic mail: daleiden@ele.kth.se

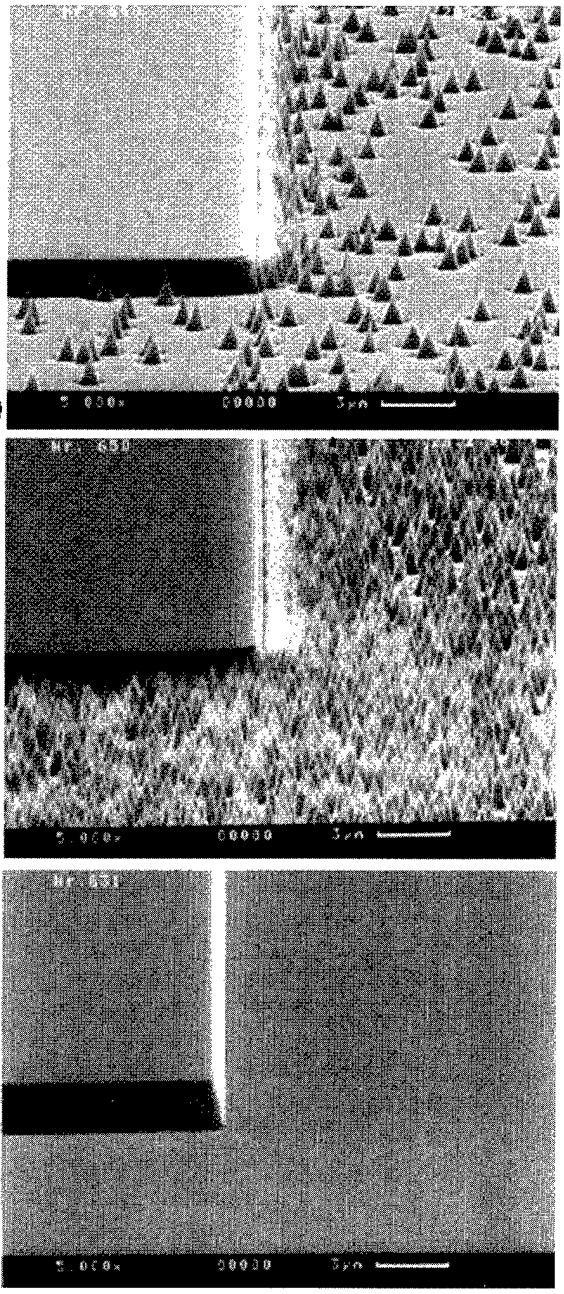


FIG. 1. Cl_2-CAIBE of InP, $U_B=400$ V, $p=3\times10^{-4}$ mbar, (a) $T_s=50$ °C, (b) $T_s=80$ °C, and (c) $T_s=-5$ °C.

III. RESULTS

Dry etching of InP-based materials layers with chlorine reactive gases generally lacks due to the low volatility of the $InCl_x$ reactive products. As an example Fig. 1(a) shows a secondary electron microscope (SEM) picture of an InP layer etched with Cl_2 at a chamber pressure of 3×10^{-4} mbar and a substrate temperature of 50 °C. Dark cones of InP arise due to a masking of the InP surface by $InCl_x$ clusters [proven by energy dispersive x-ray measurement (EDX)] during the etching.

For temperatures up to 120 °C more $InCl_x$ clusters occur and the number of InP cones strongly increases [Fig. 1(b)], $T_S=80$ °C. On the other hand, when reducing the temperature below 10 °C smooth and clean surfaces are obtained; best results for Cl_2-CAIBE are observed at low temperatures of −5 °C with a root mean square (rms) surface roughness of 0.5 nm measured by AFM on the bottom surface [Fig. 1(c)]. This can be explained as follows: By lowering the temperature the chemical etch component will be extenuated. Thus the physical component, which is kept constant, gets increasing influence. This leads to an even etching of indium and phosphorus producing a stoichiometric etched surface.

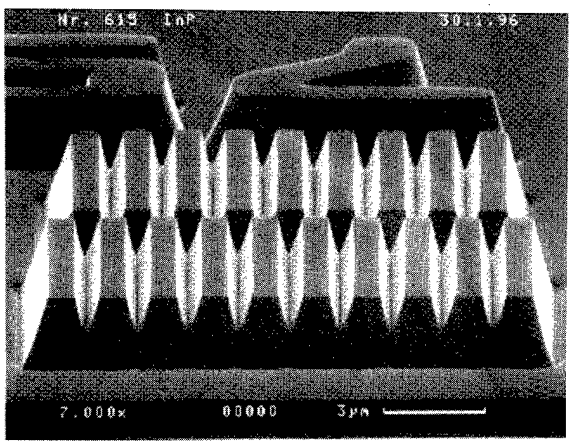


FIG. 2. BCl_3-CAIBE etched test structure in InP, $U_B=400$ V, $p=8.9\times10^{-4}$ mbar, $T_s=-5$ °C.

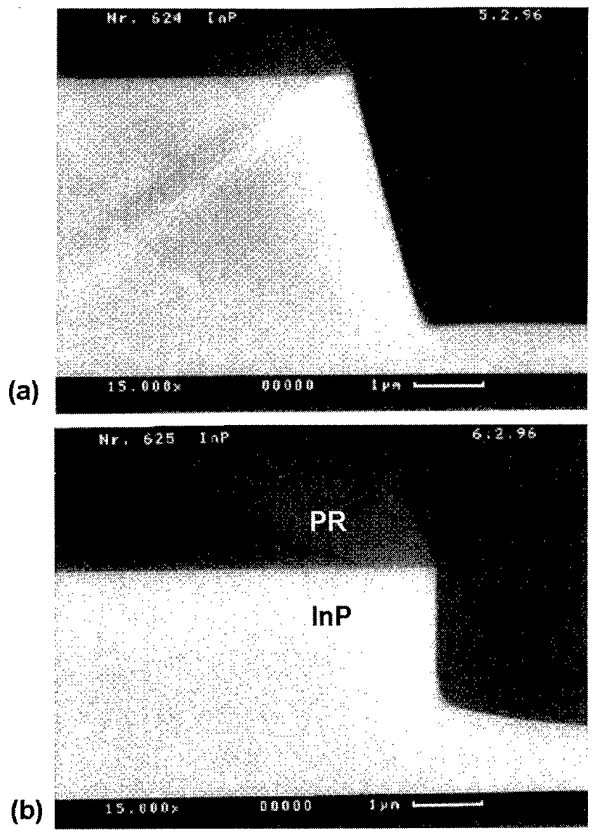


FIG. 3. Tuning of sidewall slope in InP by tilting the substrate relative to the impinging ion beam (a) tilt=0°, (b) tilt=25°, BCl_3-CAIBE, $U_B=400$ V, $p=8.9\times10^{-4}$ mbar, $T_s=-5$ °C.

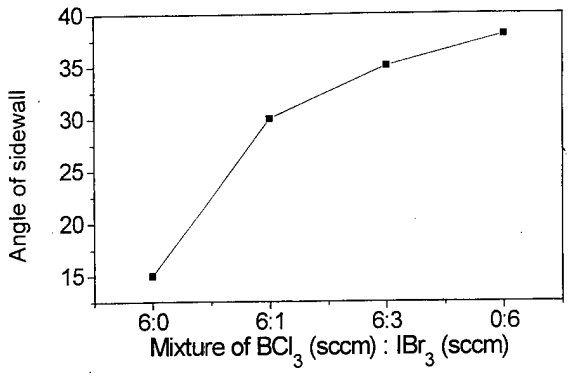


FIG. 4. CAIBE of InP:U_B=400 V, $T_{substr.}$=−5 C, dependence of sidewall angle on the mixture of BCl_3/IBr_3.

When using IBr_3 instead of Cl_2 no masking of the surface is obtained up to room temperature. Otherwise the surface roughness is rather large, 8 nm (bottom surface) in this case, even at low temperatures. Furthermore the etch rates are reduced to 40 nm/min compared to 70 nm/min when using Cl_2.

A newly developed CAIBE process with BCl_3 combines the advantages in all respects. Figure 2 shows as an example the SEM picture of an InP test structure etched at a substrate temperature of −5 °C and a chamber pressure of 8.9×10^{-4} mbar. The Ar ion beam was accelerated by 400 V. Clean and very smooth surfaces (rms roughness =0.2 nm on bottom surface, 3.6 nm on etched sidewalls) are observed; this surface quality is maintained up to room temperature. At low temperature the etch rate of the process is 75 nm/min. Similar results are obtained when etching complex (InGa)(AsP) and (AlInGa)(AsP) layers.

Low temperature CAIBE with BCl_3 yields rather steep sidewall slope of the etched structures, 15° with respect to the surface normal. By tilting the substrate relative to the impinging ion beam this slope can be tuned (Fig. 3). With a tilt of 25° the sidewall, turned towards the ion-beam is perpendicular, which is required for laser facets, the opposite sidewall has a tilt of 45°.

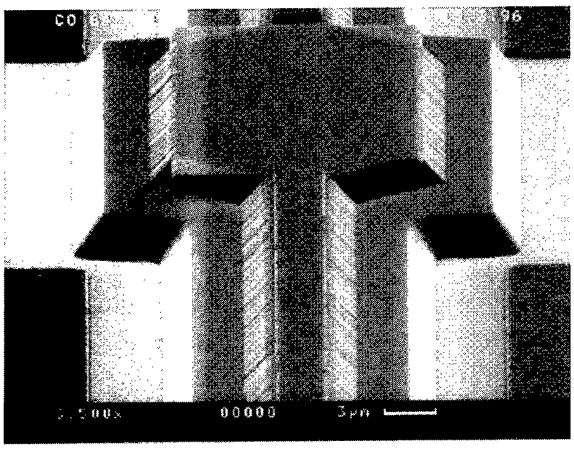


FIG. 5. BCl_3/IBr_3 (6:1) CAIBE etched waveguide mesa in (InGa)(AsP), U_B=400 V, $p=9\times10^{-4}$ mbar, T_s=−5 °C.

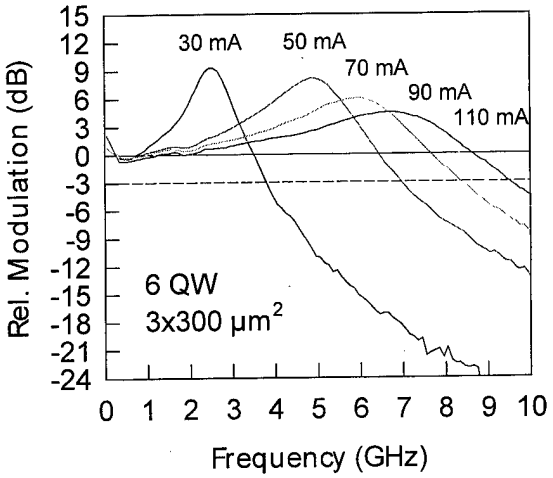


FIG. 6. Modulation response of 3×300 μm^2 6QW-(InGa)(AsP)/InP ridge waveguide laser at various bias currents (30–110 mA).

Another possibility to continuously increase the sidewall slope (symmetrical) is adding IBr_3 to the process gas. While pure IBr_3 yields a sidewall slope of 38° a 6:1 mixture of BCl_3 and IBr_3 gives a slope of 30° (Fig. 4). The addition of IBr_3 reduces the etch rate at a substrate temperature of −5 °C, the 6:1 mixture yields 55 nm/min.

Long wavelength (1.55 μm) (InGa)(AsP)/InP lasers were successfully fabricated by etching the waveguide mesa with low temperature CAIBE (Fig. 5). To obtain rather flat sidewall slopes of 30° the 6:1 mixture of IBr_3 and BCl_3 was applied.

The devices showed very low transparency current densities of 350 mA/cm^2 [number of quantum wells (QWs)=6]. Figure 6 presents small-signal direct modulation response curves of 3×300 μm^2 devices at various cw bias currents (30–110 mA). The devices achieve modulation bandwidths of 9.5 GHz at bias currents of 110 mA.

ACKNOWLEDGMENTS

This work was supported by the German Bundesministerium für Bildung und Forschung (BMBF) under Contract No. 01 BP 465/5. The authors would like to thank R. Dian for SEM characterization and W. Benz for the laser characterization.

[1] I. Adesida, K. Nummila, E. Andideh, and J. Hughes, J. Vac. Sci. Technol. B **8**, 1357 (1990).
[2] J. Singh, J. Vac. Sci. Technol. B **9**, 1911 (1991).
[3] J. W. McNabb, H. G. Craighead, H. Temkin, and R. A. Logan, J. Vac. Sci. Technol. B **9**, 3535 (1991).
[4] S. Dzioba, S. Jatar, T. V. Herak, J. P. D. Coook, J. Marks, T. Jones, and F. R. Shepherd, Appl. Phys. Lett. **62**, 2486 (1993).
[5] Y. Sugimoto, T. Yoshikawa, N. Takado, S. Kohmoto, N. Hamao, M. Ozaki, M. Sugimoto, and K. Asakawa, Nec. Res. Dev. **33**, 469 (1992).
[6] J. Daleiden, K. Eisele, R. Keller, G. Vollrath, F. Fiedler, and J. D. Ralston, Opt. Quantum Electron. **28**, 527 (1996).
[7] J. Daleiden, K. Eisele, R. E. Sah, K. H. Schmidt, and J. D. Ralston, J. Vac. Sci. Technol. B **13**, 2022 (1995).

Evaluation of trifluoroiodomethane as SiO_2 etchant for global warming reduction

F. Fracassi[a] and R. d'Agostino
CNR Centro di Studi per la Chimica dei Plasmi c/o Dipartimento di Chimica, Università di Bari, via Orabona, 4 70126 Bari, Italy

(Received 31 October 1997; accepted 22 May 1998)

The utilization of CF_3I in the plasma assisted dry etching of SiO_2 has been studied in order to reduce the environmental impact of microelectronics device fabrication. The results show that CF_3I is a promising substitute of CF_4 in oxide etching since its utilization reduces 3–3.5 times the contribution to the global warming, nevertheless it still has a consistent effect on the environment for the plasma assisted formation of perfluorocompounds. © *1998 American Vacuum Society.* [S0734-211X(98)10204-4]

I. INTRODUCTION

The dry etching of SiO_2 thin films is one of the key processes in microelectronics device fabrication. In order to obtain the proper anisotropy[1,2] and selectivity with respect to silicon and silicides, the process must be carried out in plasmas containing fluorocarbon radicals (CF_x). It is well known, in fact, that CF_x radicals play important roles in SiO_2 dry etching: they improve the oxide-to-silicon or nitride etch selectivity as well as induce anisotropy. The most popular way to generate these species is to feed glow discharges with perfluorocompounds (PFCs) such as CF_4, C_2F_6, or with CHF_3 (from the environmental point of view CHF_3 behaves as a PFC), then the most appropriate radical concentrations are achieved by selecting the proper working conditions and/or the additive-to-monomer feed ratio.

Unfortunately, most perfluorocompounds, in particular those utilized for oxide etching, are extremely long-lived species and can efficiently trap more heat than any other molecule contained in the atmosphere: i.e., they act as greenhouse gases and contribute to the warming of the earth surface (global warming). An approximate evaluation of the effect of a species on global warming can be obtained by means of the "global warming potential" on 100 years time horizon (CWP_{100})[3] which is defined as the cumulative direct effect on the atmosphere resulting from the instantaneous release of 1 kg of given gas into atmosphere relative to the release of 1 kg of CO_2 .

The GWP_{100} of PFCs utilized for oxide etching is very high, ranging from 6.000 to 12.500,[4] and therefore, even if the amount emitted in the atmosphere during microelectronics device fabrication is very low compared to other greenhouse gases, their effect of global warming is high.

Following to the international agreement signed in 1992 in Rio de Janeiro, efforts are in progress all over the world to find alternative to, or at least reduce, PFC emissions in the atmosphere. A milestone in this field is the Memorandum of Understanding between EPA and individual U.S. companies, a 5 year voluntary agreement, whose final version is dated March 1996.[5]

[a]Electronic mail: fracassi@area.area.ba.cnr.it

The above scenario has stimulated a research project at the Department of Chemistry of University of Bari, whose preliminary results are reported in this article. The ambitious aim of the overall research is to decrease the environmental impact of plasma processes by reducing the emissions in the atmosphere of PFCs and of other gases with high GWP. The present work is devoted to the dry etching of SiO_2 and has two goals: (i) characterization of the most utilized chemistry for SiO_2 dry etching (i.e., CF_4–CHF_3 fed plasma), and (ii) evaluation of other fluorinated compounds as potential substitutes to CF_4 , CHF_3 , or both. In fact a meaningful evaluation cannot simply be performed by replacing the reactants with others characterized by lower GWP_{100} , since it is necessary to quantify at the same time both the amount of unreacted feed emitted in the atmosphere as well as the PFCs produced in the glow discharge.

Among the various candidates to replace CF_4 and/or CHF_3 , we decided to start with CF_3I because, in addition to the short lifetime in atmosphere (2 days, approximately), which results in very low GWP_{100} ($\ll 1$)[6] and to the low ozone depletion potential, it was expected to behave very similarly to CHF_3 or CF_4 , at least for what SiO_2 etching is concerned.

The utilization of CF_3I and other iodofluorocarbons in SiO_2 dry etching, in completely different working conditions and apparatus, on the other hand, has been already reported in the literature.[7]

In order to simulate the typical situations of manufacturing processes and facilitate the transfer of achievements, the work has been performed with a home made etcher designed to be compatible with the operating conditions of a standard Lam Rainbow 4520 etcher, equipped with additional diagnostic techniques.

It will be shown that trifluoroiodomethane (CF_3I) is a promising candidate to replace the perfluorocompounds currently utilized for oxide processing. The etch rate of tetraethylorthosilicate (TEOS) plasma enhanced chemical vapor deposition (PECVD) deposited SiO_2 thin films in CF_3I containing plasma, as well as residue deposition, have been found to be comparable to that obtained with CF_4-based chemistries.

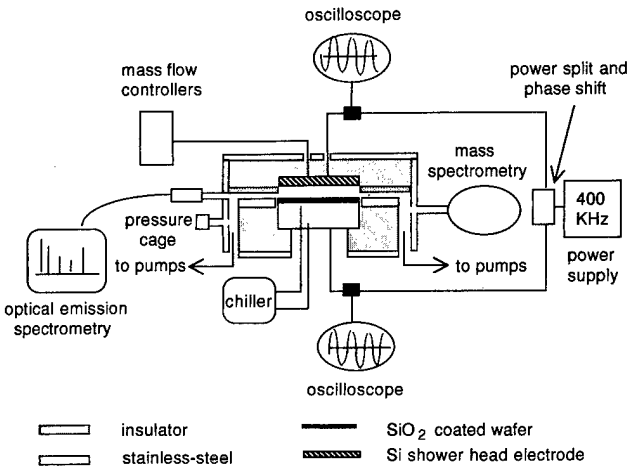


FIG. 1. Schematic of the experimental apparatus utilized.

II. EXPERIMENT

A schematic drawing of the experimental apparatus is shown in Fig. 1. The symmetrical parallel-plate reactor is equipped with two electrodes (6 in. diameter and 13 mm interelectrode space) both connected to the same 400 kHz power supply by a power split and phase shift device (Split Power Reverse Phase Technology); the electromagnetic waves on the two electrodes are shifted of 180° during operations. The lower electrode, on which a 6 in. wafer is positioned during plasma treatment, is made of stainless steel and is kept at 5 °C by a chiller.

The gas feed is controlled by means of MKS mass flow controllers and enters the vacuum system through a 6 in. silicon shower positioned on the upper electrode. The plasma confinement in the interelectrode volume (180 cc) is improved by a Vespel focusing ring. A capacitive MKS pressure gauge, a rotary backed turbomolecular pump, and a throttle valve are used to monitor and maintain the working pressure at 300 mTorr during the experiments.

The system is equipped with three diagnostics: high impedance electric probes, optical emission spectroscopy, and mass spectrometry. The electric probes, connected to a digital oscilloscope, allow control of the peak-to-peak voltage on both electrodes in order to detect variations of the electrical conditions and therefore of plasma characteristics, ion bombardment energy, in particular. It is well known, in fact, that the working frequency of 400 kHz is below the ion transit frequency[8] and therefore positive ion energy is mainly controlled by the instantaneous peak-to-peak voltage (and pressure, of course).

The optical emission of the plasma, sampled through a sapphire window by a quartz optical fiber, is analyzed with a monochromator (32 cm focal length) equipped with an intensified optical multichannel analyzer (OMA). The following spectroscopic features have been followed: CF (2022 Å),[9] CF_2 (2451, 2488, 2519 Å),[12] CF_2^+ (continuum 2400–2900 Å),[10] He (5873 Å), and Ar (7504 Å). A very intense peak, which is presumably due to iodine or iodine containing compounds, has been detected at 2050 Å in CF_3I plasma. The signal of F atoms at 7037 Å has never been detected in the experimental conditions explored. F atoms produced by fluorocarbon compound fragmentation, in fact, are expected to disappear rapidly by reacting with the upper silicon electrode (loading conditions). In order to compare the production of fluorocarbon radicals in different conditions, the relative concentration of CF and CF_2 radicals have been approximately evaluated by means of actinometry,[11,12] using Ar as actinometer. This approach needs further investigation since its validity in the experimental conditions utilized has not been checked up to now. Actinometric results can be affected by systematic errors for the different threshold energy of excitation between Ar (13 eV) and radicals (few eV),[13] when use of two different actinometers is not made. However, it has been previously shown that emission intensity ratio between CF_2^+ and Ar well represents CF_2 radical trends even when conditions are not suitable for CF_2 emissions.[12]

Mass spectrometric analyses of the plasma has been accomplished with a quadrupole mass spectrometer (QMG 511 Balzer). The gas feed is sampled at the exit of the plasma zone through an independently pumped stainless steel line (1 m long, 0.6 cm internal diameter) and enters the ionization chamber of the quadrupole through a leak valve. This arrangement allows the detection of only the stable species formed after the plasma activation since radicals, ions, and other excited species are expected to recombine or to be quenched for gas-phase or gas-surface collisions in the sampling line.

In order to get meaningful results with mass spectrometry, we adopted the following procedure: (i) to reduce species fragmentation and recombinations in the ion source, we have operated the spectrometer at low values of ionization energy and pressure: 40 eV and 5×10^{-6} Torr, respectively; (ii) in order to compensate instrumental fluctuations on the mass spectrometer and to compare the results obtained with different gas feeds, the data have been normalized to the signal of Ar ($m/z=40$), added to the gas feed at constant amount (internal standard).

Some x-ray photoelectron spectroscopy (XPS) surface analyses of treated substrates have been carried out by means of an XPS PHI-5300 apparatus in order to detect the surface contamination of etched wafers.

The etching experiments have been conducted at 300 mTorr, 400 W, 160 sccm as total flow rate (Ar 110 sccm, He 10 sccm, fluorinated compounds 40 sccm), the peak-to-peak voltage in these conditions was 780–800 V on both electrodes and the phase shift always fixed at 180°.

Etch rates have been measured *ex situ* by thickness evaluation carried by a Tencor *a*-step profilometer on 6 in. TEOS SiO_2 PECVD wafer (exposed area 30%), after 3 min of plasma treatment at the pedestal temperature of 5 °C. In order to study the interactions of CF_3I with silicon and photoresist, some experiments have also been performed on bare Si wafers and on wafers covered with positive photoresist.

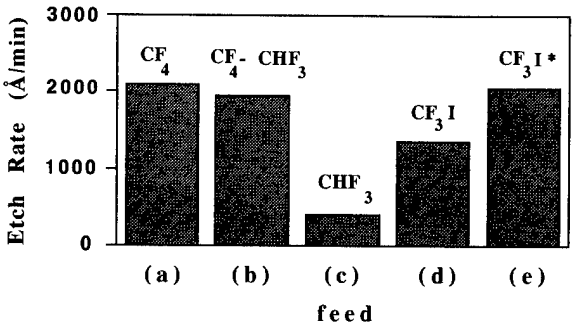


FIG. 2. SiO_2 etch rate with different feeds: (a) CF_4, 40 sccm; (b) CF_4, 15 sccm–CHF_3, 25 sccm; (c) CHF_3, 40 sccm; (d) CF_3I 40 sccm; (e) CF_3I 50 sccm. In all conditions the feed contains also Ar (110 sccm) and He (10 sccm).

III. RESULTS

A. Etch rate evaluation

The etch rates of SiO_2 in CF_4, CF_4–CHF_3, CHF_3, and CF_3I containing plasmas under the same working conditions are compared in the histogram of Fig. 2. This is a good starting point to evaluate the possibility of CF_3I utilization for oxide etching, since one of the principal requirements of a new reactant is to have etch rates similar to PFCs. As expected CF_4 leads to the highest etch rate and CHF_3 to the lowest. It is well known, in fact, that CHF_3 is added to the gas feed to improve the SiO_2/Si selectivity as well as process anisotropy by means of sidewall passivating (or sidewall inhibitor) mechanism. With CF_3I (d) a somewhat lower etch rate is obtained than with (a) CF_4 or (b) CF_4–CHF_3, nevertheless a simple increase of CF_3I flow rate to 50 sccm (e) is sufficient to increase the etch rate of CF_3I plasma almost to the value detected with 100% CF_4. The data of Fig. 2 indicate that the oxide etch rate with CF_3I can be close to those obtained with conventional chemistry and that it behaves more similarly to CF_4 than to CHF_3.

The analogy between CF_3I and CF_4 is also reflected by the similar SiO_2/Si (4.5/1) and SiO_2/photoresist (5/1) etch selectivities. Of course, it is necessary to stress at this point that the similar behaviors detected for CF_3I and CF_4 in oxide etching could be in part due to the particular experimental apparatus utilized for the study, which has been designed to improve the performances of oxide etching processes. The chemistry of the plasma phase, in fact, is strongly affected by the reaction of fluorine atoms with the silicon shower positioned on the upper electrode. As a consequence, the gas phase concentration ratio between fluorocarbon radicals-to-fluorine atoms increases and the plasma becomes more suitable to oxide etching.

Table I shows that, within the experimental error of XPS, there are not significant differences on the surface composition of SiO_2 film treated in CF_4 and CF_3I plasmas, apart, of course, the presence of some iodine (0.2%) in the latter case. Iodine content was too small to perform meaningful data treatments (with fitting procedures and chemical shift evaluation) to determine whether it is bonded to silicon, carbon or trapped in voids or in the damaged surface layers of etched SiO_2. Some differences have however been detected on the chemical composition of plasma treated silicon surfaces (Table I). With CF_3I the intensity of Si signal is very low (its surface concentration is only 2%) and the amount of carbon is significantly higher than with CF_4. This means that a thicker fluorocarbon deposit is formed on bare Si with CF_3I. The thickness of this layer can be roughly estimated to be less than 40–50 Å if the take-off angle and the escape depth of emitted photoelectrons are considered. Iodine represents a small fraction of the etched surface (only 1%) and it is probably bonded to the carbonaceous layer as it happens on the photoresist exposed to CF_3I plasma.

TABLE I. XSP analyses of samples treated in plasmas feed with different fluorine containing compounds.

Feed	Substrate	C	F	Si	O	I
CF_4	SiO_2	13	8	23	56	···
CF_3I	SiO_2	12	6	24	58	0.2
CF_4	Si	49	31	14	6	···
CF_3I	Si	68	27	2	2	1
CF_3I	Resist	52	42	···	2	4

The similarity between CF_4 and CF_3I can be partially explained with actinometric optical emission spectroscopy investigation. Preliminary, it is necessary to report here that the experimental conditions utilized in this study (low frequency plasma) are characterized by low electron temperature and/or electron density. The region of the electron energy distribution function where the threshold energy of excitation of most species falls is sparsely populated and consequently the efficiency of the excitation process is low. For this reason, the signals of He (whose excitation threshold energy is 23 eV) has never been detected even if its concentration in the gas feed was increased up to 20%.

Actinometry allows to follow the concentration trends of CF and CF_2 radicals which are extremely important species in oxide etching, although on a semiquantitative basis. They are, in fact, reported to be either etchants of SiO_2 either to act as building blocks for fluoropolymer deposition[14] and therefore, along with positive ion bombardment and substrate temperature, to participate to the well known etching-deposition competition and to control process selectivity and anisotropy. These are well known and debated topics whose discussion is beyond the scope of this presentation, nevertheless it is worth mentioning that when the effect of fluorine atoms can be ruled out, as in this case (the optical emissions of fluorine atoms have never been detected), fluorocarbon radical concentration ($[CF_x]$) can trigger the plasma from etching to deposition and *vice versa*. $[CF_x]$ must be optimized to obtain the right SiO_2 etch rate and anisotropy as well as SiO_2/Si selectivity: low $[CF_x]$ results in poor anisotropy and selectivity, while high $[CF_x]$ inhibits the etching process and favors polymer deposition.

The actinometric concentration of CF and CF_2 radicals are compared as a function of feed composition in Fig. 3 (unfortunately the very intense band at 2050 Å in CF_3I plasma

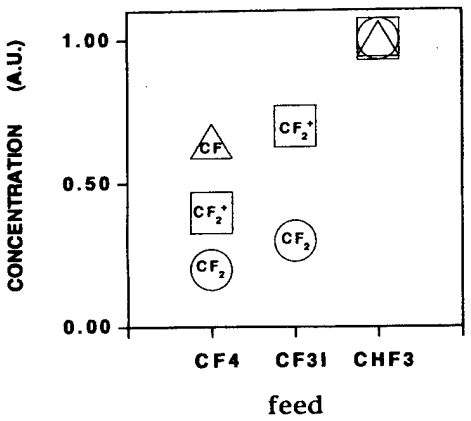


FIG. 3. Actinometric relative concentration of radicals and $I_{CF_2^+}/I_{Ar}$ vs plasma composition.

interferes with CF intensity measurements). Although the actinometric trends of Fig. 3 are semiquantitative, it can reasonably be accepted that the radical density follows the hierarchical order $CHF_3 > CF_3I > CF_4$. This is strengthened by the trend of CF_2^+-to-Ar emissions, which has been assumed to be representative of $[CF_2]$ trend in gas phase.

The similar behavior of CF_4 and CF_3I plasma with respect to CHF_3 in SiO_2 etching, at least in terms of etch rates, is most reasonably due to similar production of CF_2 (and probably CF) radicals, which is high enough to support SiO_2 etching according to reaction (1):

$$CF_x + SiO_2 = SiF_x + CO_2 \quad (1)$$

but not sufficient to have appreciable polymer deposition on silicon dioxide [reaction (2)]

$$(CF_x)_n + CF_x = (CF_x)_{n+1}. \quad (2)$$

With CHF_3 the importance of reaction (2) increases for the higher radical concentration and therefore a lower oxide etch rate is detected.

Optical emission spectroscopy can also be utilized to justify the different surface compositions detected on plasma-treated Si substrates (Table I). The low Si concentration of the surface exposed to CF_3I plasma, which has been ascribed to a 40–50 Å thick fluorocarbon layer, is compatible with the lightly higher radical concentration (and therefore improved polymerizing capability) of CF_3I with respect to CF_4 plasma. This is evident only on bare silicon (not on SiO_2) because when CF_x radicals impinge on the surface they cannot react with the silicon (as it happens with SiO_2), they can only undergo polymer formation [reaction (2)] or can be desorbed back in the plasma.

These experimental results allow to consider CF_3I as a promising alternative to PFCs in oxide etching, its properties appearing intermediate between those of CF_4 and CHF_3. Of course, the optimization of operating parameters, also in terms of feed additive concentration, is needed in order to achieve the right process characteristics in terms of selectivity and anisotropy.

B. Environmental impact evaluation

The possibility of replacing PFCs with CF_3I has been evaluated not only with etch rate measurement, but also as a function of environmental impact reduction in terms of GWP. As already mentioned in the Introduction, this study cannot be conducted by simply determining the extent of unreacted feed compounds which are emitted in the atmosphere, since the highly reactive species formed in the plasma phase recombine and lead to by-products with very high GWP, such as CF_4, C_2F_6 and others. This issue, sometime underestimated, is particularly important when dealing with reactive species as CF_3I. In this case, the high reactivity has a contrastable effect: on one side results in low GWP_{100}, and makes the effect of unreacted feed on the environment negligible, on the other side it can turn in a disadvantage for the plasma production of high GWP species. Mass-spectrometric investigation of CF_3I, CF_4 and CHF_3 fed plasmas have been performed to evaluate the quantity of high GWP compounds emitted in the atmosphere during the etching process.

The most important signals detected in the region 50–300m/e, along with their possible parent species, are reported in Table II. Of course, iodine containing fragments are present only with CF_3I. Other fragments characterized by very low intensity, between 200 and 300m/e, have been neglected. Analyses performed with the substrate electrode covered with Si (exposed area 100%), SiO_2 (exposed area 30%) or patterned SiO_2 wafer (exposed area 3%) gave the

TABLE II. Main fragments detected by mass spectrometry of rf plasmas fed with different fluorine containing compounds and their most probable parent species.

m/e	Fragment	Parent	m/e	Fragment	Parent
50	CF_2^+	CF_4, CHF_3, CF_3I	169	$C_3F_7^+$	C_3F_8
51	CHF_2^+	CHF_3	177	CF_2I^+	CF_3I
69	CF_3^+	CF_4, CHF_3, CF_3I	196	CF_3I^+	CF_3I
81	$C_2F_3^+$	C_2F_4	208	$C_2F_3I^+$	a
85	SiF_3^+	SiF_4	220	$C_3F_3I^+$	a
100	$C_2F_4^+$	C_2F_4	227	$C_2F_4I^+$	C_2F_5I
104	SiF_4^+	SiF_4	246	$C_2F_5I^+$	C_2F_5I
119	$C_2F_5^+$	C_2F_6	254	I_2^+	I_2
127	I^+	I, I_2	259	$C_3F_5I^+$	a

[a]Data do not allow any confident assignment.

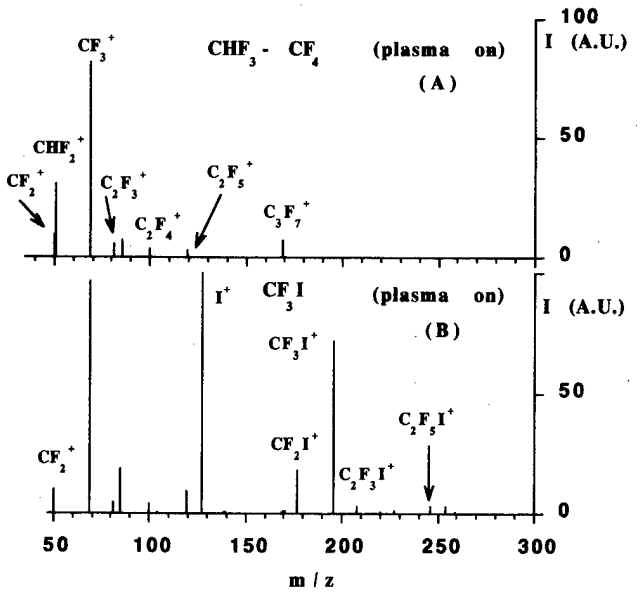


FIG. 4. Mass spectra of plasmas fed with: (a) CHF_3, 25 sccm; CF_4, 15 sccm; Ar, 110 sccm; He, 10 sccm; (b) CF_3I 40 sccm; Ar, 110 sccm; He, 10 sccm.

same results. Data of Table II indicate that in addition to the fragments of SiF_4, namely, SiF_4^+, formed for the etching of the Si upper electrode and to I^+, I_2^+ (in case of CF_3I), the main oligomers detected are C_2F_4 (GWP_{100} not known), C_2F_6 ($GWP_{100}=12.500$) and C_3F_8 ($GWP_{100}=6950$). For CF_3I fed plasma, in addition to the above species, CF_4 ($GWP_{100}=6.300$) and I-containing oligomers have also have detected (i.e., C_2F_3I and C_2F_5I). All these fragments appear in the mass spectra of CF_4-CHF_3 and CF_3I fed plasmas reported in Fig. 4.

Quantitative analysis of CF_3I, CHF_3, CF_4 and of oligomers have been performed after calibration with pure compounds. In particular the following fragments have been utilized for calibration: CF_3I^+ ($m/e=196$) for CF_3I, CHF_2^+ ($m/e=51$) for CHF_3, $C_2F_4^+$ ($m/e=100$) for C_2F_4, $C_2F_5^+$ ($m/e=119$) for C_2F_6, $C_3F_7^+$ ($m/e=169$) for C_3F_8. Quantitative analysis of CF_4 has been performed utilizing the ion CF_2^+ at $m/e=50$ after subtracting the contributions of the fragmentation of other species to this ion.

The results are reported in Table III. The total amount of iodine containing oligomers (missing in the table) is not expected to be higher than 2–4 sccm or 5–10% of the total CF_3I utilized (a simple mass balance confirms this prediction).

TABLE III. Mass spectrometric analysis of plasmas feed with different fluorine containing compounds: evaluation of PFC emissions in sccm. Figures in brackets are GWP_{100} values for pure compounds. CF_4 (40 sccm); CHF_3 (40 sccm); CF_3I (40 sccm); CF_4 (15 sccm)–CHF_3 (25 sccm). Ar (110 sccm) and He (10 sccm) are always contained in the feed.

Feed	CF_4 (6.300)	CHF_3 (12.100)	CF_3I (1)	C_2F_6 (12.500)	C_3F_8 (6.950)	C_2F_4
CF_4	29	0.6	···	2.3	0.6	0.6
CHF_3	4.5	23	···	0.9	0.2	1.6
CF_3I	1.2	···	26	3.4	0.3	1.9
CF_4-CHF_3	11	15	···	1.1	2.9	1.8

TABLE IV. Evaluation of the environmental impact of various feed with respect to CF_3I. T represent the decrease of the environmental impact utilizing CF_3I instead of CF_4, CHF_3, and CF_4-CHF_3.

Feed	T
CF_4	3.1
CHF_3	3.4
CF_4-CHF_3	3.5

Even if CF_3I considerably reduces the effect on the environment of the gas emitted in the atmosphere during oxide etching, it is far to be completely inert with respect to greenhouse effect, as its low GWP_{100} and short lifetime would indicate. By-products of CF_3I fed plasmas are CF_4 and C_2F_6 which are characterized by high GWP_{100}, the latter, in particular. The iodine containing oligomers, which have not been quantified, are not expected to contribute consistently to global warming for their short lifetime.[7]

The decrease of the total global warming effect of CF_3I with respect to a given feed can be roughly evaluated with the factor "T," calculated according to Eq. (3):

$$T=\frac{(\Sigma\Phi_i M_i \mathrm{GWP}_{100i})_x}{(\Sigma\Phi_i M_i \mathrm{GWP}_{100i})_{CF_3I}}. \tag{3}$$

Φ_i is the flow rate of i species detected by mass spectrometry (Table III) in sccm, $(GWP_{100})_i$ their global warming potential and M_i the molecular weight, x is the particular feed in consideration (e.g., CF_4, CHF_3 or CF_4-CHF_3). The sum is extended to all compounds with known GWP_{100}. In Eq. (3) Φ_i must be multiplied for M_i (expressed in g/mole) because GWP refers to unit of mass while the flow rate is proportional to moles. The results of T calculation are reported in Table IV.

CF_3I utilization reduces the effect of oxide etching on global warming by 3–3.5 times with respect to conventional feed, but nevertheless its effect is still consistent due to the high GWP compounds formed after plasma activation; C_2F_6 contribution, in particular, is approximately 80% of the total. From Table III it appears that the effect on global warming of the compounds formed in the plasma is important also with conventional feed (about 20% with CF_4 and CF_4-CHF_3).

IV. CONCLUSION

The experimental results presented in this article indicate that CF_3I is a promising alternative to conventional PFCs also from the environmental point of view since it allows to reduce the contribution of oxide etching to global warming. Etch rate determinations and XPS surface analyses show that trifluoroiodomethane is a potential substitute of CF_4 in oxide processing.

Mass spectrometric analyses have shown that a consistent amount of C_2F_6 forms in CF_3I fed plasma. C_2F_6 increases the

effect on global warming and therefore its production must be minimized to further improve the environmental impact of trifluoroiodomethane utilization. C_2F_6 generation is probably the main disadvantage connected to CF_3I utilization, while with conventional PFCs also the improvement of feed utilization must be achieved to reduce the contribute to global warming. This means that the contribution of the compounds formed by plasma assisted reactions cannot be neglected when the global warming effect is evaluated. These compounds, in fact, have been found to contribute to the total GWP for about 20%.

This contribution is strictly valid only for the particular apparatus utilized since any data with other experimental set up has been obtained up to now, nevertheless these authors are quite confident that some generalizations can be done if the working conditions are properly selected (i.e., feed composition, input power, pressure, etc.).

ACKNOWLEDGMENTS

The authors express their deep gratitude to Dr. Marco Stendardo and Mr. Vincenzo Colaprico for the numerous stimulating discussions and the technical assistance. The research has been partially supported by Lam Research-Freemont (CA) and SGS Thomson-Agrate (Italy).

[1]F. Fracassi and J. W. Coburn, J. Appl. Phys. **63**, 1758 (1988).
[2]D. L. Flamm, in *Plasma Etching, an Introduction,* edited by D. M. Mannos and D. L. Flamm (Academic, New York, 1989).
[3]J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, *Climate Change. The IPCC Scientific Assessment, 1991* (Cambridge University Press, Cambridge, 1991).
[4]S. Karecki, B. Tao, and L. R. Reif, Proceedings to SEMI-SEMICON Southwest 96, "A Partnership for PFC Emission Reduction," Austin, Texas, October 14, 1996, pp. 100–112.
[5]U. S. Environmental Protection Agency, Memorandum of Understanding, March 1996.
[6]J. T. Houghton, L. G. Meira Filho, J. Bruce, H. Lee, B. A. Callander, E. Haites, N. Harris, and K. Maskell, *Climate Change 1994, Radiative Forcing of Climate Change and An Evaluation of the IPCC IS92 Emission Scenarios* (Cambridge University Press, Cambridge, 1995).
[7]S. Karecki, L. C. Pruette, and L. R. Reif, Proceedings to MRS Fall Meeting 1996-Symposium.
[8]R. H. Bruce, J. Appl. Phys. **52**, 7064 (1981).
[9]R. d'Agostino, F. Cramarossa, S. De Benedictis, and G. Ferraro, J. Appl. Phys. **52**, 1259 (1981).
[10]R. d'Agostino, F. Cramarossa, and S. De Benedictis, Plasma Chem. Plasma Process. **2**, 213 (1982).
[11]J. W. Coburn and M. Chen, J. Appl. Phys. **51**, 3134 (1980).
[12]R. d'Agostino, F. Cramarossa, S. De Benedictis, and G. Ferraro, J. Appl. Phys. **52**, 1259 (1981).
[13]R. d'Agostino, F. Cramarossa, S. De Benedictis, and G. Ferraro, Plasma Chem. Plasma Process. **4**, 163 (1984) and references therein.
[14]R. d'Agostino, F. Cramarossa, F. Fracassi, and F. Illuzzi, in *Plasma Deposition, Treatment and Etching of Polymers,* edited by R. d'Agostino (Academic, New York, 1990).

Electrical characterization and annealing behavior of defect introduced in Si during sputter etching in an Ar plasma

P. N. K. Deenapanray,[a] F. D. Auret, and G. Myburg
Department of Physics, University of Pretoria, Pretoria 0002, South Africa

(Received 30 September 1997; accepted 15 May 1998)

We have employed current–voltage and capacitance–voltage measurements in conjunction with deep level transient spectroscopy to characterize the defects induced in n-Si during rf sputter etching in an Ar plasma. The reverse current, at a bias of 1 V, of the Schottky barrier diodes fabricated on the etched samples decreased nonmonotonically with etch time to a minimum at 6 min and, thereafter, increased. The reverse current also increased with decreasing plasma pressure. The barrier heights of the diodes followed the opposite trend. Six prominent electron traps were introduced in the substrate during Ar sputter etching. A comparison with the defects induced during high-energy alpha-particle and electron irradiation of the same material revealed that sputter etching created the VO and VP centers and $V_2^{-/0}$. The $V_2^{=/-}$ charge state of the divacancy was not detected in our plasma etched samples. We have attributed the nondetection of $V_2^{=/-}$ to the presence of defect-induced stress fields in the etched samples. A secondary defect with an energy level at $E_c - 0.219$ eV was introduced during annealing and was found to be stable at 650 °C. This defect was introduced at the expense of a sputter-etching induced defect P4, which has similar electronic and annealing properties as EAr201 ($E_c - 0.201$ eV), created in Ar-ion bombarded n-type Si. © *1998 American Vacuum Society.* [S0734-211X(98)07504-0]

I. INTRODUCTION

Plasma processes are versatile techniques which are routinely used for submicron scale device fabrication. Sputter etching, ion beam etching (IBE), and reactive ion etching (RIE) are used for anisotropic etching and the transfer of patterns to semiconductor surfaces. These plasma etching techniques also result in lattice damage at and below the surface which alters the electrical, optical and structural properties of the semiconductor.[1–3]

Traditionally, defects introduced during growth and subsequent processing have been considered undesirable. However, defect engineering[4] which relies on the controlled and reproducible introduction of defects, has been applied to control the lifetime of minority carriers,[5] tailor the barrier height of Schottky barrier diodes (SBDs),[1] and control the band offset in heterojunction devices.[6] Any successful application of defect engineering requires that the introduction rates and concentrations of processing-induced defects, together with their physical nature and electronic properties, are well characterized.

During sputter etching, the extent of damage depends on parameters such as etch time, temperature and rate, etching mode and bias conditions, as well as gas pressure and species. Extensive studies have been conducted to investigate the extent and type (donor or acceptor) of defects created during plasma etching and their effects on the barrier height modification of SBDs fabricated on the etched surfaces.[7–10] The sputter yield and the structural and bonding changes induced in Si by Ar-ion etching have also been studied.[11–13] Despite these studies, little information is available on the electronic properties, structure and thermal stability of defects introduced during sputter etching of Si in an Ar plasma. In this article, the electronic properties and annealing behavior of sputter-etching-induced (SEI) defects are reported. To better understand the nature of these defects, we have compared their electronic properties to those of defects created during high-energy (MeV) alpha-particle, proton and electron irradiation, and low-energy (1 keV) He- and Ar-ion bombardment of the same material. We have also investigated the extent of SEI damage by monitoring the rectifying behavior of the SBDs fabricated on the etched surfaces.

II. EXPERIMENTAL PROCEDURE

(111) oriented n-type Si layers of thickness 4 μm and doped to 4.8×10^{15} P/cm^3 were epitaxially grown on n^+ substrates by chemical vapor deposition (CVD). The samples were chemically cleaned before being sputter etched in a Leybold-Heraeus Universal Sputtering machine using an rf-excited (13.56 MHz) plasma. One batch of samples were sputter etched for 2–10 min with the plasma pressure kept constant at 9×10^{-3} mbar. A second batch of samples was etched at plasma pressures of 2×10^{-3}–2×10^{-2} mbar for 6 min. For both experiments, the dc bias on the rf electrode was fixed at 400 V. The samples were kept at room temperature (295 K) by mounting them on a water cooled substrate holder.

Circular Pd Schottky contacts of 0.77 mm diameter and 100 nm thickness were resistively deposited on the etched samples through a metal contact mask. As a control, Pd contacts were also deposited on chemically cleaned but unetched

[a]Present address: Department of Electronic Materials Engineering, Research School of Physical Sciences and Engineering, Australian National University, Canberra, ACT, 0200, Australia; electronic mail: pnk109@rsphysse.anu.edu.au

Si. Ohmic contacts were formed on the n^+ substrates using a liquidus In–Ga alloy.

Current–voltage ($I-V$) and capacitance–voltage ($C-V$) measurements were used to monitor the quality of the SBDs. The $C-V$ barrier height (ϕ_b^{CV}) was calculated from a graph of $1/C^2$ vs V_R by using a reverse bias, V_R, of between 0 and 1 V. The SEI defects were characterized using deep level transient spectroscopy (DLTS).[14] The energy levels, E_T, and apparent capture cross sections (i.e., thermal dependence of cross sections were not determined), σ_a, of the defects were determined from DLTS Arrhenius plots of $\log(e/T^2)$ vs $1/T$, where e is the emission rate at a temperature T. The defect depth profiles were determined using the method of Zohta and Watanabe.[15] Isochronal annealing experiments were conducted to 650 °C, at 50 °C intervals, for periods of 20 min under Ar flow and zero bias. DLTS was employed to monitor changes in the peak signal intensities of the prominent SEI defects as well as the introduction of secondary defects during annealing.

III. RESULTS

A. $I-V$ and $C-V$ characteristics of SBDs

Figure 1(a) shows the $I-V$ characteristics, measured at 250 K, of the SBDs deposited on the etched Si samples as a function of etch time. The reverse current of the control diode at 250 K (not shown) was of the order of 10^{-9} A for a bias of -1 V. The high ideality factors ($n>1.1$) of the diodes suggested that the current transport over the barrier was not dominated by thermionic emission,[16] so that the measured $I-V$ barrier heights (ϕ_b^{IV}) were not meaningful. It can be seen from Fig. 1(a) that the reverse current decreased nonmonotonically with etch time to reach a minimum at 6 min, and increased thereafter. Plots of $\ln(J/T^2)$ vs $1000/T$ at a forward voltage of 0.1 V are depicted in Fig. 1(b) for diodes fabricated on samples etched for 2, 6, and 10 min. It can be seen from this graph that the temperature dependence of the forward current of the diodes shows two activation energies, suggesting that two different mechanisms are responsible for current transport over the Schottky barrier in the temperature range examined (120–300 K). The low temperature mechanism is dominant for a 2 min etch time whereas the high temperature mechanism is dominant for a 6 min etch time. The low values of activation energies [less than $(E_g - V_F)/2$] suggest that the transport mechanisms involve a combination of generation recombination and tunneling.[17]

The above results are further substantiated by observing the change in barrier heights of the diodes as a function of etch time (Fig. 2). The barrier heights were extracted from $C-V$ measurements and compared with values calculated from Eqs. (1) and (2) using diode currents at a reverse bias of 1 V:

$$J_0=\frac{J}{\exp(qV/kT)-1}, \tag{1}$$

$$\phi_b^{\text{eff}}=-\frac{kT}{q}\ln(J_0/A^{**}T^2). \tag{2}$$

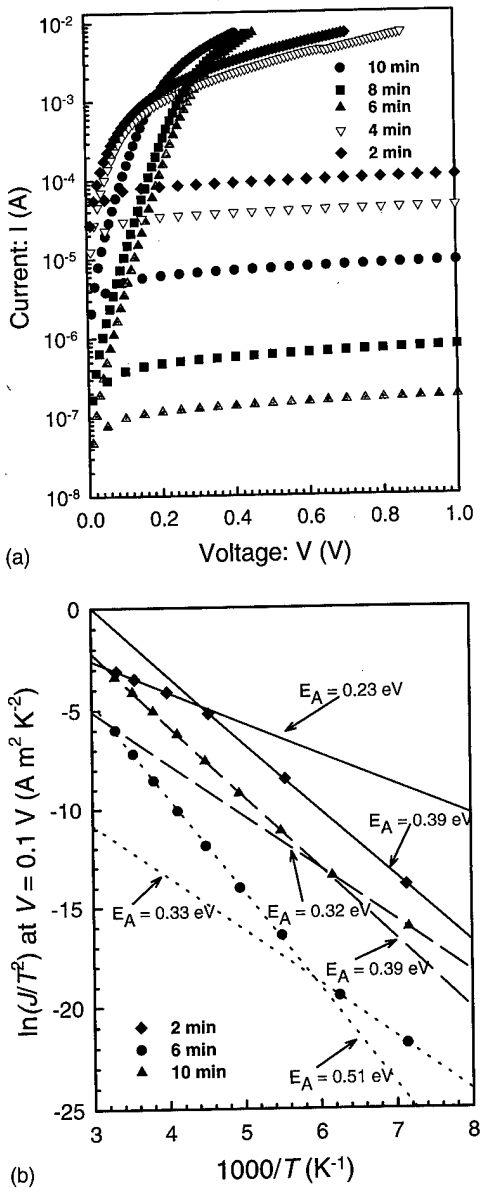


FIG. 1. Etch time dependence of: (a) the $I-V$ characteristics measured at 250 K, and (b) plots of $\ln(J/T^2)$ against $1000/T$ at a forward voltage of 0.1 V and measured between 120 and 300 K, for SBDs fabricated on plasma-etched n-Si at a pressure of 9×10^{-3} mbar and dc bias of 400 V.

In the above equations, J is the current density, V is the reverse bias, A^{**} is the effective Richardson constant (112×10^4 A m^{-2} K^{-2}—assumed to be constant between 120 and 300 K), T is the measurement temperature, ϕ_b^{eff} is the effective barrier height, and the remaining symbols have their usual meaning. As anticipated, the barrier heights follow the opposite trend to the reverse current except for the relatively higher ϕ_b^{CV} of the diode fabricated on the 2 min sputter-etched sample. $C-V$ depth profiling (not shown) demonstrated that the free carrier compensation at depths <0.5 μm below the interface after a 2 min etch was approximately 10%, and hence ϕ_b^{CV} of the corresponding diode is higher than expected. In general, the differences between ϕ_b^{CV} and ϕ_b^{eff} can be attributed to: (a) the image-force lowering of barrier height was not taken into account in calculating ϕ_b^{eff},

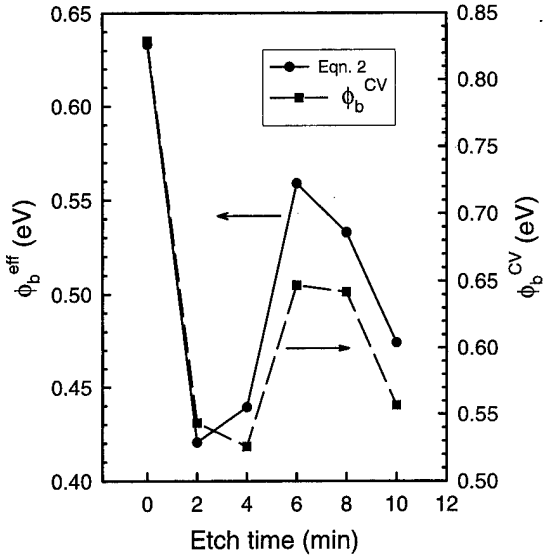


FIG. 2. Variation of $C-V$ barrier heights [ϕ_b^{CV}] and barrier heights calculated from Eqs. (1) and (2) [ϕ_b^{eff}] as a function of etch time. The measurements were made using Schottky barrier diodes fabricated on samples etched for 2, 4, 6, 8, and 10 min at room temperature.

and (b) the measured ϕ_b^{CV} values could be overestimated if the traps introduced during plasma-etching ionize under $C-V$ bias excursions.[18] The significant lowering in ϕ_b^{CV} of the diodes fabricated on the etched samples compared to that of the control diode further confirmed that the rectifying properties of Pd/n-Si SBDs fabricated on sputter-etched surfaces were poorer than those on unetched surfaces.

The variations in ϕ_b^{IV} and current, I_R, at a reverse bias of 1 V are plotted as a function of plasma pressure in Fig. 3(a), while Fig. 3(b) shows the pressure dependence of free carrier concentration of the etched samples. The dashed lines in Fig. 3(a) represent the barrier height (solid circles) and reverse current (solid squares) of the control (unetched) Schottky diode. The lowering in barrier heights with decreasing plasma pressure observed in Fig. 3(a) is in agreement with the results obtained during sputter deposition of metals on n-Si[1,19] and with the proposal that ion-induced etching introduces donor-type levels at and below the semiconductor surface.[8] On the other hand, the reverse current is found to decrease with increasing Ar pressure. The two following pertinent remarks can be made concerning the $C-V$ depth profiles shown in Fig. 3(b). First, the level of free carrier compensation increased with decreasing plasma pressure. This free carrier compensation could be caused either by the introduction of acceptor-type defects or be due to dopant passivation in the form of dopant-defect pairs. Second, these sputter-etching induced defects migrated to depths exceeding the projected range of ~20 Å of 400 eV Ar ions in Si (TRIM).

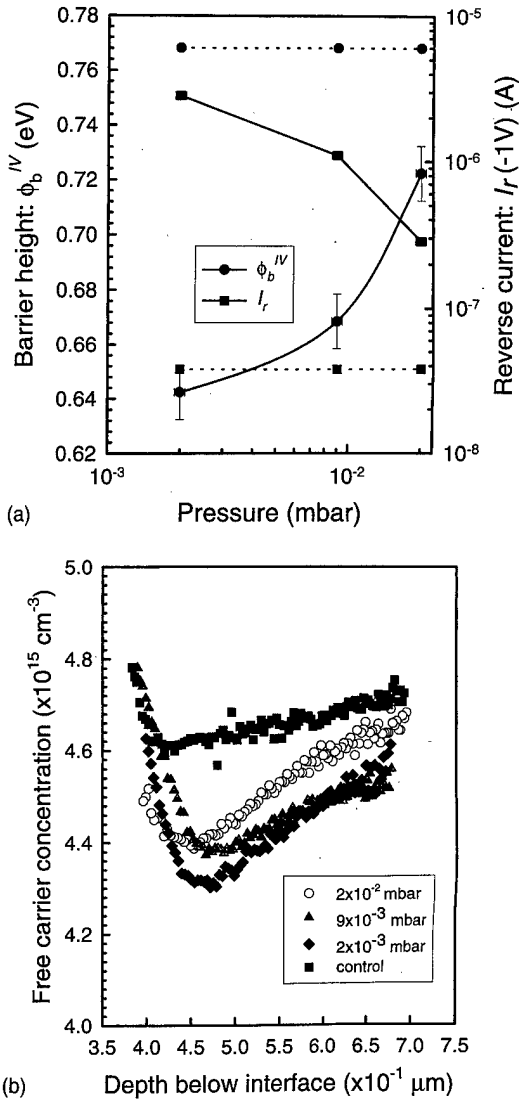


FIG. 3. Plots of (a) $I-V$ barrier heights (ϕ_b^{IV}) and reverse current at a bias of -1 V (I_R), and (b) free carrier concentration of sputter etched samples as a function of plasma pressure (2×10^{-3}, 9×10^{-3}, and 2×10^{-2} mbar). The dashed lines in (a) correspond to the barrier height (solid circles) and reverse current (solid squares) of the control sample. The carrier compensation observed in the control sample in (b) could be a result of hydrogen passivation of dopant atoms during the chemical cleaning step prior to metallization.

B. DLTS results

A typical DLTS spectrum for a sputter-etched sample is shown in curve (b) of Fig. 4. Since the sputter-etched SBDs were leaky at temperatures above 250 K, the DLTS spectra depicted in Fig. 4 were recorded between 40 and 250 K. The defects are labeled P1 through to P8, and are identified in the form, for instance EP088 (P1), where "E" denotes an electron trap, "P" stands for plasma-etching related defect and "088" corresponds to the position in meV of the trap below the conduction band. Three well resolved peaks EP088 (P1), EP143 (P2) and EP454 (P8) were observed under normal conditions (filling pulses of 0.2 ms width). Figures 4(c) and 4(d) show that EP183 (P3) and EP310 (P6) could be fully resolved after using filling pulse widths of 20 ns, suggesting that their capture cross sections are larger than those of P4, P5 and P7, respectively. We have also used current (I)–DLTS to monitor the presence of a shallow electron trap with a level at $E_c-0.056$ eV, which could not be detected

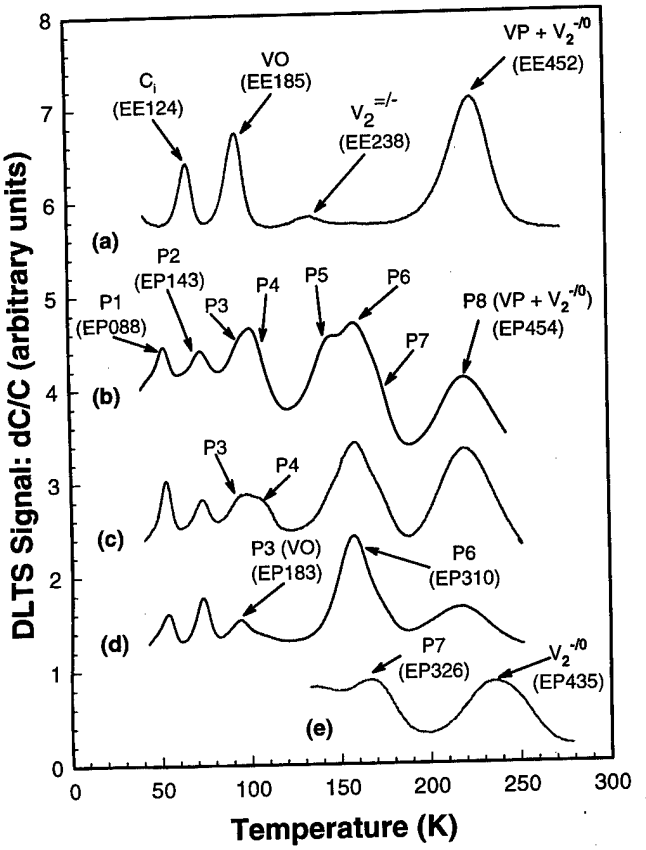


FIG. 4. DLTS spectra of epitaxially grown n-Si doped to 4.8×10^{15} cm^{-3} with P, bombarded with high energy (MeV) electrons [curve (a)], and sputter etched in an Ar plasma [curve (b)] using 0.2 ms filling pulses. Curves (c) and (d) from the same sample were obtained using filling pulse widths of 500 and 25 ns, respectively. Curve (e) was obtained after annealing at 250 °C. All curves were recorded at a lock-in amplifier frequency of 46 Hz, $V_r=1$ V and $V_p=1.4$ V.

using conventional capacitance (C)–DLTS. The Arrhenius plots from which the activation energy, E_t, and apparent capture cross section, σ_a, of the SEI defects were extracted are illustrated in Fig. 5.

Figures 6(a) and 6(b) illustrate the peak DLTS signal intensities of the prominent SEI defects as functions of etch

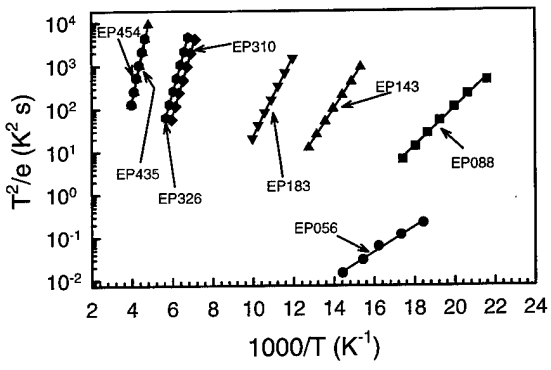


FIG. 5. Arrhenius plots of $\log(e/T^2)$ vs $1000/T$ from which the activation energy, E_T, and apparent capture cross section, σ_a, of the sputter-etching induced defects were extracted. The thermal dependence of the capture cross sections were not taken into account during the measurements.

time and plasma pressure, respectively. Figure 6(a) shows that the intensities of the SEI defects decreased nonmonotonically with etch time to reach a minimum at 6 min, and thereafter increased. Due to the high degree of carrier compensation, the DLTS signals for $t=2$ min are expected to be grossly underestimated and are, therefore, lower than the signals corresponding to $t=4$ min. Similarly, the relatively lower intensities of the SEI defects for 2×10^{-3} mbar compared to 9×10^{-3} mbar [Fig. 6(b)] have also been attributed to the heavy carrier compensation sustained at the lowest pressure [Fig. 3(b)]. These DLTS observations, therefore, support the $I-V$ and $C-V$ results shown in Figs. 1, 2, and 3. It is pointed out here that since the shallow dopant profiles vary with depth [as evidenced in Fig. 3(b)], the DLTS peak intensities shown in Fig. 6 cannot be used as equivalent measures for the defect concentrations.

DLTS depth profiling of EP310 and EP454 were done using the samples etched for 6 min at different plasma pressures. The defect profiles (not shown) revealed that the concentrations of the two prominent defects decreased exponentially in the region extending beyond 0.45 μm from the metal-semiconductor interface. We have fitted this exponential decay using Eq. (3),

$$N_T(x)=N_{T0}\exp(-x/L) \tag{3}$$

and extracted an "apparent" or extrapolated interface ($x=0$) concentration, N_{T0}, of the defects as well as their characteristic decay length, L. The extrapolated surface concentrations and characteristic lengths of EP310 and EP454 are summarized in Table I for the lowest and highest plasma pressures. The values of N_{T0} at 2×10^{-3} mbar were found to be marginally smaller than their corresponding values at 9×10^{-3} mbar for both EP310 and EP454.

EP088 and EP143 were completely removed after annealing at 100 °C, whereas the overlapping peaks of P3 and P4 were annihilated above 200 °C. The prominent defects EP310 and EP454 were annealed out at 300 and 250 °C, respectively. Isochronal annealing proved to be an efficient technique to resolve some of the overlapping DLTS defect peaks. The "signature" of P7 (EP326) was extracted after annealing an etched sample at 250 °C [curve (e) in Fig. 4]. The peak of P5, which has a level at $E_c-0.219$ eV in the band gap, was identified after removing both P6 and P7 at 300 °C. Figure 7 depicts the DLTS spectra from an annealed sputter-etched sample at 100, 400, 450, and 550 °C [curves (a), (b), (c), and (d), respectively]. It can be seen from curve (e) of Fig. 4 that EP454 (P8) was removed above 250 °C, after which an electron trap EP435, which was thermally stable up to 550 °C, could be observed (Fig. 7). EP203 was the main defect after annealing at 450 °C and could not be detected above 550 °C. The intensity of a defect EP219 was observed to increase above 400 °C, while that of EP203 decreased. EP219 was the main defect above 450 °C and its magnitude decreased when annealing above 550 °C. This defect was however found to be thermally stable at 650 °C. The energy levels, E_T, and apparent capture cross sections, σ_a, of the prominent SEI defects are summarized in Table II.

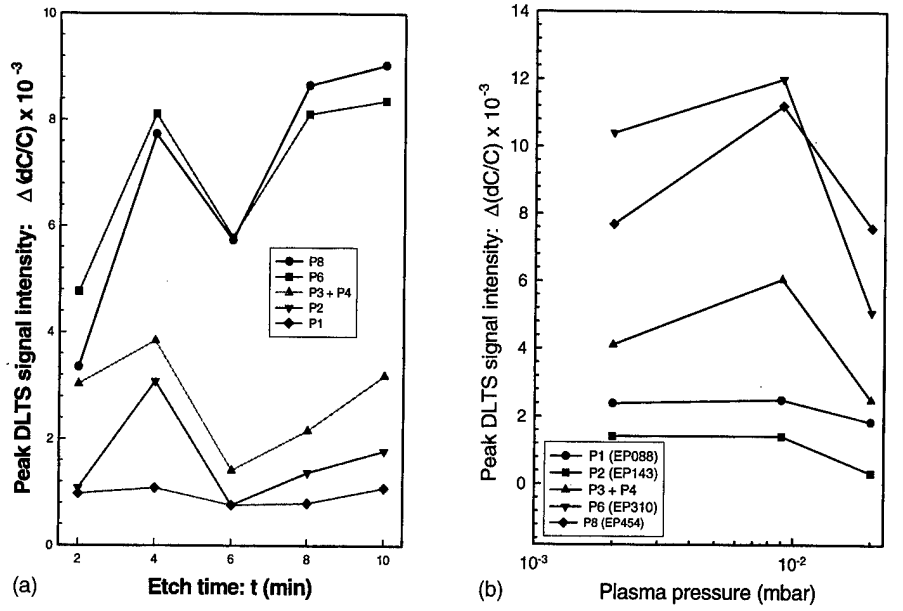


FIG. 6. Illustrates the variations in peak DLTS signal intensities of the prominent SEI defects as a function of etch time [curve (a)] and plasma pressure [curve (b)]. The data shown in curve (a) were obtained from sputter-etched epitaxially grown n-Si (doped to 4.8×10^{15} P/cm^{-3}) for 2, 4, 6, 8, and 10 min at a constant pressure of 9×10^{-3} mbar and a fixed bias of 400 V. The pressure dependence curve was extracted from samples etched at 2×10^{-2}, 9×10^{-3}, and 2×10^{-3} mbar at a fixed sputter time and bias of 6 min and 400 V, respectively.

This table also lists the electronic properties of the defects introduced in the same material by 1 keV He and Ar ions, and 5.4 MeV alpha-particle irradiation. The probable structure of the defects are given in the last column of the table.

IV. DISCUSSION

A. Electrical characteristics of SBDs and extent of SEI damage

The results presented above have shown that the rectifying properties of Pd SBDs fabricated on sputter-etched n-Si were poorer than those fabricated on a chemically cleaned but unetched substrate. We can explain the differences between the electrical properties of the diodes deposited on Si etched for different time periods by considering the following two competing time-dependent processes: (a) defect removal by Ar-ion sputtering and (b) defect in-diffusion. During the etching process, low-energy Ar ions sputter the Si surface whilst coming to rest, on average, at a depth of ~21 Å below the instantaneous surface. The projected range of the ions depends on their energy, which will be assumed to be equal to qV_{dc} eV throughout this discussion (V_{dc} being the dc bias on the electrode and q the electron charge). The low-energy implanted Ar ions transfer energy to the Si lattice predominantly by nuclear stopping and create damage in the crystal by displacing substrate atoms. Sputtering also results in the simultaneous erosion of the topmost atomic layers, and hence the removal of ion-induced damage. The thickness of damage material removed always lags the mean ion penetration (~21 Å) and the radiation enhanced diffusion of existing ion beam induced defects after ion etching has started. For relatively small etch times, the in-diffusion of defects supersedes their removal leading to accumulation of damage below the semiconductor surface. Furthermore, the etch rate requires a definite time to equilibrate, such that for small etch times the accumulation of damage is predominant over the sputtering process. As shown in Figs. 1 and 2, the quality of the diodes improved with etch time up to 6 min. Based on a glancing-angle Rutherford backscattering spectroscopy/channeling (RBS/C) study of the subsurface structural changes of sputter-etched Si (to be published elsewhere), we

TABLE I. Surface concentrations, N_{T0}, and characteristic lengths, L, of EP310 and EP454 in 6 min sputter-etched n-Si in an Ar plasma at 2×10^{-2} and 9×10^{-3} mbar.

Defect	Pressure (mbar)	N_{T0} (cm^{-3})	L (nm)
EP310	2×10^{-2}	$\sim8\times10^{12}$	95±16
	9×10^{-3}	$\sim8\times10^{13}$	95±16
EP454	2×10^{-2}	$\sim4.6\times10^{13}$	143±2
	9×10^{-3}	$\sim7.3\times10^{13}$	143±2

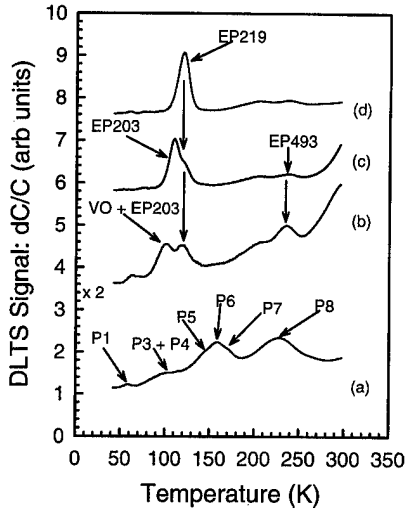


FIG. 7. DLTS spectra of annealed sputter-etched n-Si at 100 °C [curve (a)], 400 °C [curve (b)], 450 °C [curve (c)], and 550 °C [curve (d)]. The defect annealing stages were similar for all the sputter-etched samples. The spectra show that post-etching high temperature treatment of the samples removed most of the prominent sputter etching induced defects, but nonetheless introduced a main secondary defect EAr219.

TABLE II. Summary of the electronic properties (energy level, E_T, and apparent capture cross section, σ_a) of sputter etching induced defects in epitaxially grown n-Si, together with the properties of the defects introduced in the same material by 1 keV He and Ar ions, and 5.4 MeV alpha-particles.

Ion and energy	Defect	E_t (meV)	σ_a (cm^2)	T_{out} (°C)	Similar defects
Sputter etching	EP088	088±2	3.7×10^{-16}	100	C_i-C_s
	EP143	143±2	5.7×10^{-15}	100	$C_s-Si_i-C_s$
	EP183	183±2	3.4×10^{-15}	400	VO
	EP310	310±3	1.6×10^{-15}	300	EAr310
	EP203	203±2	1.7×10^{-15}	450	EHe203/EAr201
	EP454	454±4	4.2×10^{-16}	250	$VP+V_2^{-/0}$
He: 1 keV	EHe087	087±2	8.2×10^{-16}	100	C_i-C_s
	EHe113	113±2	3.6×10^{-15}	100	C_i-C_s
	EHe178	178±2	7.5×10^{-16}	400	VO
	EHe203	203±3	1.1×10^{-15}	450	EAr201/EP203
	EHe341	341±4	1.0×10^{-15}		
	EHe390	390±4	6.3×10^{-17}		
Ar: 1 keV	EAr098	098±2	6.6×10^{-17}	100	C_i-C_s
	EAr126	126±2	2.2×10^{-16}	100	$C_s-Si_i-C_s$
	EAr178	178±2	7.5×10^{-16}	400	VO
	EAr201	201±3	9.7×10^{-16}	450	EHe203/EP203
	EAr219	219±4	1.4×10^{-16}	>650	EP219
	EAr288	288±4	1.1×10^{-15}		
	EAr310	310±4	1.6×10^{-16}	?	EP310
	EAr377	377±5	1.1×10^{-17}		
He: 5.4 MeV	Eα178	178±2	7.5×10^{-16}	>400	VO
	Eα251	251±3	1.3×10^{-15}	>450	$V_2^{=/-}$
	Eα415	415±4	1.9×10^{-16}	>450	$V_2^{-/0}$
	Eα437	437±4	2.4×10^{-15}	200	VP

have tentatively attributed the lower rectifying characteristics of the diodes corresponding to etch times larger than 6 min to the formation of an amorphous layer with a corresponding continuous distribution of electronic states in the band gap.[20] The amount of etching-induced damage were further substantiated by our DLTS results. Figure 6(a) showed that the peak DLTS signal intensities of the main SEI defects followed similar trends as the $I-V$ characteristics of the diodes. Furthermore, the DLTS depth profiles of EP454 and EP310[20] revealed that the defect migrated to depths larger than the projected range of 400 eV Ar ions (21 Å), suggesting that diffusion of the plasma-etching induced defects could have been assisted by recombination-enhanced diffusion or transient enhanced diffusion.[21,22]

The plasma pressure dependence of the rectifying characteristics of the diodes is explained by considering the mean free path, λ, between collision of Ar ions with neutral Ar atoms in the cathode sheath. Since λ is inversely proportional to the plasma density, the frequency of collision of an Ar ion in a denser plasma is higher. The ion is thus accelerated over a shorter distance inside the discharge and its kinetic energy before a collision is lower compared to another ion in a less dense environment. The Ar ions in a low-pressure plasma, therefore, bombard the Si surface with relatively higher energies compared to Ar ions in a denser plasma. The result is that more damage is introduced in the semiconductor lattice at lower plasma pressures. As shown in Fig. 3(a), the electrical properties of the diodes became poorer, while Fig. 3(b) illustrated that the compensation in free carrier concentration of the etched samples increased, with decreasing plasma pressure. The DLTS plots illustrated in Fig. 6(b) can be used to further validate our results. Since the DLTS signal is proportional to the free carrier concentration in the region probed, a lower than expected defect intensity is measured in a sample which incurred heavier carrier compensation. Hence, taking into account the higher carrier concentration in the region below 0.45 μm in the sample etched at a plasma pressure of 2×10^{-3} mbar compared to 9×10^{-3} mbar [Fig. 3(b)], it can be expected that the data points corresponding to a pressure of 2×10^{-3} mbar in Fig. 6(b) are higher than those corresponding to a pressure of 9×10^{-3} mbar. In this case, the results shown in DLTS results depicted in Fig. 6(b) are in agreement with those shown in Fig. 3(b).

Figure 3(b) also revealed that there was partial carrier compensation in the control sample. We have attributed this to dopant passivation by hydrogen incorporation in the samples during the premetallization chemical cleaning step. It can, however, be inferred from the same graph that sputter etching introduced additional compensating centers. Since ultrahigh purity argon gas was used during sputter etching and given that both the etched and control sample were subject to the same chemical cleaning prior to metallization, we have proposed that these compensating centres are most likely to be acceptor-type defects. Alternatively, the observed free carrier compensation was due to dopant passivation in the form of dopant-defect pairs.

B. Electronic, structural and annealing properties of SEI defects

Curve (a) of Fig. 4 shows the DLTS spectrum of the structurally identified defects induced in n-Si during MeV electron irradiation. We have observed that high-energy alpha-particle and proton irradiation introduced levels at $E_c-0.08$ eV and $E_c-0.14$ eV, in addition to the VO and VP centers, and the two charge states of the divacancy ($V_2^{=/-}$ and $V_2^{-/0}$). $V_2^{-/0}$ is only detected after removing the VP center at 180 °C. EE124 (the second "E" in the nomenclature refers to "electron irradiated") was removed and levels at $E_c-0.141$ eV and $E_c-0.085$ eV, which have similar electronic "signatures" as EP143 and EP088, respectively, emerged after annealing our electron-irradiated sample at 180 °C for 30 min. These strongly suggest that EE124 is structurally identical to the carbon interstitial (C_i). Since EP088 was reversibly removed after cooling under zero bias ($V_R=0$ V) and reintroduced under a reverse bias excursion ($V_R=-2$ V) and noting that the intensity of EP143 increased and decreased, respectively, under these cycles, the two defects have been assigned the $C_s-Si_i-C_s$ and C_i-C_s structures, respectively.[23] EP435 could be detected after removing EP454 at 250 °C [Fig. 4(e)] and its intensity was found to be approximately six times less than that of EP454 [Fig. 7(b)]. Furthermore, EP435 was observed to be thermally stable up to 500 °C [Fig. 7(b)]. The "signature" of EP454 together with its annealing behavior suggest that it is

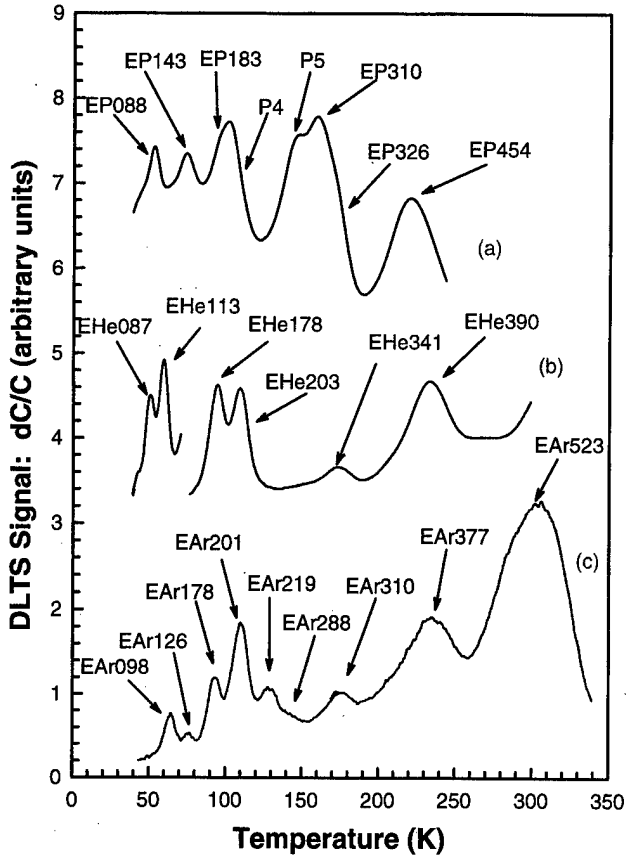


FIG. 8. Comparison of DLTS spectra of epitaxially grown n-Si doped to 4.8×10^{15} cm^{-3} with P. (a) Sputter etched in an Ar plasma at a pressure of 9×10^{-3} mbar for 6 min (dc bias of 400 V), (b) bombarded with 1 keV He ions and (c) 1 keV Ar ions. The 1 keV bombardments were done to a dose of 1×10^{12} cm^{-2} using a sputter gun. A lock-in amplifier frequency of 46 Hz was used to extract the data.

the superposition of the VP center and $V_2^{-/0}$ (EP435). No direct evidence could be obtained from our results to show the detection of $V_2^{=/-}$. Since $V_2^{-/0}$ is observed after the VP center anneals out, we argue that the high level of SEI damage in the region probed by DLTS created stress fields which prevented the complete filling of $V_2^{=/-}$.[24] Current DLTS further showed that a shallow electron trap with a level at $E_c - 0.056$ eV was also created during the plasma-etching process.

During any sputter-etching process, the substrate is subjected to bombardment by energetic noble gas ions from the plasma. In order to further characterize the SEI defects, we have compared them with those introduced during 1 keV Ar-ion bombardment of the same material. Since photoluminescence experiments have demonstrated that ion beam etching of Si using low-energy noble gas ions created noble-gas-related defects,[25,26] we have also compared the SEI defects to those created by 1 keV He-ion bombardment. We have recently shown that 1 keV He-ion bombardment of Si introduced a different set of defects compared to 5.4 MeV alpha-particle irradiation of the same material.[27] Figure 8 shows the DLTS spectra of defects created during 1 keV Ar- and He-ion bombardment of n-Si [curves (c) and (b), respectively]. From the curves in Figs. 7 and 8, P4 is matched to EAr201 and EHe203, because they have similar "signatures" and annealing properties. This comparison further concludes that P4 is not a noble gas related defect. Since the electronic properties of EP310 and EP326 are, within experimental error, similar to that of EAr310, we are currently conducting annealing studies on 1 keV Ar-ion bombarded Si samples to elucidate which of the two could be similar to EAr310. Furthermore, EP493, which is observed after annealing at 400 °C, has a similar "signature" to EAr377, while EP219 (dominant above 550 °C) is electronically similar to EAr219.

V. CONCLUSIONS AND SUMMARY

The electronic and annealing properties of defects introduced in n-Si during sputter etching, together with the electrical characteristics of the Schottky barrier diodes fabricated on the etched samples, have been studied as a function of etch time and plasma pressure. Our results have shown that sputter etching caused carrier compensation in n-Si. These compensating centers have been attributed to the etching-induced acceptor-type defects or dopant passivation through the creation of dopant-defect complexes. Capacitance–voltage depth profiles further revealed that the compensating centers migrated beyond the penetration depth of the low-energy Ar ions from the plasma. The rectifying properties of the SBDs were less degraded for relatively long etch periods and higher plasma pressure. Our results have demonstrated that sputter etching introduced electrically active defects which altered the electrical properties of Schottky barrier diodes fabricated on the etched surfaces. Some of the SEI defects were successfully identified as being structurally identical to C_i–C_s, the VO and VP centres and $V_2^{-/0}$. The nondetection of $V_2^{=/-}$ has been attributed to the presence of stress fields in our sputter-etched samples. Annealing could remove the SEI defects but also introduced a prominent defect EP219 which was thermally stable at 650 °C. EP203 has a similar DLTS "signature" and annealing behavior as EAr201, produced by 1 keV Ar-ion bombardment. A similar defect (EHe203) was also observed in low-energy He-ion bombarded n-type Si, which suggests that EP203 is not a noble gas related defect.

ACKNOWLEDGMENTS

The financial assistance of the FRD and CEFIM are gratefully acknowledged, as well as the technical assistance of J. P. le Roux and C. M. du Toit. The authors also thank C. Schutte for making the Universal Sputter Machine at DETEK available to us. The authors are also thankful to Mark C. Ridgway for critically reviewing the article and to J. B. Malherbe for the 1 keV He- and Ar-ion bombardments.

[1]F. H. Mullins and A. Brunnschweiler, Solid-State Electron. **19**, 47 (1976).
[2]S. J. Pearton, W. S. Hobson, U. K. Chakrabarti, G. E. Derkits, Jr., and A. P. Kinsella, J. Electrochem. Soc. **137**, 3892 (1990).
[3]G. S. Oehrlein, Mater. Sci. Eng., B **4**, 441 (1989).
[4]M. Lannoo, Mater. Res. Soc. Symp. Proc. **262**, 775 (1992).
[5]S. D. Brotherton and P. Bradley, J. Appl. Phys. **53**, 5720 (1982).

[6]F. Capasso, A. Y. Cho, K. Mohammed, and P. W. Foy, Appl. Phys. Lett. **46**, 664 (1985).
[7]S. Ashok, T. P. Chow, and B. J. Baliga, Appl. Phys. Lett. **42**, 687 (1983).
[8]S. J. Fonash, S. Ashok, and R. Singh, Appl. Phys. Lett. **39**, 423 (1981).
[9]M. Hirai, H. Iwakuro, J. Ohno, and T. Kuroda, IEEE Trans. Compon., Hybrids, Manuf. Technol. **13**, 629 (1990).
[10]O. O. Awadelkarim, T. Gu, P. I. Mikulan, R. A. Ditizio, and S. J. Fonash, Appl. Phys. Lett. **62**, 958 (1993).
[11]M. Balooch, M. Moalem, Wei-E. Wang, and A. V. Hamza, J. Vac. Sci. Technol. A **14**, 229 (1996).
[12]R. Singh, S. J. Fonash, S. Ashok, P. J. Caplan, J. Shappirio, M. Hage-Ali, and J. Ponpon, J. Vac. Sci. Technol. A **1**, 334 (1983).
[13]J. R. Troxell, Solid-State Electron. **26**, 539 (1983).
[14]D. V. Lang, J. Appl. Phys. **45**, 3014 (1974).
[15]Y. Zohta and M. O. Watanabe, J. Appl. Phys. **53**, 1809 (1982).
[16]E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts*, 2nd ed. (Clarendon, Oxford, 1988), p. 38.
[17]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), p. 263.
[18]D. Bauza, J. Appl. Phys. **73**, 1858 (1993).
[19]P. N. K. Deenapanray, F. D. Auret, G. Myburg, M. Hayes, W. E. Meyer, and C. Schutte, Mater. Sci. Forum **248–249**, 249 (1997).
[20]P. N. K. Deenapanray, F. D. Auret, C. Schutte, G. Myburg, W. E. Meyer, J. B. Malherbe, and M. C. Ridgway, Mater. Res. Soc. Symp. Proc. **442**, 87 (1997).
[21]J. L. Benton, J. Michel, L. C. Kimerling, B. E. Weir, and R. A. Gottscho, J. Electron. Mater. **20**, 643 (1991).
[22]B. G. Svensson, C. Jagadish, and J. S. Williams, Phys. Rev. Lett. **70**, 3756 (1993).
[23]J. L. Benton, M. T. Asom, R. Sauer, and L. C. Kimerling, Mater. Res. Soc. Symp. Proc. **104**, 85 (1988).
[24]B. G. Svensson, B. Mohadjeri, A. Allen, J. H. Svensson, and J. W. Corbett, Phys. Rev. B **43**, 2292 (1991).
[25]N. Bürger, K. Thonke, R. Sauer, and G. Pensl, Phys. Rev. Lett. **52**, 1645 (1984).
[26]J. Weber, Physica B **170**, 201 (1991).
[27]F. D. Auret, P. N. K. Deenapanray, W. E. Meyer, S. A. Goodman, and G. Myburg, J. Appl. Phys. (to be published).

Low frequency noise in heavily doped polysilicon thin film resistors

M. J. Deen, S. Rumyantsev,[a] and J. Orchard-Webb[b),c)]
School of Engineering Science, Simon Fraser University, Vancouver, British Columbia V5A 1S6, Canada

(Received 17 April 1998; accepted 22 April 1998)

Low frequency noise and current–voltage measurements in several heavily doped polysilicon resistors of varying geometry and both p and n type, and over a limited range of temperatures from −60 to 50 °C were conducted for the first time. We found that the noise in p-type polysilicon was independent of temperature, but not the n-type polysilicon. For the p-type resistors, linear current–voltage characteristics were observed, and the relative noise spectral density was independent of bias and inversely proportional to frequency. For the n-type resistors, linear current–voltage characteristics were observed, and the relative noise spectral density was independent of bias. Finally, the normalized noise level in the linear n-type resistors was almost an order of magnitude lower than for the p-type resistors. We believe that this difference is because n-type dopants segregate to the grain boundaries, thus passivating some of the traps there. Boron (p-type dopant), on the other hand, does not segregate to the grain boundaries, leaving more unpassivated grain-boundary traps which capture and emit more carriers, resulting in more low frequency noise.
© *1998 American Vacuum Society.* [S0734-211X(98)01704-1]

I. INTRODUCTION

Low frequency noise in semiconductor materials and passive and active elements is important in determining the quality of the materials,[1–8] or in determining the amount of noise the elements will contribute to the overall noise of a circuit or system.[9] In some applications, the low frequency noise sets a low limit on the sensitivity of the circuit, for example, in avalanche photodiodes that are used to detect and convert optical radiation into an electrical signal.[2,3]

There are three main types of low frequency noise. First, white noise (shot noise or thermal noise) whose magnitude is independent of frequencies up to very high frequencies. Second, $1/f^{\gamma}$-type noise, and for γ between 0.8 and 1.2, it is generally called $1/f$ noise. Third, generation-recombination noise where the noise is constant up to a corner frequency, and then it decreases at 20 dB/decade as the frequency increases.[10,11] The origins of shot noise, thermal noise, and generation-recombination noise are based on well-understood physical effects. However, the origin of $1/f$ noise is still a subject of intensive research and no single theory has been able to explain it for all the semiconductor materials or devices in which it has been observed.[12]

Low frequency noise has previously been studied in avalanche photodiodes,[2] resonant tunneling diodes,[1,7] metal–oxide–semiconductor field effect transistors (MOSFETs),[8,10,13–15] and bipolar transistors.[4,5,10] We have also studied high-temperature effects in heavily doped silicon[15] and polycrystalline titanium silicide[16] resistors. To date, there are only few publications about noise in silicon grain boundaries[17,18] and polycrystalline thin films and polycry stalline silicon resistors.[19] In these papers the noise properties of low and moderately doped polysilicon samples were investigated. In Ref. 19 the noise in heavily doped polysilicon with doping level $\sim 10^{19}$ cm^{-3} was measured. Here, the purpose of this research is not to study the origin of $1/f$ noise, but rather to quantify the low frequency noise in several very heavily doped polysilicon resistors as a function of type of resistor, biasing, temperature, and geometry. These resistors are typically used in telecommunications, mixed signal circuits, general analog circuits, or high-voltage circuits.

II. DEVICE AND EXPERIMENTAL DETAILS

The n-type polysilicon resistors were fabricated using a heavy POCL (liquid source of phosphorus) deposition and diffusion. The resulting n-type dopant concentration is $\sim 5 \times 10^{20}$ cm^{-3}. The p-type polysilicon films, on the other hand, are implanted with boron, and then annealed by heat treatment. The resulting p-type dopant concentration is $\sim 2 \times 10^{19}$ cm^{-3}. The films are then patterned to different length-to-width ratios for study.

Polysilicon resistors R_p in series with a metal resistor R_0 were biased with 12 or 24 V batteries, as shown in Fig. 1. The voltage fluctuations across resistors were amplified with a PAR 113 low noise amplifier and then measured with an HP 3561 dynamic signal analyzer. For each noise spectra, 200 averages using the dynamic signal analyzer were made. Noise measurements were made on at least four samples of each of the resistor geometries for both n- and p-type resistors. The results that will be presented are averages from these samples in which their noise spectra were almost identical. All samples were in a ceramic packages. The noise spectra were measured for frequencies between 1 Hz and 100 kHz. However, data were collected in the computer for the frequency points shown for each of the graphs. For some spectra, all the measured frequency points were collected (for example, Fig. 6), while for others, fewer frequency points were collected (for example, Fig. 2). The instrument

[a)]Present address: Ioffe Institute, 194021, St. Petersburg, Russia.
[b)]Present address: Mitel Telecomms., Mitel Business Park, Portskewett, Gwen, Wales NP64YR.
[c)]Electronic mail: jamal@cs.sfu.ca

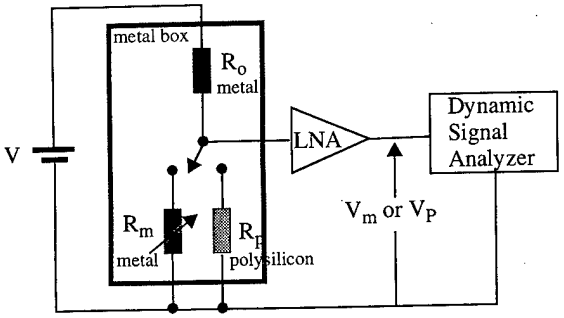


FIG. 1. Experimental setup for noise measurements.

used is a dynamic signal analyzer which makes a Fourier transform of the signal to display the power spectral density.

The gain of the low noise amplifier was set so as not to overload the dynamic signal analyzer, and to sufficiently amplify the voltage fluctuations. The frequency range of the dynamic signal analyzer was set from 1 Hz to 100 kHz and the measurement mode selected to measure noise voltage power spectral density. After the noise voltage density measurements (S_p normalized to 1 Hz) with the polysilicon resistor R_P were made, the polysilicon resistor was replaced with a metal resistor R_m of same resistance value of R_P (i.e., $R_m = R_P$ and now R_0 is in series with R_m) and the entire measurement was repeated using the test setup shown in Fig. 1. It is known that the noise of the amplifier depends on the loading resistance. Therefore, this procedure allowed us to determine the noise of the amplifier under measurement conditions with high accuracy. Using this procedure, we can calculate the spectral noise density of voltage fluctuations S_V (normalized to 1 Hz)[20,21] as

$$S_V \cong \left(\frac{R_0 + R_p}{R_0}\right)^2 \cdot S_p \tag{1}$$

for the case when the noise of the polysilicon resistor is much larger than the noise of the metal resistor.

III. MEASUREMENT RESULTS

Polysilicon n-type resistors with dimensions $L/W = 150/75\ \mu$m, $L/W = 225/75\ \mu$m, and $L/W = 400/4\ \mu$m were measured. However, only for samples with relatively low area did the noise level significantly exceed the background noise of the amplifier.

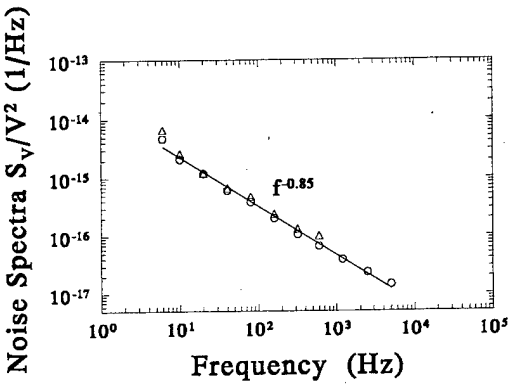


FIG. 2. Normalized noise spectra for n-type polysilicon resistors for two biases $V_p = 5$ V (triangles) and $V_p = 10$ V with $L/W = 400/4\ \mu$m.

If the noise observed is determined by bulk or surface mechanisms and not by the contacts, then the spectral noise density should be inversely proportional to the volume or area, for bulk or surface noise mechanisms, respectively, of the resistor. The polysilicon resistors with $L/W = 150/75\ \mu$m and $L/W = 225/75\ \mu$m have significantly higher area than the polysilicon resistor with dimensions $L/W = 400/4\ \mu$m, and therefore, both were of low noise because of their larger areas (their thicknesses were all the same, 0.25 μm).

The noise spectra of both n- and p-type samples were fitted to the expression

$$S_{V,n} = \frac{K_F \cdot V_p^2}{f^\gamma}. \tag{2}$$

Then their geometry was included in the expression to normalize the K_F values. In this way, for a given technology, circuit designers can select an appropriate geometry (width and length) to minimize the noise level. For all samples measured, the average K_F values normalized to the geometry will be given. The standard deviations of the K_F values were less than 5% of their average values.

Figure 2 shows the relative noise spectra for polysilicon n-type resistor with dimensions $L/W = 400/4\ \mu$m. It is seen that for $V_p = 5$ or 10 V, the noise spectra practically overlap one another. The independence of relative noise spectra on bias conditions is a property of linear metallic and semiconductor resistors. The noise spectra are of the form $1/f^\gamma$ with $\gamma = 0.85$. The spectral noise density of polysilicon n-type resistors may be estimated (from four samples) using the following expression:

$$S_{V,n} = \frac{2.5 \times 10^{-11} V_p^2}{LWf^{0.85}}. \tag{3}$$

This relatively low noise, we believe, is a result of segregation of the n-type dopants (phosphorus) to the grain boundary,[22] thus possibly passivating some of the grain-boundary traps which produce noise by trapping and detrapping of mobile carriers. Second, the n-type samples are heavily doped ($\sim 5 \times 10^{20}$ cm^{-3}), thus the potential barrier at the grain boundary is very low, making the noise from the grain boundary proportionally lower. Also, the potential barrier at the grain boundary is higher for the p-type samples because of their lower doping, thus the noise originating form the grain boundary is relatively higher.

Figures 3, 4, and 5 show the relative noise spectra for p-type polysilicon resistors. Figure 4 shows the noise spectra at three bias conditions. It is seen that the relative noise spectra do not depend on bias conditions, similar to what was observed for n-type resistors. From Fig. 4, we note that the level of noise depends only on the resistor's dimensions. For p-type polysilicon resistors, the noise spectra is given by

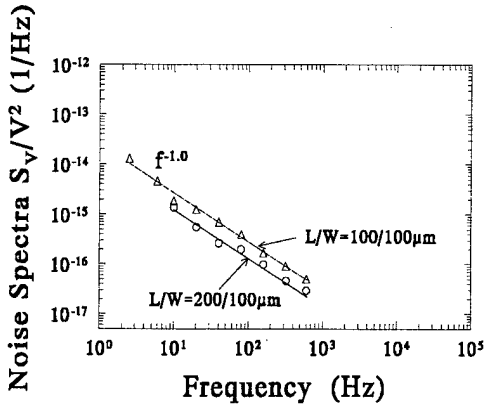


FIG. 3. Normalized noise spectra for p-type polysilicon resistors of different dimensions for $V_p = 10$ V.

$$S_{V,p} = \frac{5 \times 10^{-10} V_p^2}{LWf}. \tag{4}$$

Therefore, polysilicon p-type resistors are characterized by a higher level of noise than polysilicon n-type resistors. This is expected since the p-type dopants (boron) do not segregate to the grain boundaries. Thus there are many unpassivated grain-boundary traps which capture and emit carriers, resulting in a higher level of noise than the n-type polysilicon, as explained above.

Only one of the polysilicon p-type resistors measured exhibited a very high level of noise. The relative noise spectra of this sample is shown in Fig. 5. At low biases, this resistor was characterized by a linear current–voltage characteristics, in common with all other resistors measured. The noise spectral density was proportional of V_p^2. Only at higher biases above 1 V was a superlinear $I-V$ dependence observed and a sublinear deviation from $S_V - V_p^2$ dependence obtained. We believe that this feature of the resistor is determined by the poor quality of the contact, or possible by heating effects at biases above 1 V. If the nonlinearity is due to poor contact, then we note that this poor contact condition may be determined by noise measurements at low biases (because of the higher noise obtained compared to other similar resistors), even if the $I-V$ relation of the contact is linear at these biases.

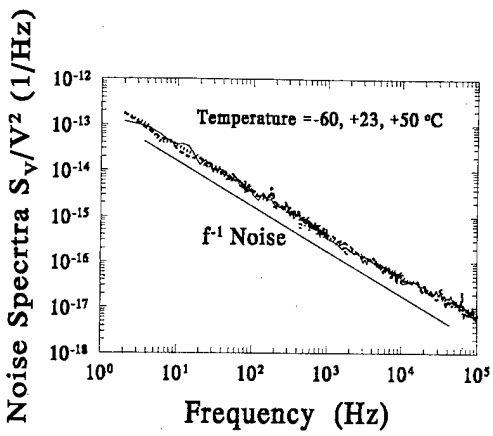


FIG. 5. Frequency dependence of normalized noise at different three temperatures of −60, 23, and 50 °C for p-type polysilicon resistors.

The temperature dependence of low frequency noise in some polysilicon resistors was also measured. The relative spectral noise density for p-type polysilicon resistor over the temperature range from −60 to +50 °C, shown in Fig. 6, did not depend on temperature. Figure 7 shows the temperature dependence of the noise in polysilicon n-type resistors at different frequencies. It is seen that n-type polysilicon resistors under investigation are characterized by a monotonic and rather weak temperature dependence of the noise.

Expressions (3) and (4) allow us to estimate the noise level in different polysilicon resistors. Another way of noise level characterization is using of the Hooge parameter α, given by

$$\alpha = \frac{S_V}{V_p^2} f^{\gamma} N, \tag{5}$$

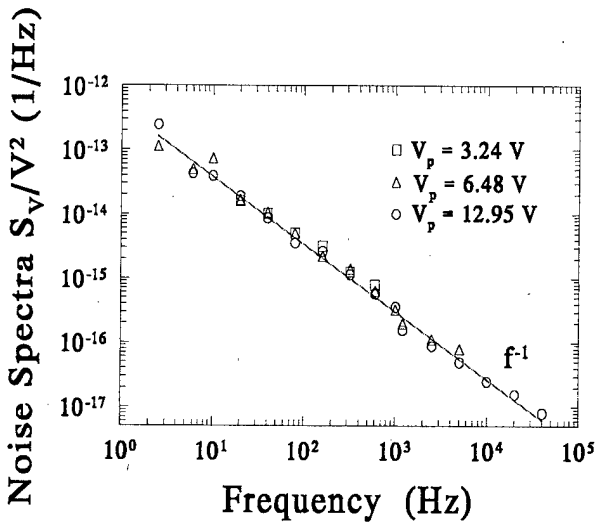


FIG. 4. Normalized noise spectra for p-type polysilicon resistors for three different biases with $L/W = 400/4$ μm.

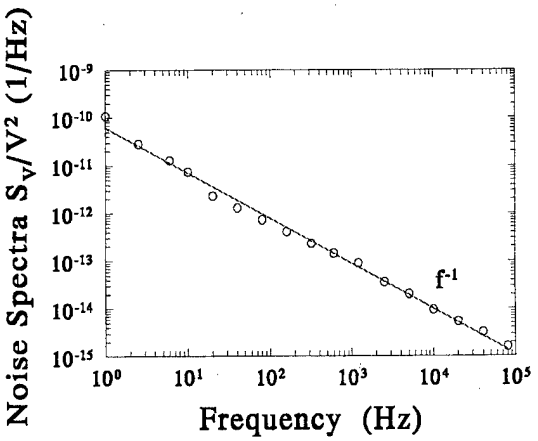


FIG. 6. Normalized noise spectrum for p-type polysilicon resistors with very high level of the noise for $V_p = 1$ V.

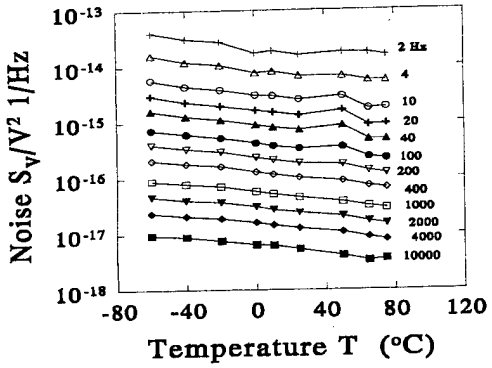


FIG. 7. Temperature dependence of normalized noise at different frequencies for n-type polysilicon resistors.

where N is the total number of the carriers in the sample and γ is typically unity or very close to it. In Refs. 18 and 19, the models of the $1/f$ noise (there, $\gamma=1$) in polysilicon resistors were developed on the basis of Eq. (5) and the values of α were calculated. In Refs. 23 and 24, the frequency exponent γ was allowed to vary, but were still close to unity, similar to our results reported here. In the usual expression, $\gamma=1$, and this is the case for our results except for the n-type resistors. Note also that the relation (5) is empirical and does not carry information about the origin of the noise. That is why the use of it as a basis for noise calculations in such a complicated system as polysilicon resistors, from our point of view, is not justified.

On the other hand, Eq. (5) is very convenient for comparison and estimation of the noise level in different materials. Using this approach, we estimated α taking into account that $N=nV$ (n is doping concentration per cm^{-3}, V is volume of the sample). The α values are $\alpha=3\times10^{-3}$ $\mathrm{Hz}^{-0.15}$ and $\alpha=15\times10^{-3}$ for n- and p-type resistors, respectively. Therefore the noise level for n-type resistors is rather small and typical for metals. The noise level for p-type resistors is higher, that is why this kind of resistors can cause additional noise in integrated circuits.

IV. CONCLUSIONS

Almost all n- and p-type samples of heavily doped polysilicon exhibited linear $I-V$ behavior. The polysilicon resistors under investigation exhibit $1/f^{\gamma}$ with $\gamma=0.85$ and $\gamma=1.0$ for n- and p-type resistors, respectively. The noise level of n-type resistors may be characterized by the expression (2) or by the Hooge parameter $\alpha=3\times10^{-3}$. Polysilicon p-type resistors are characterized by a higher level of the noise with $\alpha=15\times10^{-3}$ [see also expression (3)]. One reason for the higher noise in p-type polysilicon resistors is the larger number of unpassivated grain-boundary traps. In addition, for n-type polysilicon with very high doping, the grain-boundary potential barrier is lower compared to the (relatively) lower doped p-type films, thus making the noise from the grain boundary proportionally lower.

One sample had nonlinear $I-V$ characteristics at higher biases (above 1 V) and a very large $1/f$ noise. Deviation from $S_V-V_p^2$ was also observed at these higher biases. These results most probably indicate a poor contact. No temperature dependence of noise in p-type polysilicon resistors was observed. Polysilicon n-type resistors are characterized by the weak temperature dependence of the noise. Finally, increasing the grain size, passivating more traps in the grain boundary, and reducing the grain-boundary potential barrier would result in reduced low frequency noise.

ACKNOWLEDGMENTS

This research was supported in part by grants from Mitel Semiconductor, Micronet, a federal network center of excellence in microelectronics and NSERC of Canada. The authors would also like to thank Alain Comeau of Mitel AB, Sweden for his valuable comments on this research.

Presented at the 8th Canadian Semiconductor Technical Conference, 12–15 August 1997.

[1]M. J. Deen, Mater. Sci. Eng., B **20**, 207 (1993).
[2]J. J. Goedbloed and E. Smeets, Electron. Lett. **14**, 67 (1978).
[3]X. Zhao, M. J. Deen, and L. Tarof, IEE Electron. Lett. **32**, 250 (1996).
[4]A. Ng, M. J. Deen, and J. Ilowski, Can. J. Phys. **70**, 949 (1992).
[5]M. Doan, Z. Buffet, and M. J. Deen, in *Proceedings of the Symposium on the Degradation of Electronic Devices Due to Device Operation as Well as Crystalline and Process-Induced Defects,* Proceedings Vol. 94-1, edited by H. J. Quessnier, J. E. Chung, K. E. Bean, T. J. Shaffner, and H. Tsuya (The Electrochemical Society, Penington, NJ, 1994), p. 235.
[6]A. Van Der Ziel and H. Tong, Electronics **95**, (1966).
[7]M. J. Deen, 23rd European Solid-State Device Research Conference (ESSDERC '93), Grenoble, France, 1993, p. 355.
[8]M. J. Deen and C. Quon, *7th Biennial European Conference—Insulating Films on Semiconductors (INFOS 91), Liverpool, United Kingdom, 2–5 April, 1991,* edited by W. Eccleston and M. Uren (IOP, Bristol, UK, 1991), p. 295.
[9]R. P. Jindal, IEEE Trans. Electron Devices **34**, 305 (1987).
[10]M. J. Deen, J. I. Ilowski, and P. Yang, J. Appl. Phys. **77**, 6278 (1995).
[11]Y. Zhu, M. J. Deen, and T. M. Kleinpenning, J. Appl. Phys. **72**, 5990 (1992).
[12]D. A. Evans and P. T. Landsberg, Proc. R. Soc. London A **267**, 464 (1962).
[13]F. N. Hooge, IEEE Trans. Electron Devices **41**, 1926 (1994).
[14]J. Chang, A. Abidi, and C. R. Vishwanathan, IEEE Trans. Electron Devices **41**, 1965 (1994).
[15]F. Grabowski, Solid-State Electron. **32**, 909 (1989).
[16]M. J. Deen, A. A. Naem, and L. Y. Chee, J. Appl. Phys. **76**, 5253 (1994).
[17]A. A. Naem, M. J. Deen, and L. Chee, J. Appl. Phys. **76**, 1071 (1994).
[18]A. Madenbach and J. Werner, Phys. Rev. **38**, 13150 (1988).
[19]M. H. Luo and G. Bosman, IEEE Trans. Electron Devices **37**, 768 (1990).
[20]H. C. DeGraaf and M. T. M. Huybers, J. Appl. Phys. **54**, 2504 (1983).
[21]X. M. Li, M. J. Deen, S. Stapleton, R. H. S. Hardy, and O. Berolo, Cryogenics **30**, 1140 (1990).
[22]M. M. Mandurah, K. C. Sarawat, C. R. Helms, and T. I. Kamins, J. Appl. Phys. **51**, 5755 (1980).
[23]X. C. Chu, A. Van der Ziel, and K. H. Duh, Solid-State Electron. **28**, 325 (1985).
[24]M. Tacano, Y. Sugiyama, T. Taguchi, and H. Soga, Solid-State Electron. **31**, 1215 (1988).

Damage-free cleaning of Si(001) using glancing-angle ion bombardment

Jose Gregorio C. Labanda[a] and Scott A. Barnett[b]
Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208

L. Hultman
Thin Film Physics Division, Department of Physics, Linköping University, S-581 83 Linköping, Sweden

(Received 8 October 1997; accepted 8 May 1998)

The effects of glancing-angle argon ion bombardment on air-contaminated Si(001) surfaces were studied. Bombarding at substrate temperature $T_s = 730$ °C, impingement angle $\phi = 3-15°$ relative to the surface plane, ion energy $E = 1$ keV and dose $D = 3\times10^{15}$ ions cm^{-2} gave high-quality reflection high energy electron diffraction (RHEED) patterns and contaminant-free surfaces as observed by ion scattering spectroscopy. Atomic force microscopy images showed roughness value ≤ 0.5 nm under these conditions, but the roughness increased and RHEED patterns became spotty for higher doses or energies. Secco etching of samples bombarded at $T_s = 730$ °C showed etch pits with a density of 10^6-10^7 cm^{-2} that increased with increasing D and E. Room-temperature bombardment with $E = 1$ keV, $D = 3\times10^{15}$ ions cm^{-2} and $\phi = 3°$, followed by a 730 °C anneal, yielded a lowest roughness value of 0.2 nm. Secco etching showed no resolvable pits, indicating a dislocation density $<4\times10^4$ cm^{-2}. © *1998 American Vacuum Society.*
[S0734-211X(98)04704-0]

I. INTRODUCTION

Ion bombardment has been widely used for cleaning Si surfaces, but can lead to defect formation,[1–3] surface roughening,[4–6] and recoil implantation of surface contaminants.[7] Bean *et al.*[1] showed that 1.0 keV Ar ion bombardment of Si at low temperature causes amorphization, and after annealing to recrystallize the surfaces, dark-spot defects were present in transmission electron microscope images. Lowest defect densities were observed for an ambient-temperature ion bombardment followed by an 800 °C anneal which regrew the amorphous layer. At higher temperatures, damage is produced in the form of subsurface dislocation loops. Garverick *et al.*[7] showed that damage can also be minimized by sputter cleaning at elevated temperatures with very-low-energy ($\leq$100 eV) Ar ions at low doses.[7,8] Device-quality Si has been grown on Si cleaned using $\approx$25 eV Ar ions.[8] These very-low-energy sputter cleaning experiments were carried out by immersing the Si in a glow discharge plasma typically operating at pressures of $>10^{-3}$ Torr.

Glancing-angle Ar ion bombardment has recently been shown effective for producing clean, flat, nearly defect-free GaAs(001) surfaces.[9,10] The damage was found to decrease dramatically with decreasing angle ϕ relative to the surface plane. The low damage was explained by increased ion reflection, along with decreased ion implantation range, damage range, and ion channeling. Selective sputtering of asperities at low ϕ was found to smoothen preroughened surfaces.[9] To our knowledge, this technique has not been attempted for Si.

In this article, we describe the effects of glancing-angle Ar ion bombardment of Si(001) surfaces. Surface structure and morphology were studied using reflection high energy electron diffraction (RHEED) and atomic force microscopy (AFM) as a function of substrate temperature T_s, ion dose D, energy E and impingement angle ϕ relative to the surface plane. Surface composition was measured using *in situ* ion scattering spectroscopy. Defect densities were observed by chemical etching using the Secco formulation and cross-sectional transmission electron microscopy (XTEM). Best results were obtained using an ambient-temperature bombardment with $E = 1$ keV, $D = 3\times10^{15}$ ions cm^{-2} and $\phi = 3°$ followed by a 730 °C anneal, where the surface roughness $R = 0.2$ nm and no defect-etch pits were observed.

II. EXPERIMENTAL PROCEDURE

The experiments were carried out in a turbomolecular-pumped molecular beam epitaxy (MBE) chamber with a base pressure of $\sim10^{-9}$ Torr.[11] RHEED observations were made with a 25 keV electron beam incident at 7 mrad from the surface. A rasterable ion gun provided typical target currents of 0.1–5 μA at $E = 0.1-5.0$ keV. Ion dose values quoted are ions per unit substrate area. Two 70 mm ports at 3° and 18° relative to the nominal substrate surface plane combined with $\pm10°$ substrate tilting[12] allowed ϕ values from 0° to 28°. The source gases were 99.999% pure Ar or He further purified using a Ti getter. The chamber pressure with either Ar or He was typically 1×10^{-6} Torr with differential pumping using an 80 ℓ/s turbomolecular pump.

The polished Si(001) substrates were cleaned with acetone and methanol prior to insertion into the MBE chamber, and then annealed for 10 min at T_s=450 to 730 °C. Bombardment was then carried out with Ar ions impinging in the $\langle 110\rangle$ direction while maintaining the substrate at the same annealing temperature unless otherwise noted. T_s was measured via the silicon substrate resistivity after direct calibration with a thermocouple and observing the melting points of InSb and Al.

[a]Electronic mail: labanda@princeton.wireline.slb.com
[b]Electronic mail: s-barnett@nwu.edu

The surface topography and roughness were characterized using an atomic force microscope (AFM) operated in air. The roughness R was defined as the average difference in height between the five highest peaks and five lowest valleys, relative to the mean surface position, over a 500 nm profile. The roughness values given represent the average of four arbitrarily chosen azimuthal profile directions with the error bars indicating the scatter in the measurements. Surface root-mean-square roughness values from the AFM images were also obtained; these are two to three times smaller in magnitude with a nearly identical roughness dependence on the various bombardment conditions used in this study.

In situ measurements of surface composition were made using ion scattering spectrometry (ISS) with the same ion gun used for bombarding the surfaces. ISS data were obtained using He ions incident at $\phi=18°$, $E=1$ keV, $J=6.5\times10^{13}$ ions cm^{-2} s^{-1} and scattering angle $\theta=75°$. Scattered ions were energy analyzed using a 160° sector, 36.5-cm-radius electrostatic energy analyzer with 1-mm-diam entrance and exit apertures, and detected using a 10^7-gain dual microchannel plate detector. A typical spectrum was obtained using an ion dose of $\sim3\times10^{16}$ ions cm^{-2}. ISS spectra shown are normalized to the Si peak. An As peak is observed in all the scans indicating contamination from the MBE chamber, which is also used for GaAs growth.

Estimates of the defect densities after ion bombardment were obtained by chemical etching using the Secco formulation:[13] a 5 min etch in a 2:1 solution of concentrated HF and (0.15 M) $K_2Cr_2O_7$. Secco etch pits have been shown to form at dislocations in n-and p-type Si(001) substrates. Prior to the Secco etch, the samples were first dipped in concentrated (49%) HF acid and then rinsed with an overflow of deionized water. Etching was done with ultrasonic agitation to avoid surface roughening arising from bubble formation. After etching, the samples were rinsed with an overflow of deionized water and blown dry using high purity nitrogen. The surfaces of the etched samples were then observed using AFM.

XTEM observations were made using a Philips CM 20 UT microscope after sample preparation using a combination of mechanical and ion thinning, as described previously.[10]

III. EXPERIMENTAL RESULTS

A. Temperature effects

Initial experiments were carried out to determine an appropriate T_s value for the bombardment studies. Si substrates were annealed for 10 min at 450, 620, and 730 °C and then bombarded at the same T_s with Ar ions incident at $\phi=15°$, $E=1.0$ keV and $D=3\times10^{15}$ ions cm^{-2}. The sample annealed and bombarded at 450 °C gave very weak RHEED patterns with high background as shown in Fig. 1(a). A more distinct and streaky pattern appears on the sample bombarded at 620 °C [Fig 1(b)]. Samples bombarded at $T_s=730$ °C gave streaky RHEED patterns showing 2×1 reconstruction, strong Kikuchi lines, and higher-order spots indicating reasonably flat and well-ordered surfaces [Fig. 1(c)]. Samples annealed in vacuum at $T_s=730$ °C without any ion bombardment showed very weak RHEED patterns. This agrees with prior studies showing that annealing at this temperature does not provide clean surfaces.[14]

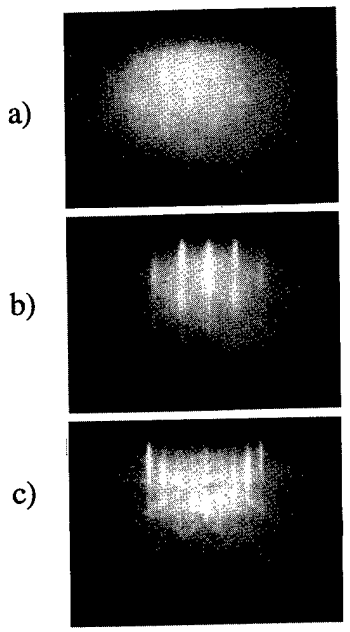


FIG. 1. RHEED patterns from Si(001) bombarded with 1 keV Ar ions at substrate temperatures of (a) 450, (b) 620, and (c) 730 °C.

ISS spectra were also observed after the above sputtering cleaning steps with the Si maintained at the same T_s. As shown in Fig. 2, ISS spectra obtained for the sample bombarded at $T_s=450$ °C showed a broad peak at the positions expected for C and O (note that the mass resolution was too low to resolve these). At $T_s=620$ °C, the C+O peak was smaller. At $T_s=730$ °C, the C and O peaks were barely resolved from the Si peak. A substrate temperature of $T_s=730$ °C was chosen for most of the experiments below, since flat, clean and well-ordered surfaces were obtained.

B. Ion bombardment parameter effects

Glancing-angle He ions produced distinct changes in RHEED patterns for Si(001) substrates after bombardment to a dose $D=3\times10^{15}$ ions cm^{-2} at $T_s=730$ °C and $\phi=15°$. Initially diffuse RHEED patterns evolved into elongated spots arranged in a bulk-like pattern, with some reconstruction evident. No higher-order Laue spots were observed. ISS scans did not show a C+O peak, indicating a clean but rough surface. Prolonged He bombardment to a dose $D>1.2\times10^{16}$ ions cm^{-2} was necessary to give RHEED patterns comparable to those obtained using Ar ion bombardment at 3×10^{15} ions cm^{-2} [see Fig. 1(c)]. These results are different than those for GaAs, where glancing-angle He ion bombardment yielded no observable cleaning over a wide range of parameters.[10]

Figure 3 shows RHEED patterns and AFM images from Si(001) surfaces bombarded by Ar ions with different doses, maintaining $E=1.0$ keV, $T_s=730$ °C and $\phi=15°$. Without ions, the surfaces showed hazy, featureless RHEED patterns

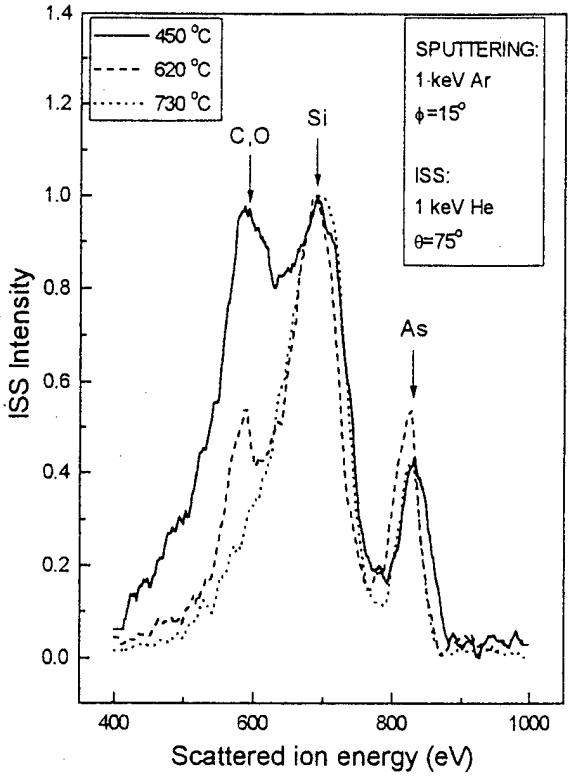


FIG. 2. ISS scans corresponding to the conditions given in Fig. 1 are shown. The As peak is present due to the background arsenic present in the chamber, which is also used for GaAs growth.

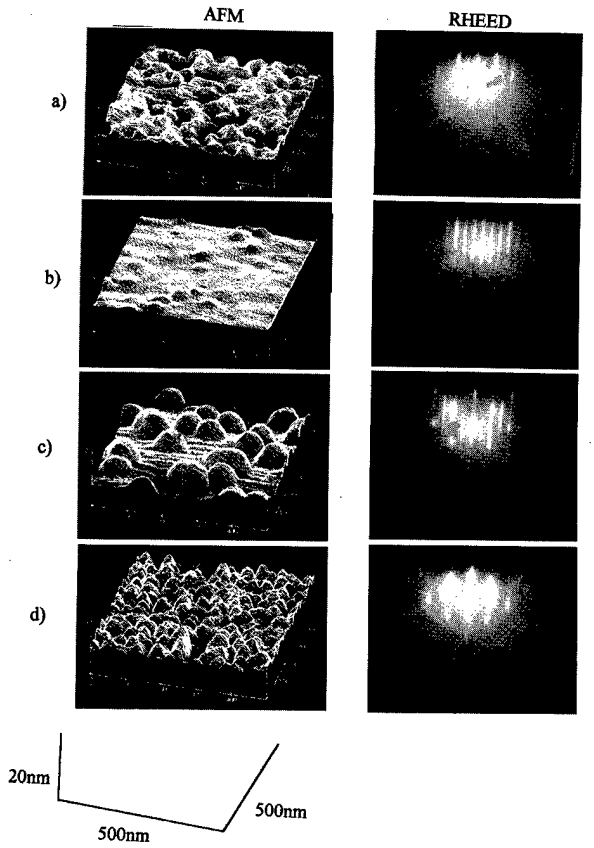


FIG. 3. RHEED patterns and AFM images from Si(001) bombarded with 1 keV Ar ions incident at $\phi=15°$ at doses of (a) 1, (b) 3, (c) 6, and (d) 12 $\times 10^{15}$ ions cm^{-2}.

but were relatively flat with $R=0.3$ nm. At the lowest D, 1.3×10^{15} ions cm^{-2}, the AFM roughness had increased to $R=2.7$ nm. RHEED patterns had streaks with bulk-like intensity modulations and indicated weak reconstruction. $D=3\times10^{15}$ ions cm^{-2} gave the best RHEED pattern, with 2×1 reconstruction streaks, higher-order diffraction spots, and strong Kikuchi lines. The surface was also relatively smooth ($R=0.5$ nm). Bombardment at higher D yielded spotty RHEED patterns, while the AFM images showed rougher surfaces with distinct hillock formation. The largest dose, $D=1.2\times10^{16}$ ions cm^{-2}, resulted in a surface with $R=5.3$ nm. ISS scans showed a negligibly-small C+O peak, even for the lowest D value.

Figure 4 shows the RHEED patterns and AFM images of Si(001) samples bombarded with different energies E at $D=3\times10^{15}$ ions cm^{-2}, $\phi=15°$, and $T_s=730$ °C. The sample bombarded at $E=500$ eV showed a RHEED pattern that had streaks with bulk-like intensity modulations and weak reconstruction streaks. The AFM image showed considerable surface roughness ($R=3.1$ nm). This surface was similar to the lowest dose sample in Fig. 3. $E=1.0$ keV gave the smoothest surface ($R=0.5$ nm) and highest-quality RHEED pattern. For E above 1.0 keV, there was little change in the RHEED pattern although a slight increase in R was observed. ISS scans did not show C and O peaks even at the lowest E.

Lower ϕ values, 7° and 3°, were also tested with $E=1000$ eV and $D=3\times10^{15}$ ions cm^{-2}. For these values, streaky RHEED patterns with 2×1 reconstruction and higher-order Laue spots (similar to the best patterns shown in Figs. 3 and 4) were obtained. Table I shows the roughness values for $\phi=3°$ and 15° versus D. R values were consistently lower for $\phi=3°$, with a minimum $R=0.3$ nm observed over a wider range of doses.

Finally, some samples were prepared in which the bombarded at room temperature, followed by a 730 ° anneal. The sample and sample-holder were first outgassed for 10 min at 730 °C. After cooling to room temperature, Ar bombardment was carried out with $E=1000$ eV, $D=3\times10^{15}$ ions cm^{-2}, and $\phi=3°$. Hazy, featureless RHEED patterns were observed after bombardment, indicating an amorphous surface. However, after annealing for 10 min at 730 °C, RHEED patterns emerged that were comparable to the best patterns in Figs. 3 and 4. Figure 5 shows a typical AFM image obtained after this procedure; the value $R=0.2$ nm obtained was the lowest observed in this study. This peak-to-peak roughness is comparable to the thickness of a single monolayer of Si, 0.135 nm, indicating that these surfaces were close to the flatness limit achievable for a vicinal surface.

C. Ion damage

Estimates of the defect densities of the bombarded samples were obtained from 50 μm×50 μm AFM images of the surfaces after Secco etching. Samples bombarded at T_s

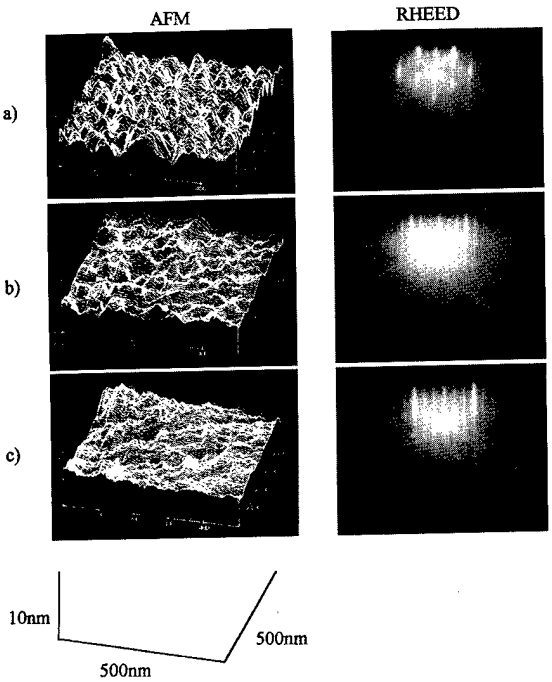


FIG. 4. RHEED patterns and AFM images from Si(001) bombarded with Ar ions incident at $\phi = 15°$ to a dose of 3×10^{15} ions cm^{-2} at ion energies of (a) 0.5, (b) 1.5, and (c) 2 keV.

$=730$ °C showed 0.5–2.0 μm diameter pits that were 20–50 nm deep with a density of 10^6-10^8 cm^{-2}. The pit density increased with increasing D, E, and ϕ, as indicated in Table II. Typical AFM images of etched surfaces are shown in Fig. 6. Figure 6(a) shows that no etch pits were observed on an unbombarded Secco-etched Si(001) surface. Decreasing the bombardment angle from $\phi = 15°$ to 3° [Figs. 6(b) and 6(c)] decreased the defect density from $\approx4\times10^7$ to $\approx4\times10^6$ cm^{-2}. Figure 6(d) shows that after ambient-temperature bombardment and annealing, no etch pits were present, indicating a defect density $<4\times10^4$ cm^{-2}.

Attempts were made to image the ion-induced defects using XTEM. Even with a heavily damaged sample (an etch pit density of $\approx6\times10^7$ cm^{-2}), no defects were detectable as shown in Fig. 7. This was due to the relatively small interface area sampled by XTEM.

TABLE I. Surface roughness of Si(001) bombarded with 1 keV ions at different doses and angles.

Ion dose (10^{15} cm^{-2})	AFM roughness (nm)	
	$\phi = 15°$	$\phi = 3°$
1	2.7±0.3	0.5±0.1
3	0.5±0.1	0.3±0.1
6	4.2±0.6	0.3±0.1
12	5.3±0.6	0.6±0.1

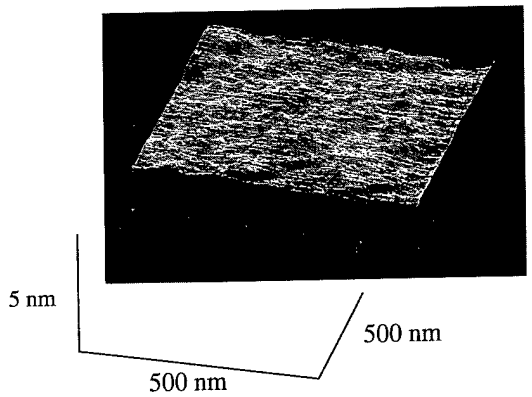


FIG. 5. AFM image of a Si(001) surface bombarded at ambient temperature using 1 keV Ar ions at 3° from the surface to a dose of 3×10^{15} ions cm^{-2}, followed by a 730 °C, 10 min anneal.

IV. DISCUSSION

The effects of glancing-angle ion bombardment of air-contaminated Si surfaces at elevated temperatures are qualitatively similar to those described previously for GaAs.[10] That is, the initially flat surfaces first roughened during ion bombardment, then smoothened, and finally roughened again (see Fig. 3). ISS results showed that C and O were effectively removed before the first roughness maximum. We believe that inhomogeneous removal of the contaminant layer causes the initial surface roughening. The smoothening effect of glancing-angle ions, previously demonstrated for GaAs and explained by selective sputtering from surface protrusions combined with surface diffusion,[9] causes a rapid decrease in roughness after the contaminant layer is removed. The roughening observed at high doses is attributed to preferential sputtering at the defects created by ion bombardment. A similar hillock morphology has previously been observed on Si(001) bombarded by normally incident 500 eV Ar ions at doses similar to those used here.[15] They attributed the hillocks to step-bunching at defects. Etch-pit densities approached 10^8 cm^{-2} for higher doses, about an order of magnitude less than the island densities that can be inferred from the high dose images in Fig. 3. As can be seen in Fig. 6(b), however, the etch pits are large (500–2000 nm) and closely spaced, such that there was likely considerable un-

TABLE II. Etch pit densities of Si(001) bombarded with different doses, energies, and angles. Typical etch pit diameter was 0.5–2 μm.

Ion dose (10^{15} cm^{-2})	Ion energy (eV)	ϕ (deg.)	Etch pit density (10^7 cm^2)
1	1000	15	1
3	1000	15	4
6	1000	15	6
12	1000	15	2[a]
3	500	15	0.4
3	1500	15	1
3	2000	15	2
3	1000	3	0.4

[a]Etch pit diameter was two to four times bigger than the norm. The low density may be due to pit coalescence.

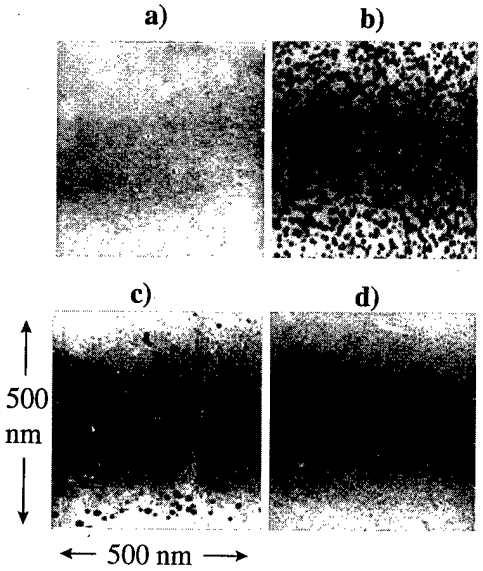


FIG. 6. AFM images of Secco-etched Si(001) that had been exposed to differing ion bombardment treatments. (a) shows an unbombarded surface, (b) shows a surface bombarded at $T_s=730$ °C, $E=1$ keV, $D=3\times10^{15}$ ions cm^{-2}, and $\phi=15°$, (c) shows the same as (b) but with $\phi=3°$, and (d) shows the same as (c) but at ambient temperature followed by a 730 °C anneal. White-to-black represents a depth change of 500 nm.

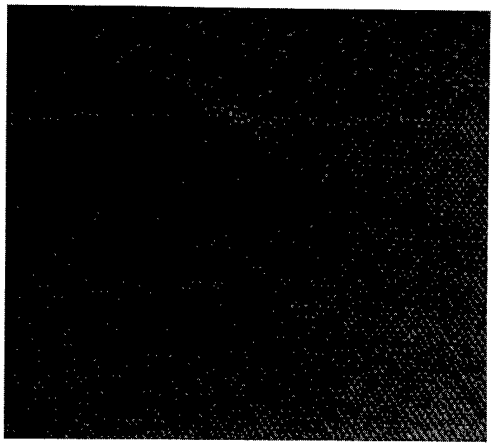


FIG. 7. Cross-sectional transmission electron microscope image of a Si(001) surface after ion bombardment at $T_s=730$ °C, $E=1$ keV, $D=1.2\times10^{16}$ ions cm^{-2}, and $\phi=15°$.

dercounting of defects at the higher densities. The numbers are therefore consistent with the idea that hillock formation was related to ion-induced defects.

The broader minimum in the roughness and lower R values for $\phi=3°$ vs 15° (see Table I) can be explained by the lower damage production rate at the lower angle (see Fig. 6), which should lead to a lower roughening rate. Lowering ϕ also maximizes the sputtering of protrusions, and minimizes sputtering at flat regions, which explains in part the lower roughnesses obtained at $\phi=3°$.

The effect of ion energy is analogous to the dose effect noted above. Decreasing E decreases the sputtering yield, causing a similar effect as decreasing the dose. This can be seen by noting that the surface obtained for $E=500$ eV and $D=3\times10^{15}$ ions cm^{-2} (Fig. 4) was nearly identical to that obtained with $E=1$ keV and $D=1.3\times10^{15}$ ions cm^{-2} (Fig. 3); in both cases, the bombardment was insufficient to smoothen the surface. It is not clear, however, why increasing E above 1 keV (Fig. 4) did not increase the roughness as rapidly as increasing D (Fig. 3), since both sputter yield and damage production rate increase with increasing E.

The dose required to clean and smoothen Si(001) was a factor of $\approx$ 10 times less than that required for GaAs(001),[10] for the same E and ϕ. Furthermore, Si(001) was effectively sputter cleaned with Ar ions at $\phi=3°$, which had little effect on GaAs(001). It seems unlikely that these differences are due to differences in the air-contamination layer on the two materials: the surface oxides typically have similar thicknesses (a few nm) and SiO_2 is more stable than arsenic and gallium oxides, suggesting that the sputtering rate should, if anything, be lower for Si. The difference may be a mass effect. That is, Ar (40 AMU) ions are heavier than Si (28 AMU), but lighter than Ga (70 AMU) or As (75 AMU). Thus, Ar ions are readily reflected from a GaAs surface in binary collisions, but will retain some forward momentum even in head-on collisions with Si atoms. It was also found that He ions provided significant cleaning of Si, but had little effect on GaAs. This is probably due to the better mass match of He with Si. For example, the maximum fraction of the energy that can be transferred from He to Si in a head-on collision is 0.44 vs 0.20 for He and Ga.

The substrate temperature during bombardment of Si(001) played a key role in determining surface quality. The presence of roughness for bombardment at $T_s=450$ °C is indicated by the spotty nature of the RHEED pattern in Fig. 1. The roughness likely originates from the surface crater formation that has been observed to occur upon energetic ion impact.[16] The streaky patterns at 730 °C indicate that this can be overcome by sufficient surface diffusion. The high background intensity for low T_s (Fig. 1), indicating poor crystalline perfection, was eliminated by annealing at higher T_s. The ISS data in Fig. 2 suggest that contamination was present after low-T_s sputtering, but we believe that this is an artifact; i.e., the surface was recontaminated during the ISS measurement. Indeed, if the initial surface-contamination layer were still present after low-T_s sputtering, then it would not have been possible to obtain good surfaces by ambient-temperature bombardment and annealing. Rather, we believe that the sputtering completely cleans the surface even at low T_s, and that any recontamination is removed by the 730 °C anneal.

Ion damage decreased with decreasing ion energy, dose, and impingement angle. The dependencies on E and D are not surprising. A decrease in damage with decreasing ϕ was previously shown for Ar ion bombardment of GaAs and explained by increased ion reflection, decreased channeling, and decreased projected damage range. The latter two factors both helped confine point defects very near the surface, enhancing their rate of self and surface annihilation.

The best results, both in terms of surface flatness and low defect density, were obtained by ion bombarding at ambient

temperature and then annealing at 730 °C. A similar procedure was employed by Bean *et al.* for Si, except that the Ar bombardment was at normal incidence.[1] They concluded that the ambient-temperature bombardment results in amorphization of a thin surface layer, which is then epitaxially regrown in the subsequent anneal. They observed a reduction in defect density compared with sputtering at elevated temperatures which is explained by a reduction in channeling, but the defects were not completely eliminated. The present results show a complete elimination of defects, at least to the present detection limits. This may be attributed to the glancing-angle geometry, which further reduced ion channeling. The lower roughness after ambient-temperature bombardment may also be due to the amorphous layer which should sputter more uniformly than the crystalline surfaces with defects present during elevated-temperature sputtering.

The present ion cleaning technique compares well with non-ion cleaning methods, including high temperature ($>\sim$1000 °C) thermal annealing,[14] UV-H_2/O_2 cleaning,[17,18] and recent variations of the standard HF dip technique.[19,20] It is undesirable to expose Si wafers to the temperatures used in thermal desorption cleaning. Furthermore, residual hydrocarbons form very stable SiC clusters at $T_s = 550$ °C that require temperatures >1200 °C to remove.[14] UV/oxygen[17] and UV/hydrogen[18] cleaning procedures remove adsorbed hydrocarbons from silicon surfaces. The UV/hydrogen method in particular has been successful for *in situ* cleaning as-received wafers without any prior wet chemical treatment. However, it is ineffective in removing chemisorbed carbon atoms. Both methods require close-proximity UV lamps and in the UV hydrogen method, hydrogen pressures of $\sim 10^{-3}$ Torr are required, which may not be compatible with many MBE systems. Also, both require ~900 °C final anneals to remove the oxide layer, higher than in the present technique. Finally, the present technique has practical advantages in terms of versatility, reproducibility, and the avoidance of wet chemical treatments and the associated waste disposal problems.

V. SUMMARY AND CONCLUSIONS

Si(001) prepared by ambient-temperature bombardment with 1 keV Ar ions to a dose of 3×10^{15} ions cm^{-2} at 3° from the surface plane, followed by annealing at 730 °C for 10 min, provided clean, defect-free surfaces with 0.2 nm peak-to-peak roughness. The lack of damage was explained by the formation of an amorphous layer, along with increased ion reflection, reduced ion channeling, and lower projected damage range for glancing ion incidence compared with normal incidence. The very smooth surfaces were explained by selective sputtering of surface protrusions. This technique allows *in situ* cleaning of as-received wafers without any prior wet chemical etching at a lower temperature than most other cleaning methods, and can easily be incorporated in most UHV systems.

Ion bombardment at elevated substrate temperatures yielded rougher surfaces and measurable defect densities. During bombardment, surfaces first roughened due to nonuniform removal of the oxide layer, then smoothened due to preferential sputtering of asperities, and then roughened gradually due to selective sputtering at defects. The amount of damage increased with increasing ion dose, energy, and impingement angle. He ion bombardment was less effective for sputter-cleaning Si(001), requiring doses ≈4 times higher to achieve results comparable to Ar bombarded surfaces.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of the National Science Foundation Grant No. DMR-9523553, the use of the facilities of the National Science Foundation Materials Research Laboratory at Northwestern University, and support from the Swedish Natural Science Foundation.

[1] J. C. Bean, G. E. Becker, P. M. Petroff, and T. E. Seidel, J. Appl. Phys. **48**, 907 (1977).
[2] K. T. Sung and S. W. Pang, J. Vac. Sci. Technol. A **12**, 1346 (1994).
[3] D. G. Armour and A. H. Al-Bayati, Nucl. Instrum. Methods Phys. Res. B **67**, 279 (1992).
[4] E. Chason, T. M. Mayer, and A. Payne, Appl. Phys. Lett. **60**, 2353 (1992).
[5] E. Chason and T. M. Mayer, Appl. Phys. Lett. **62**, 363 (1993).
[6] E. A. Eklund, R. Bruinsma, and J. Rudnick, Phys. Rev. Lett. **67**, 1759 (1991).
[7] L. M. Garverick, J. H. Comfort, T. R. Uyeh, R. Reif, F. A. Baiocchi, and H. S. Luftman, J. Appl. Phys. **62**, 3398 (1987).
[8] T. Shibata and T. Ohmi, J. Electron. Mater. **19**, 1065 (1990).
[9] J. G. C. Labanda, S. A. Barnett, and L. Hultman, Appl. Phys. Lett. **66**, 3114 (1995).
[10] J. G. C. Labanda, S. A. Barnett, and L. Hultman, J. Vac. Sci. Technol. B **13**, 2260 (1995), and references therein.
[11] C.-H. Choi, R. Ai, and S. A. Barnett, Phys. Rev. Lett. **67**, 2829 (1991), and references therein.
[12] J. G. C. Labanda and S. A. Barnett, J. Vac. Sci. Technol. A **14**, 485 (1996).
[13] D. G. Schimmel, J. Electrochem. Soc. **123**, 734 (1976).
[14] R. C. Henderson, W. J. Polito, and J. Simpson, Appl. Phys. Lett. **16**, 15 (1970).
[15] M. R. Cohen, R. J. Simonson, M. M. Altamirano, K. L. Critchfield, W. T. Kemp, and J. A. Meinhardt, J. Vac. Sci. Technol. A **11**, 971 (1993).
[16] P. Bellon, S. Jay Chey, J. E. van Nostrand, M. Ghaly, D. G. Cahill, and R. S. Averback, Surf. Sci. **339**, 135 (1995).
[17] S. R. Kasi and M. Liehr, Appl. Phys. Lett. **57**, 2095 (1990).
[18] G. Lippert, H. J. Thieme, and H. J. Osten, J. Electrochem. Soc. **142**, 191 (1995).
[19] T.-H. Park, Y.-S. Ko, T.-E. Shim, J.-G. Lee, and Y.-K. Kim, J. Electrochem. Soc. **142**, 571 (1995).
[20] S. G. D. S. Fiho, C. M. Hasenack, L. C. Salay, and P. Mertens, J. Electrochem. Soc. **142**, 902 (1995).

Inductively coupled plasma etching of $(Ba,Sr)TiO_3$ thin films

H. M. Lee,[a] D. C. Kim, W. Jo, and K. Y. Kim
Division of Devices and Materials, LG Corporate Institute of Technology, Seoul 137-140, Korea

(Received 23 July 1997; accepted 17 April 1998)

Etching behaviors of $(Ba,Sr)TiO_3$ (BST) thin films were studied by an inductively coupled plasma (ICP). Using Cl_2 gas with Ar gas, etch rate of the BST thin films and their selectivity for SiO_2 hard masks were systematically investigated over a wide range of gas mixing ratio and ICP power. The films were etched at rates of 400 Å/min using 20% of Cl_2 flow and 80% of Ar flow under ambient pressure below 10 mTorr and at a reactive ion etch power of 150 W and at an ICP power of 700 W. Chemical analysis by Auger electron spectroscopy showed that postetch residues had Ba- and Sr-rich phases which were not found on the surface of the SiO_2 hard mask. It was found that the residues of the BST films can be removed with the addition of SF_6 gas in the etchant gas mixture.
© *1998 American Vacuum Society.* [S0734-211X(98)02404-4]

I. INTRODUCTION

Recently, much attention has been paid to $(Ba,Sr)TiO_3$ (BST) thin films due to its high permittivity for application of giga-bit level dynamic random access memories[1–3] and integrated decoupling capacitors.[4,5] For fabrication of BST thin films into integrated semiconductor devices, it is essential to develop a reliable dry-etching process.[6] Etching of BST thin films has been usually performed by reactive ion etching (RIE) process, but there are some drawbacks such as slow etch rates less than 250 Å/min, postetch residues, and physical damages after etching.[7,8]

Inductively coupled plasma (ICP) etching offers an attractive alternative dry-etching technique. The general belief is that ICP sources are efficient to generate the high density plasma and easier to scale up than electron cyclotron resonance sources, and economical in terms of cost and power requirements. At low pressures below 10 mTorr, ICP is expected to produce low damage while achieving a high etch rate, because the plasma diffuses from the generation region and drifts to the substrate at relatively low energy.[9] In this article, ICP etching behavior of BST thin films is reported using a Cl_2 based gas chemistry with Ar. Furthermore, it was demonstrated that addition of SF_6 gas in the etchant gas mixture shows a dramatic reduction of postetch residues. In addition, an etching mechanism of the BST thin film is proposed based on Auger electron spectroscopy (AES) analysis.

II. EXPERIMENT

BST thin films used in this study were prepared on *p*-Si (100) by rf-magnetron sputtering. Growth temperature of the BST thin films was 490 °C and other deposition conditions were summarized in Table I.[10] From x-ray diffraction and Rutherford backscattering spectrometry, the BST thin films were confirmed to be a single phase and stoichiometric. The crystalline nature of the films was found to be polycrystalline with a grain size less than 1000 Å.

An ICP reactor with an Al_2O_3 insulation encircled by an inductive coil has been used in this study.[11] The plasma generated by a rf generator connected to the substrate at 13.56 MHz is intensified by a magnetic field induced by the inductive coil tuned at 2 MHz. Cl_2 and SF_6 gases were chosen as etchants,[12] while Ar gas was incorporated to use its milling effect on oxide materials. Conventional photoresist could not be applied as a mask material because its etch rate is faster than that of a BST film and its properties change from the heat generated during the etching process. Instead of photoresist, SiO_2 masks were prepared by low pressure chemical vapor deposition. The SiO_2 masks were patterned by RIE technique with CHF_3/O_2 chemistry. Selectivity, a ratio of etch rate of BST thin films to that of SiO_2 masks, is an important parameter to evaluate the effectiveness of the masks. Surface morphology and cross-sectional profile were investigated by scanning electron microscopy (SEM). Compositions of surfaces and postetch residues were examined by AES.

III. RESULTS AND DISCUSSION

Figure 1 shows etch rate of BST thin films and selectivity as a function of % Cl_2 flow in Ar/Cl_2 chemistry. The total flow rate of the gases was maintained at 40 sccm. During the etching process, pressure of the chamber, rf power, and ICP power were fixed at 10 mTorr, 150 W, and 700 W, respectively. The etch rates of BST thin films in pure Ar and in pure Cl_2 were found to be 220 and 280 Å/min, respectively. The etch rate and the selectivity of the films show a maximal value when Cl_2 in the gas mixture was near 20%. The etch rate of the film at 20% of Cl_2 was about 400 Å/min, which was two times faster than that of the film etched in pure Ar ambient. This result implies reactive etching of BST films. Likewise, selectivity shows a peak around 20% of Cl_2. The maximal value of the selectivity was found to be about 0.5, indicating that it is necessary to deposit SiO_2 masks more than two times thicker than the thickness of the BST thin films to be etched.

Figure 2 shows etch rate of the BST thin films and selectivity as a function of ICP power. As rf power applied to the ICP antenna was raised, the etch rate of the films increased. On the other hand, selectivity for the SiO_2 mask shows no

[a]Electronic mail: hmlee@lgcit.com

TABLE I. Growth conditions of BST thin films on *p*-Si (100) by rf-magnetron sputtering.

Parameters	Values
Sputtering target	Sintered $Ba_{0.5}Sr_{0.5}TiO_3$ (4 in.)
Working pressure	20 mTorr
Ratio of flow rate ($Ar:O_2$)	4:1
Total flow rate ($Ar + O_2$)	50 sccm
Radio frequency-power density	2.4 W/cm^2
Substrate temperature	490 °C
Presputter time	20 min
Target-substrate distance	13 cm
Film thickness	1500

dependence on ICP power. This can be explained by the fact that the etch rate of the SiO_2 mask increases faster than that of the BST thin film. From the dependencies of etch rate on $\%Cl_2$ and selectivity on ICP power, it is suggested that the etching of the BST thin films is sputtering dominated reactive etching with Cl_2 gas.

In Fig. 3(a), a SEM micrograph shows etched surface of the BST thin film using 8 sccm of Cl_2 flow and 32 sccm of Ar flow under ambient pressure below 10 mTorr and at a RIE power of 150 W and at an ICP power of 700 W. Postetch residues whose size was about 1 μm were seen only on the etched surface of the film. The residues were not found on the surface of the SiO_2 hard mask. The film shows a rough surface, which might originate from etching process damage or redeposition of the etched materials. The etch rate under the above conditions was about 400 Å/min. The etched surface of BST thin film using 4 sccm of Cl_2 flow, 4 sccm of SF_6 flow, and 32 sccm of Ar flow is shown in Fig. 3(b). The ambient pressure and the rf powers were set at the same vales as those of Fig. 3(a). The figure shows few postetch residues and clean surface morphology. The root mean square surface (rms) roughness was found to be about 26 Å and the etch rate was about 150 Å/min.

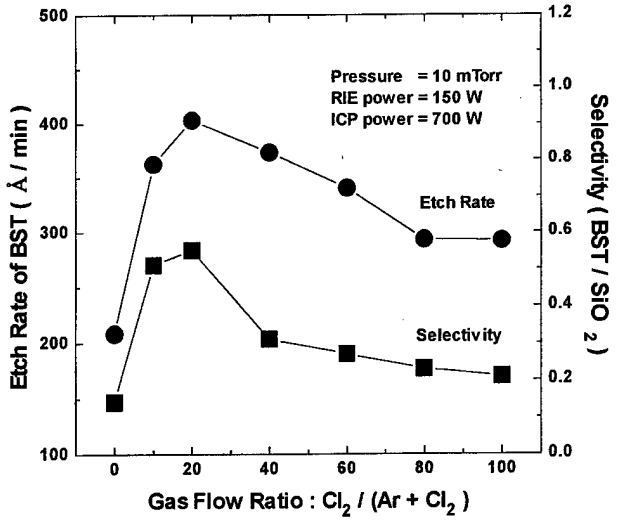


FIG. 1. Etch rate of BST thin films and selectivity as a function of % Cl_2 in an ICP-generated Ar/Cl_2 plasma.

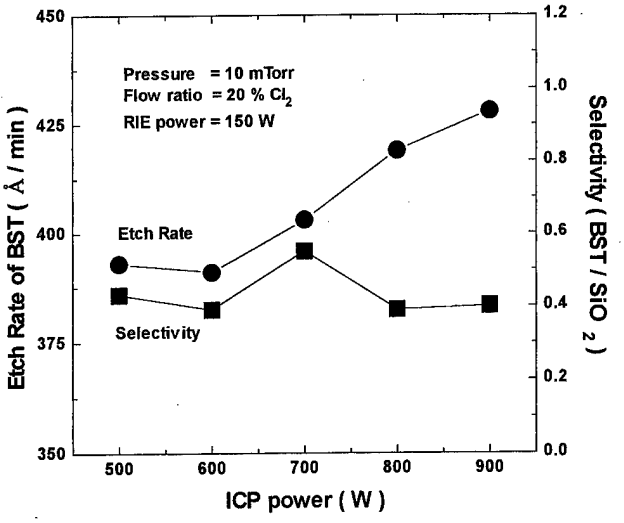
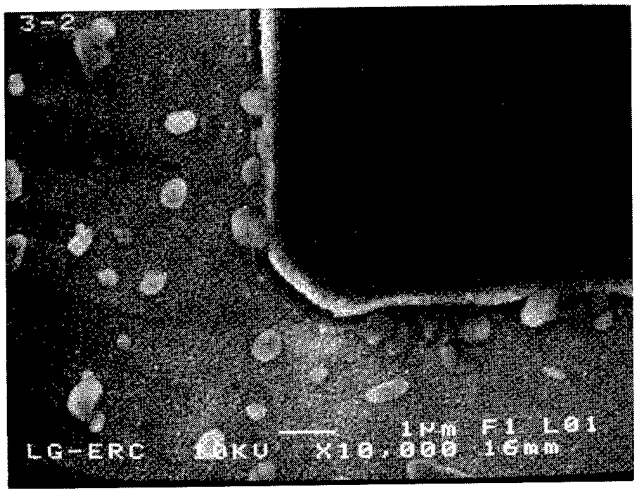
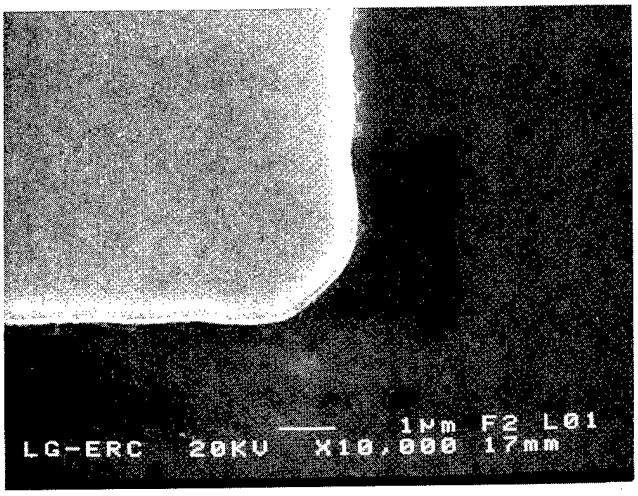


FIG. 2. Etch rate of BST thin films and selectivity as a function of ICP power in an ICP-generated Ar/Cl_2 plasma.



(a)



(b)

FIG. 3. SEM micrographs of the BST thin film etched (a) in Ar/Cl_2 chemistry and (b) in $Ar/Cl_2/SF_6$ chemistry.

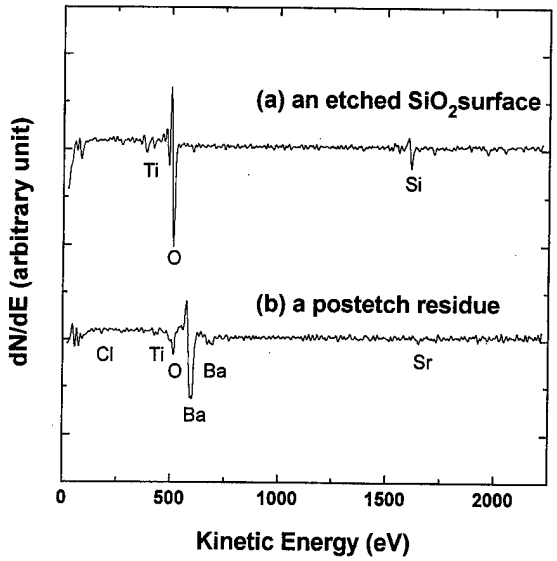


FIG. 4. AES spectra after ICP etching in Ar/Cl_2 plasma. The spectra for (a) an etched SiO_2 surface and (b) a postetch residue, which is indicated in Fig. 3(a).

To elucidate this phenomenon, chemical analysis by AES was performed. Figure 4 displays the AES spectra of an etched SiO_2 hard mask surface and a postetch residue after Ar/Cl_2 plasma treatment, which is found in Fig. 3(a). The spectra for the SiO_2 surface show Ti and Si peaks, denoting that there are Ti atoms or compounds from the BST film. On the other hand, the spectra for a residue show the barely detectable Cl and Ti signals and the much bigger Ba and Sr signals. This implies that the residues consist of mostly elemental Ba and some Sr and residual Cl from the gas phase. Considering that some compounds based on fluorine are more volatile than their chloride counterparts,[12] it can be explained that SF_6 added gas mixture is effective to reduce the postetch residues.

IV. CONCLUSION

ICP etching behaviors of the BST thin films were studied with Cl_2 based gas system as a function of plasma chemistry and ICP power. The chemistry dependence of the etch rate of the BST thin films suggested that the BST thin films should be etched through Ar ion assisted chemical reaction. The etch rate of BST thin films greater than 400 Å/min was measured at 20% Cl_2 and 80% Ar, ambient pressure of 10 mTorr, and a RIE power of 150 W and an ICP power of 700 W. Furthermore, it was found that by using the $Ar/Cl_2/SF_6$ gas mixture, the residues can be removed. It was determined that they consist mainly of Ba and Sr elements by AES analysis of the postetch residues.

[1]W. Y. Hsu, J. D. Luttmer, R. Tsu, S. Summerfelt, M. Bedekar, T. Tokumoto, and J. Nulman, Appl. Phys. Lett. **66**, 2975 (1995).
[2]C. S. Hwang, S. O. Park, H.-J. Cho, H.-K. Kang, S. I. Lee, and M. Y. Lee, Appl. Phys. Lett. **67**, 2819 (1995).
[3]P.-Y. Lesaicherre, H. Yamaguchi, Y. Miyasaka, H. Watanabe, H. Ono, and M. Yoshida, Integr. Ferroelectr. **8**, 201 (1995).
[4]D. Dimos, S. J. Lockwood, T. J. Garino, H. N. Al-Shareef, and R. W. Schwartz, Mater. Res. Soc. Symp. Proc. **433**, 305 (1996).
[5]W. Williamson III, B. K. Gilbert, H. D. Chen, K. R. Udayakumar, L. E. Cross, and C. M. Bozler, Integr. Ferroelectr. **10**, 335 (1995).
[6]H. Aoki, T. Hashimoto, E. Ikawa, T. Kikkawa, K. Takeuchi, S. Yamamichi, T. Sakuma, and Y. Miyasaka, Jpn. J. Appl. Phys., Part 1 **32**, 376 (1993).
[7]S. W. Pan, Solid State **27**, 249 (1984).
[8]D. L. Flamm, P. L. Cowan, and J. A. Golovchenko, J. Vac. Sci. Technol. **17**, 1341 (1980).
[9]R. J. Shul, G. B. McClellan, S. A. Casalnuovo, and D. J. Rieger, Appl. Phys. Lett. **69**, 1119 (1996).
[10]W. Jo, D. C. Kim, H. M. Lee, and K. Y. Kim, Mater. Res. Soc. Symp. Proc. **433**, 57 (1996).
[11]W. L. Johnson, in *High Density Plasma Source*, edited by O. A. Popov (Noyes, Park Ridge, NJ, 1995), p. 114.
[12]*CRC Handbook of Chemistry and Physics*, edited by R. C. Weast (CRC, Boca Rotan, 1988), p. B-48, 1st student ed.

Reactive ion etching of $Pb(Zr_xTi_{1-x})O_3$ thin films in an inductively coupled plasma

Chee Won Chung[a)]
Electronic Materials Laboratory, Materials and Device Sector, Samsung Advanced Institute of Technology, Suwon 440-600, Korea

(Received 21 August 1997; accepted 22 May 1998)

Reactive ion etching of $PbZr_xTi_{1-x}O_3$ (PZT) thin films was studied by using $Cl_2/C_2F_6/Ar$ gas chemistry in an inductively coupled plasma (ICP). PZT films were deposited on $Pt/Ti/SiO_2/Si$ substrates by the sol–gel process. PZT films were etched by varying the etching parameters including coil rf power, dc-bias voltage to the substrate, and gas pressure. Etching characteristics of PZT films were investigated in terms of etch rate, etch selectivity, etch profile and etching mechanism. Etch profile along with etch anisotropy was observed as a function of etching parameter by field emission scanning electron microscopy. For understanding of the etching mechanism, x-ray photoelectron spectroscopy and ICP analysis for film composition were utilized. Finally, the pattern transfer of PZT films with fine geometry was successfully achieved at the optimum etching condition. © *1998 American Vacuum Society.* [S0734-211X(98)09404-9]

I. INTRODUCTION

Recently, ferroelectric thin films have received great attention for the application to high density memory devices[1] such as dynamic and nonvolatile random access memory devices, infrared sensors,[2] electro-optical devices,[3] etc. In particular, much research of $PbZr_xTi_{1-x}O_3$ (PZT) and PLZT (La-doped PZT) ferroelectric films as a dielectric material for storage capacitors of highly integrated memory devices has been carried out since these films have a high dielectric constant and remanant polarization. Etching processes for both ferroelectric films and electrode materials of the capacitors must be developed in order to accomplish integration of these devices. Various etching processes including wet etching,[4] ion beam etching,[5] reactive ion beam etching,[6] plasma etching,[7] and reactive ion etching (RIE)[8–10] have been studied to define the patterns on ferroelectric materials such as PZT and PLZT thin films. Conventional RIE may not be applied to etch these materials with photoresist masks due to very poor selectivity to photoresist films. Therefore, high density plasma systems such as inductively coupled plasma (ICP),[11] transformerly coupled plasma (TCP), and electron cyclotron resonance (ECR)[12] should be employed. Little study on the etching characteristics of PZT films including etch profile and the etching mechanism has been done with high density plasma.

In this study, reactive ion etching of PZT films with $Cl_2/C_2F_6/Ar$ gas combination was studied by using high density ICP. An etching study was performed as a function of etching parameters including coil rf power, dc-bias voltage to the substrate, and gas pressure. Etching characteristics have been investigated in terms of etch rate, etch selectivity, etch profile and etch anisotropy. Understanding of the etching mechanism of PZT films has been attempted by analyzing the surface compositions of PZT films etched at various etch times under the same etching condition. Finally, delineation of the fine patterns on PZT films was performed for application to high density memory devices.

II. EXPERIMENT

Reactive ion etching of PZT films was investigated by using a multiplex inductively coupled plasma system (STS Ltd., UK), which provides high density, high ion current plasma at low pressure. The schematic diagram of the ICP etcher is shown in Fig. 1. The ICP system with an optical end-point detector consists of a loadlock chamber, a process chamber, and a substrate susceptor assembly. The substrate susceptor is cooled by He through chilled deionized water. The coil, which was connected to a 13.56 MHz rf power supply, was wound around the ceramic chamber to generate a high density plasma. A rf-bias voltage induced by rf power at 13.56 MHz was capacitively coupled to the substrate susceptor to control ion energy.

$Pb(Zr_xTi_{1-x})O_3$ thin films of 2500 Å in thickness were deposited on Pt coated $Ti/SiO_2/Si$ substrates by the sol–gel process. Pt electrodes (2500 Å) were deposited by using dc magnetron sputtering and a Ti layer of 300 Å was used as an adhesion layer between the Pt and SiO_2 films. The sol–gel solution for the PZT films was prepared by using precursors of lead acetate-trihydrate, zirconium *n*-propoxide, and titanium isopropoxide. Acetic acid and *n*-propanol were used as solvents. The solution was spun on the substrates at 2500 rpm for 35 s. The coated PZT films were annealed at 650 °C for 30 min to form the PZT perovskite phase. The etching samples were patterned by using conventional photoresists with a thickness of 1.2 μm.

Dry etching of the PZT films was studied by varying the etching parameters including coil rf power (400–800 W), dc-bias voltage to substrate (100–400 V), and gas pressure (1–10 mTorr) with $Cl_2/C_2F_6/Ar$ gas chemistry. The gases containing chlorine and fluorine are very effective in making volatile compounds with PZT films, depending on the etching conditions. The photoresist masks after etching were re-

[a)]Electronic mail: cwchung@saitgw.sait.samsung.co.kr

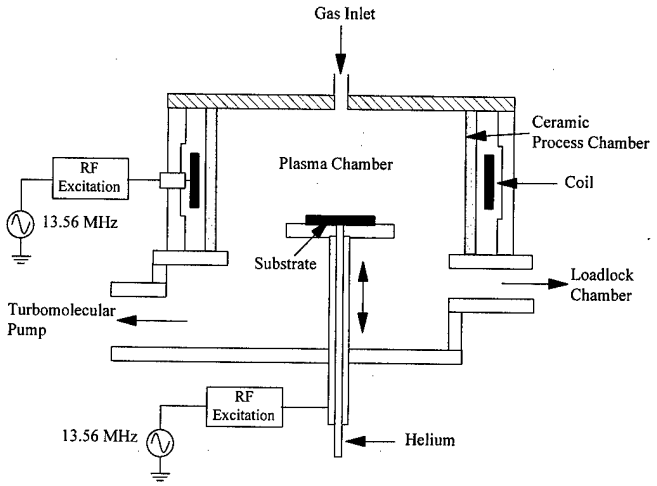


FIG. 1. Schematic diagram of the ICP etcher.

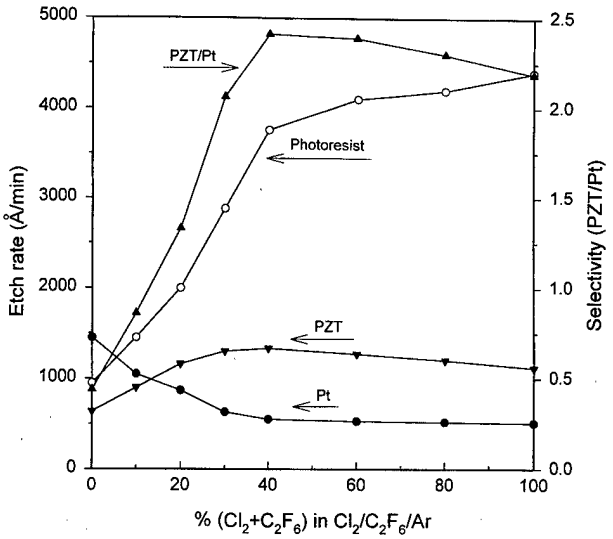


FIG. 2. Effect of the concentration of an etch gas on the etch rates of PZT and Pt films. Total flow rate: 30 sccm, coil rf power: 600 W, dc-bias voltage to substrate: 300 V, and gas pressure: 5 mTorr.

moved by both oxygen plasma ashing and wet stripping. Etching of Pt films was also carried out under the same etching conditions as in PZT etching to examine the selectivity of the PZT film to the Pt film. First, the composition of the etch gas (Cl_2/C_2F_6/Ar) was varied to find the optimum gas concentration of giving a high etch rate of the PZT film and a high selectivity of PZT to Pt. The ratio of Cl_2 to C_2F_6 was fixed at 9:1 since the chlorine compounds produced from the etching process were more volatile than the fluorine compounds[13] and the fluorine gas could retard the erosion of the photoresist by the chlorine gas.

The etch rates were measured by using a Dektak surface profilometer, and the etch profiles and etch slopes were observed by using field emission scanning electron microscopy (FESEM) as a function of the etching parameter. X-ray photoelectron spectroscopy (XPS) and ICP analysis for film composition were employed for investigation of the etching mechanism of PZT films.

III. RESULTS AND DISCUSSION

Reactive ion etching of PZT thin films was systematically investigated as a function of coil rf power, dc-bias voltage to substrate, and gas pressure for the characterization of the etching process in an ICP system. The effect of the concentration of an etch gas on the etch rates of PZT and Pt films is shown in Fig. 2. The etch gas used in this study was a mixture of Cl_2, C_2F_6, and Ar. When only Ar gas is used as an etch gas, the etch rate of Pt is much higher than that of PZT. This confirms that sputtering by Ar ions is more effective on Pt film than PZT film. As the concentration of (Cl_2+C_2F_6) gas increases up to about 40%, the etch rate of PZT increases. It indicates that there is a chemical enhancement in the etching of the PZT film. However, as the amount of (Cl_2+C_2F_6) gas exceeds over 40%, the etch rate of PZT begins to decrease. This seems to imply that a certain amount of physical bombardment by Ar ions is required to enhance the etch rate of PZT despite the chemical effect by an increase of (Cl_2+C_2F_6) gas.

In the case of Pt etching, as the concentration of (Cl_2+C_2F_6) gas increases, the etch rate of Pt decreases rapidly and becomes constant. This may be due to the fact that thin polymer films from the etch gas are formed on the Pt surface so that they inhibit sputtering of Ar ions to Pt film and/or Ar sputtering is more effective than the sputtering by (Cl_2+C_2F_6) gas.[14] In addition, it is clear that there is no chemical enhancement in the etching of Pt. As can be seen in Fig. 2, the high etch rate of the PZT film was obtained at this relatively mild condition and a high selectivity of PZT to Pt could also be achieved. Therefore, with consideration of the selectivity of PZT to the photoresist, all experiments for etching of PZT films were carried out with an etch gas of 30% (Cl_2+C_2F_6) in Cl_2/C_2F_6/Ar.

Figure 3 shows the effect of coil rf power on the etch rates

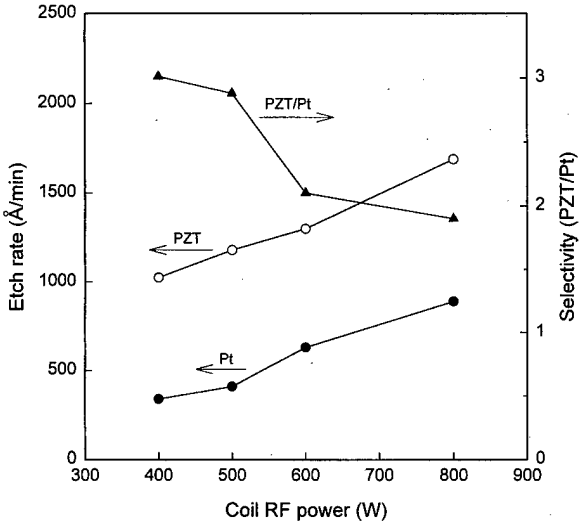


FIG. 3. Etch rates of PZT and Pt films as a function of coil rf power. Total flow rate: 30 sccm, etching gas: 30% (Cl_2+C_2F_6) in Cl_2/C_2F_6/Ar, dc-bias voltage to substrate: 300 V, and gas pressure: 5 mTorr.

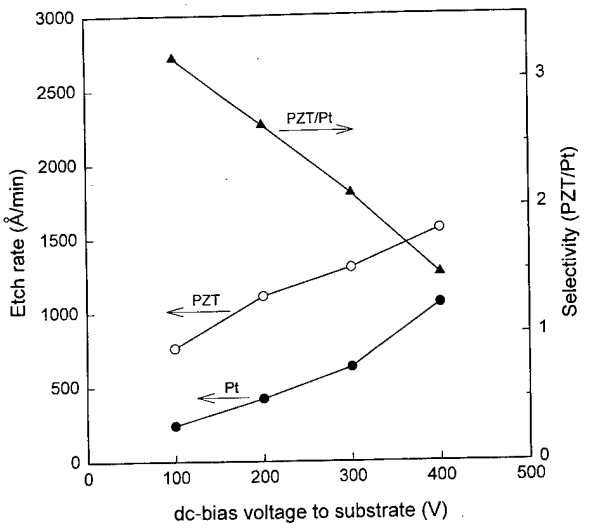


FIG. 4. Etch rates of PZT and Pt films as a function of dc-bias voltage to the substrate. Total flow rate: 30 sccm, etching gas: 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$, coil rf power: 600 W, and gas pressure: 5 mTorr.

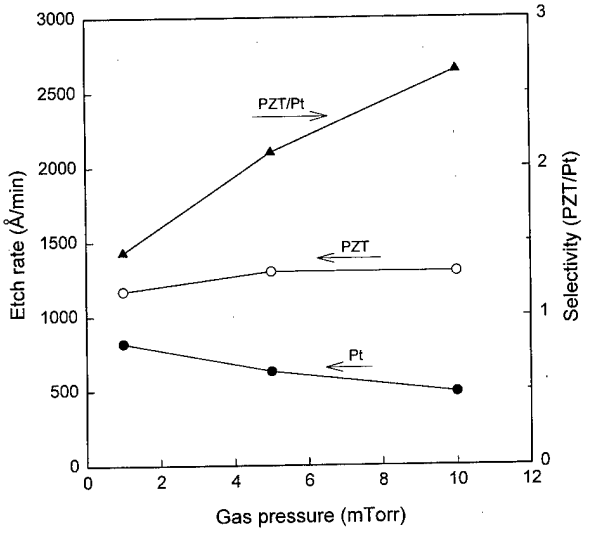


FIG. 5. Etch rates of PZT and Pt films as a function of gas pressure. Total flow rate: 30 sccm, etching gas: 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$, coil rf power: 600 W, and dc-bias voltage to the substrate: 300 V.

of PZT and Pt films under 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$. As coil rf power increases, the etch rates of both films increase almost linearly but the selectivity of PZT to Pt decreases. With increasing coil rf power, the plasma density increases so that the increased reactive free radicals and ions enhance the etch rates of both films. The etch rates of PZT and Pt films are shown in Fig. 4 as a function of dc-bias voltage. As dc-bias voltage increases, the etch rates of both

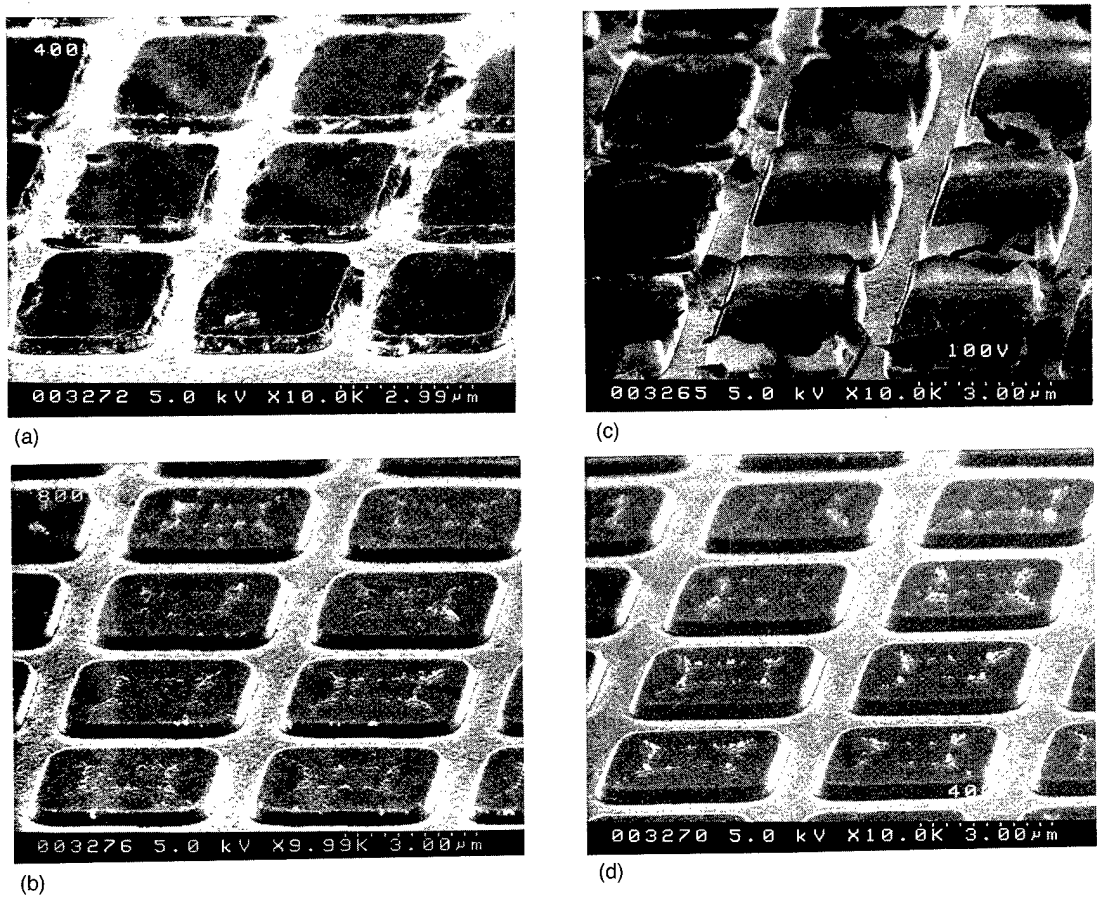
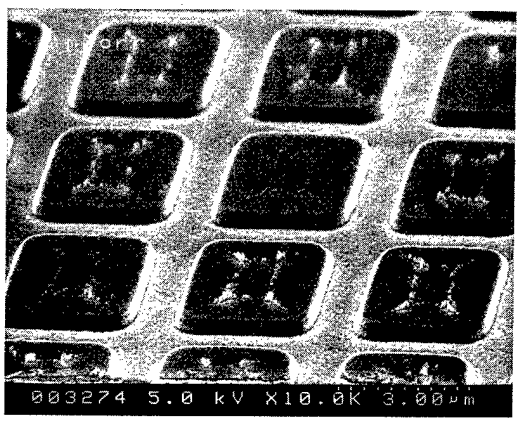
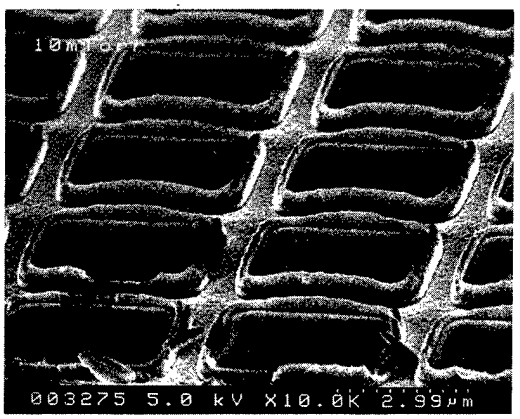


FIG. 6. FESEM photographs of etched PZT films as a function of etching parameter. Total flow rate: 30 sccm and etching gas: 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$. (1) coil rf power variation: (a) 400 W and (b) 800 W; dc-bias voltage to the substrate: 300 V and gas pressure: 5 mTorr. (2) (dc-bias voltage to the substrate) variation: (c) 100 V and (d) 400 V; coil rf power: 600 W and gas pressure: 5 mTorr. (3) Gas pressure variation: (e) 1 mTorr and (f) 10 mTorr; coil rf power: 600 W and dc-bias voltage to the substrate: 300 V.



(e)

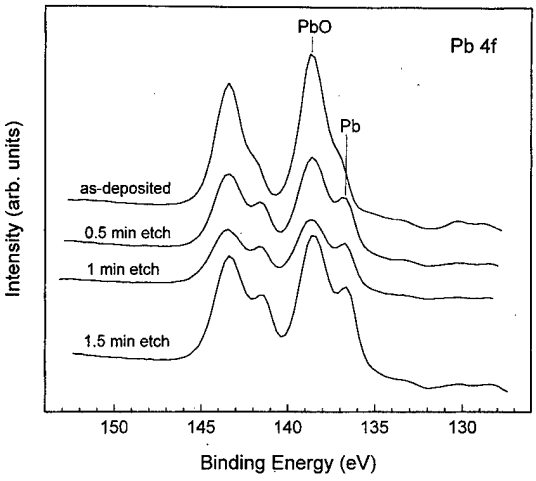


(f)

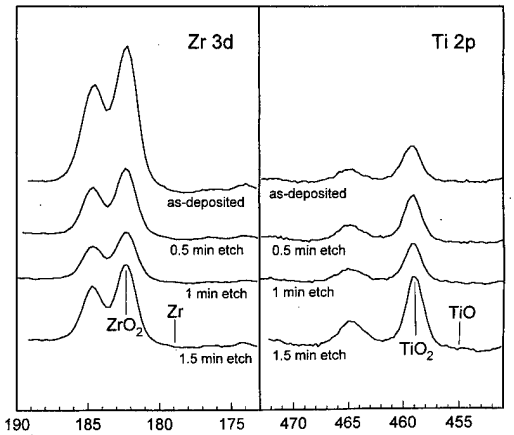
FIG. 6. (*Continued.*)

films increase but the selectivity of PZT to Pt rapidly decreases. With increasing dc-bias voltage, the bombarding ion energy is considered to increase, resulting in decreasing the selectivity of PZT to Pt since the ion bombardment to the substrate has more influence on the etching of the Pt film than the PZT etching. The effect of gas pressure on the etch rate is shown in Fig. 5. As gas pressure increases, the etch rate of Pt decreases while the etch rate of PZT seems to be constant so that the selectivity of PZT to Pt increases. With variation of gas pressure, the plasma density changes depending on the gas under study.[15] In case of 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$ as an etching gas, the plasma density is thought to increase with increasing gas pressure.[15] The Pt etching process, however, becomes ineffective (slow etching) in spite of the increase of the plasma density as the gas pressure increases. This is believed to be due to ineffective removal of sputtered Pt atoms by collision with incoming Ar ions and/or less sputtering of Ar ions onto Pt films with increasing pressure. As can be seen in Figs. 3–5, the high selectivity of PZT to Pt can be obtained at low coil rf power, low dc-bias voltage, and high gas pressure.

Figure 6 shows the FESEM photographs of etched PZT thin films with photoresist stripped off as a function of coil rf



(a)



(b)

FIG. 8. XPS spectra of PZT films etched at various etch times. (a) Pb 4*f*, and (b) Zr 3*d* and Ti 2*p*. Total flow rate: 30 sccm, etching gas: 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$, coil rf power: 800 W, dc-bias voltage to the substrate: 300 V, and gas pressure: 1 mTorr.

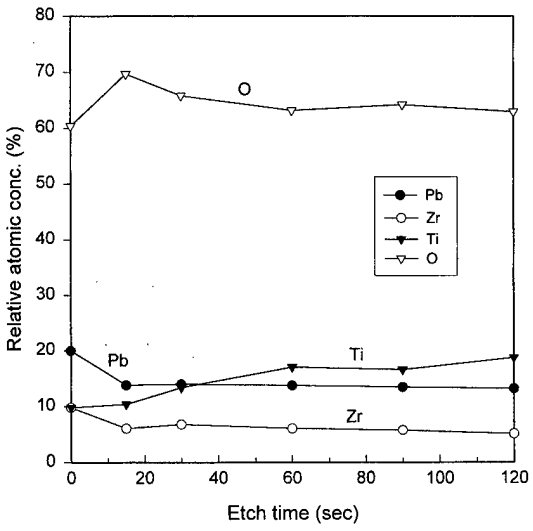


FIG. 7. Composition change of the PZT film surface as a function of etch time determined form XPS spectra. Total flow rate: 30 sccm, etching gas: 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$, coil rf power: 800 W and dc-bias voltage to the substrate: 300 V, and gas pressure: 1 mTorr.

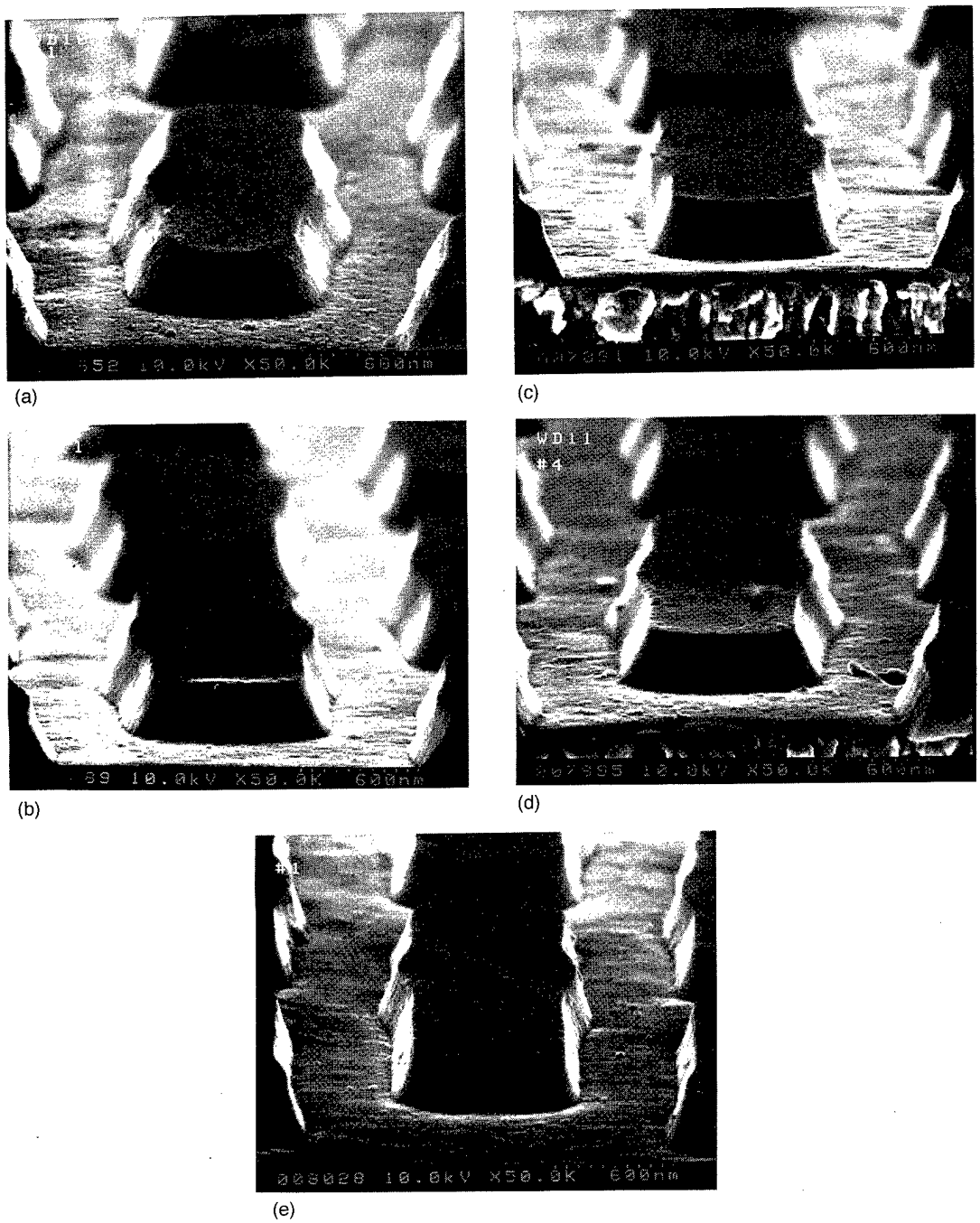


FIG. 9. FESEM photographs of etched PZT films as a function of etching parameter. Total flow rate: 30 sccm and etching gas: 30% ($Cl_2+C_2F_6$) in Cl_2/C_2F_6/Ar. (a) Coil rf power: 600 W, dc-bias voltage to the substrate: 300 V, and gas pressure: 5 mTorr. (b) Coil rf power: 800 W, dc-bias voltage to the substrate: 300 V, and gas pressure: 5 mTorr. (c) Coil rf power: 600 W, dc-bias voltage to the substrate: 300 V, and gas pressure: 1 mTorr. (d) Coil rf power: 600 W, dc-bias voltage to the substrate: 400 V, and gas pressure: 5 mTorr. (e) Coil rf power: 800 W, dc-bias voltage to the substrate: 300 V, and gas pressure: 1 mTorr.

power, dc-bias voltage to the substrate, and gas pressure. The PZT films [Figs. 6(a) and 6(c)] etched at low coil rf power (400 W) and low dc-bias voltage (100 V) have etch residues and particles along the etch patterns while the PZT films [Figs. 6(b) and 6(d)] etched at high coil rf power (800 W) and high dc-bias voltage (400 V) are clean without any residue. The case etched at low dc-bias voltage (100 V) has produced thicker polymer films along the patterns than that etched at low coil rf power (400 W). Therefore, it is obvious that dc-bias voltage to the substrate has a greater influence on the etch profile than coil rf power. At a high gas pressure of 10 mTorr [Fig. 6(f)], polymer residues which should be removed for the next step were deposited on the sidewall of the films while at low gas pressure of 1 mTorr [Fig. 6(e)], clean etch profiles were obtained. From Fig. 6, the etching conditions (high coil rf power and high dc-bias voltage), which

provide high density plasma and high ion energy, result in good etch profiles. However, low gas pressure is preferred to obtain clean etch profiles relative to high gas pressure due to the effective ion bombardment onto the substrates at low gas pressure. As a result, clean etch profiles can be achieved at high coil rf power, high dc-bias voltage, and low gas pressure. One can notice that these etching conditions are inconsistent with those for obtaining the high selectivity of PZT to Pt. From etch rate, etch selectivity, and etch profile, the optimum etching conditions which provide good etch profiles and high selectivity of PZT to Pt are found to be a coil rf power of 800 W, dc-bias voltage of 300 V and gas pressure of 1 mTorr with an etch gas of 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$.

Since the PZT film consists of three components (PbO, ZrO_2 and TiO_2) of the PZT solid solution, it is meaningful to compare the relative compositions of each component upon etching for an understanding of the etching mechanism. Figure 7 shows the composition changes of PZT film surface as a function of etch time. The etch samples for this study were bare PZT films without photoresist masks. The etch conditions used here gave clean etch profiles without any etch residues so that the measured compositions were those of etched PZT film surfaces. The relative atomic concentrations of Pb, Zr and Ti elements were determined from XPS spectra. The sensitivity factors of each element for XPS analysis were calibrated with the film composition obtained by the ICP analysis of as-deposited PZT film. All samples were scanned after 1 min sputtering in order to remove the effect of C contamination from the film surface. The main etching products in this study are expected to be chlorine compounds (e.g., $PbCl_2$, $ZrCl_4$, $TiCl_4$) since most of the chemical etching gas is chlorine, and chlorine compounds are more volatile than fluorine compounds.[13] The relative atomic concentration of Pb decreases rapidly at an initial stage of etching and seems to slightly decrease as the etching proceeds. This is because PbO in PZT is easily attacked by ion bombardment in plasma and transformed to Pb, relative to ZrO_2 and TiO_2 in PZT. The variation of Pb $4f$ peaks of etched PZT surfaces is shown in Fig. 8(a) for various etch times. It is evident from Fig. 8 that the chemical bond of PbO is broken to react with Cl gas and the peak of metal Pb in a Pb $4f$ peak begins to appear upon etching, decreasing the Pb content faster than Zr and Ti. The relative Zr concentration also exhibits a tendency of decreasing with increasing etch time while the relative Ti concentration slowly increases (Fig. 7). In comparison with Pb $4f$ peaks little changes in Zr $3d$ and Ti $2p$ peaks were found as shown in Fig. 8(b). From comparison of the relative atomic concentrations of Pb, Zr and Ti on the etched PZT surface, the relative Ti concentration increases while the Pb, Zr concentrations decrease upon etching. This means that the etch rate of TiO_2 component in PZT film is slowest. Therefore, it can be thought that the etching of TiO_2 in PZT is the rate limiting step in this etching study.

Figure 9 shows the FESEM photographs of the etched pattern with a line of 0.7 μm as a function of etching parameter, that is, by varying only one parameter from the standard etching condition for comparison of the etch slopes. The degree of anisotropy of the etching process can be obtained from the slope of the etched pattern. The pattern profile etched at the standard etching condition of this study (coil rf power of 600 W, dc-bias voltage of 300 V, gas pressure of 5 mTorr) is shown in Fig. 9(a). The etch slope of the pattern in Fig. 9(a) is about 62°. Figure 9(b) shows the pattern profile etched at a coil rf power of 800 W, keeping the other parameters the same as the standard condition. Figures 9(c) and 9(d) show the etch profiles etched by varying gas pressure (1 mTorr) and dc-bias voltage to the susceptor (400 V), respectively, from the standard etching condition. From the FESEM photographs, the etch slopes of the patterns etched at a coil rf power of 800 W [Fig. 9(b)] and dc-bias voltage of 400 V [Fig. 9(d)] are about 65° while the etch slope from the gas pressure of 1 mTorr [Fig. 9(c)] is about 70°. It is found from Fig. 9 that gas pressure is a more important factor in improving the degree of anisotropy of PZT etching than coil rf power and dc-bias voltage. It is considered that this is due to (1) a broadening of the ion energy distribution with increasing pressure[16] (2) and an increase of residence time and collisional scattering of the plasma species at high pressure.[17] Finally, the pattern transfer of PZT films with the 0.7 μm line was achieved with a high degree of anisotropy (etch slope of about 78°) at the optimum etching condition (coil rf power of 800 W, dc-bias voltage of 300 V and gas pressure of 1 mTorr) as shown in Fig. 9(e).

IV. CONCLUSIONS

Reactive ion etching of $PbZr_xTi_{1-x}O_3$ thin films was studied by using a $Cl_2/C_2F_6/Ar$ gas mixture in an inductively coupled plasma. A high etch rate of PZT films (>1300 Å /min) and good selectivity of PZT to Pt could be obtained with an etch gas of 30% $(Cl_2+C_2F_6)$ in $Cl_2/C_2F_6/Ar$ by using photoresist masks. While chemical enhancement was found in etching of PZT films, it appeared that ion bombardment was responsible for Pt etching. It was observed that the etching conditions (high coil rf power, high dc-bias voltage to the substrate, and low gas pressure), which provide high density plasma and high ion energy, result in good etch profiles. The dc-bias voltage to substrate has a greater effect on the etch profile than the coil rf power and the gas pressure. It was found from XPS analysis that PbO in PZT film was readily broken and transformed to Pb relative to ZrO_2 and TiO_2. It resulted in faster consumption of PbO than ZrO_2 and TiO_2. Comparison of the relative atomic concentrations of Pb, Zr and Ti determined from XPS spectra suggests that the etching of TiO_2, in this study, is a rate limiting step in the etching of PZT films.

The degree of anisotropy of PZT etching was enhanced with increasing coil rf power and dc-bias voltage to the substrate, and decreasing gas pressure. Among the three etching parameters used in this study, decreasing gas pressure appears to be most effective for obtaining a high degree of anisotropy. Finally, a pattern with a fine geometry was successfully delineated with a high degree of anisotropy and good selectivity at the optimum etching condition.

ACKNOWLEDGMENTS

The author would like to thank Dr. I. K. Yoo, Dr. I. S. Chung, and the FOS team colleagues for their encouragement and useful discussions, and I. Y. Song, J. S. Lee, and J. H. Lee for their most helpful support of this work.

[1]L. Parker and A. Tasch, IEEE Circuits Devices Mag. **6**, 17 (1990).
[2]K. K. Deb, K. W. Bennett, P. S. Brody, and B. M. Melnick, Integr. Ferroelectr. **6**, 253 (1995).
[3]G. Teowee, J. M. Boulton, E. K. Franke, S. Motakef, T. P. Alexander, T. J. Bukowski, and D. R. Uhlmann, Integr. Ferroelectr. **15**, 281 (1997).
[4]S. Mancha, Ferroelectrics **135**, 131 (1992).
[5]T. Kawaguchi, H. Adachi, K. Setsune, O. Yamazaki, and K. Wasa, Appl. Opt. **23**, 2187 (1984).
[6]M. A. Title, L. M. Walpita, W. Chen, S. H. Lee, and W. Chang, Appl. Opt. **25**, 1509 (1986).
[7]M. R. Poor and C. B. Fleddermann, J. Appl. Phys. **70**, 3385 (1991).
[8]K. Saito, J. H. Choi, T. Fukuda, and M. Ohue, Jpn. J. Appl. Phys., Part 2 **31**, L1260 (1992).
[9]J. J. van Glabbeek, G. A. C. M. Spierings, M. J. E. Ulenaers, G. J. M. Dormans, and P. K. Larson, Mater. Res. Soc. Symp. Proc. **310**, 127 (1993).
[10]D. P. Vijay, S. B. Desu, and W. Pan, J. Electrochem. Soc. **140**, 2635 (1993).
[11]C. W. Chung, W. I. Lee, and J. K. Lee, Integr. Ferroelectr. **11**, 259 (1995).
[12]B. Charlet and K. E. Davies, Mater. Res. Soc. Symp. Proc. **310**, 363 (1993).
[13]D. R. Lide, *Handbook of Chemistry and Physics*, 71st ed. (CRC, Boston, MA, 1991), pp. 4-72–4-118.
[14]F. Fracassi, R. d'Agostino, and A. Cacucci, J. Vac. Sci. Technol. A **13**, 63 (1995).
[15]Y. Ra, S. G. Bradley, and C. H. Chen, J. Vac. Sci. Technol. A **12**, 1328 (1994).
[16]K. W. Whang, S. H. Lee, and H. J. Lee, J. Vac. Sci. Technol. A **10**, 1307 (1992).
[17]R. J. Shul, G. B. McClellan, R. D. Briggs, D. J. Rieger, S. J. Pearton, C. R. Abernathy, J. W. Lee, C. Constantine, and C. Barratt, J. Vac. Sci. Technol. A **15**, 633 (1997).

Structural study of $YSi_{1.7}$ layers formed by channeled ion beam synthesis

M. F. Wu[a] and Shude Yao
Department of Technical Physics, Peking University, Beijing, 100871, People's Republic of China and Instituut voor Kern- en Stralingsfysika, University of Leuven, B-3001 Leuven, Belgium

A. Vantomme,[b] S. Hogg, H. Pattyn, and G. Langouche
Instituut vorr Kern- en Stralingsfysika, University of Leuven, B-3001 Leuven, Belgium

Qinqing Yang and Qiming Wang
NIOE Laboratory, Institute of Semiconductors, Chinese Academy of Science, Beijing, 100083, People's Republic of China

(Received 30 January 1998; accepted 17 April 1998)

High quality $YSi_{1.7}$ layers (χ_{min} of Y is 3.5%) have been formed by 60 keV Y ion implantation in Si (111) substrates to a dose of $1.0\times10^{17}/cm^2$ at 450 °C using channeled ion beam synthesis (CIBS). It shows that, compared to the conventional nonchanneled ion beam synthesis, CIBS is beneficial in forming $YSi_{1.7}$ layers with better quality due to the lower defect density created in the implanted layer. Rutherford backscattering/channeling and x-ray diffraction have been used to study the structure and the strain of the $YSi_{1.7}$ layers. The perpendicular and parallel elastic strains of the $YSi_{1.7}$ epilayer are $e^{\perp}=-0.67\%\pm0.02\%$ and $e^{\parallel}=+1.04\%\pm0.08\%$. The phenomenon that a nearly zero mismatch of the $YSi_{1.7}$/Si (111) system results in a nonpseudomorphic epilayer with a rather large parallel strain relative to the Si substrate ($\epsilon^{\parallel}=+1.09\%$) is explained, and the model is further used to explain the elastic strain of epitaxial $ErSi_{1.7}$ and $GdSi_{1.7}$ rare-earth silicides.
© *1998 American Vacuum Society.* [S0734-211X(98)02904-7]

I. INTRODUCTION

Following the studies on transition metal silicides, the rare-earth (RE) silicides have attracted considerable attention because they can be grown epitaxially on Si (111) and form the lowest known Schottky barrier height (~0.3–0.4 eV) on *n*-type Si. In the meantime, the development of techniques such as molecular beam epitaxy and ion beam synthesis provides the possibilities to prepare the epitaxial RE silicide layers with much better quality, which is very important both in applications and in fundamental research. Yttrium silicide shares many of the characteristics of the RE silicides and is usually included as a RE silicide.[1,2] Moreover, its nearly zero mismatch with Si (111) at room temperature,[1] the occurrence of only one silicide phase during thermal annealing[3] and the good Y-oxide dielectric properties (dielectric constant ~14–17, which is much higher than the value of 3.9 for SiO_2)[4] have resulted in extensive studies[5–15] of $YSi_{1.7}$. As is the case for many RE silicides, $YSi_{1.7}$ has a hexagonal phase based on the AlB_2 structure with ~15% vacancies in the Si sublattice. The vacancies are either ordered or not ordered, depending on the annealing temperatures[12] and the growth conditions.[2,13]

Like all RE metals yttrium is very reactive, so the crystalline quality and the interface abruptness of the Y-silicide layer formed by conventional deposition techniques in a high vacuum system were usually unsatisfactory. Koleshko *et al.*[15] reported that a polycrystalline Y-silicide film of the AlB_2 structure type was formed after annealing at 673 K for 30 min. By using rapid electron beam heating, Knapp *et al.*[1,2] succeeded in epitaxial growth of $YSi_{1.7}$ films on Si (111), however the reported 26% χ_{min} value is rather high. They pointed out that further improvement of the quality of the epitaxial films can be achieved using substrate cleaning and film preparation under ultrahigh vacuum conditions. Y-silicide with very low χ_{min} (~3%) has been grown in ultrahigh vacuum using a template technique and postannealing at 900 °C,[5] however the Rutherford backscattering (RBS) spectra show that the quality of the interface is still rather poor, probably due to the contamination problems involved in the deposition techniques. Alternatively, ion beam synthesis (IBS) can effectively avoid the aforementioned problems. Ion beam synthesized $YSi_{1.7}$ layers with χ_{min} values of 11% (Ref. 6) and 7.1% (Ref. 7) have been reported. However, these results are not as good as IBS-formed $CoSi_2$ films which can reach a χ_{min} value as low as 3%.[16] One of the possible reasons for this is that, by using the conventional nonchanneled IBS, the heavier mass of the yttrium atoms (compared to Co) results in a heavier defect density leading to a much higher χ_{min} value ($\geqslant$75%) of the Y peak[6] compared to the χ_{min} value of ~50% of the Co peak[17] for the RBS spectra of the as-implanted samples. The higher damage level of the as-implanted samples will then influence the crystalline quality of the IBS-formed silicide layers after subsequent annealing.

We have shown that channeled ion beam synthesis (CIBS) is indispensable for ion beam synthesis of $ErSi_{1.7}$ and $GdSi_{1.7}$ layers at energies ~90 keV.[18–21] However, it is expected (and confirmed in the present work) that the much lower defect density produced by CIBS compared to the nonchanneled IBS is helpful to improve the crystalline quality of IBS-formed $YSi_{1.7}$ films, even if it is not indispensable.

Moreover, a study of the elastic strain of a $YSi_{1.7}$/Si (111)

[a]Electronic mail: Wu%KSF%FYS@vinip.cc.kuleven.ac.be
[b]Postdoctoral researcher FWO (Fund for Scientific Research, Flanders, Belgium).

heterostructure is of particular interest. In principle, a small lattice mismatch is an important prerequisite for the growth of an epilayer with good crystalline quality and abrupt interface. Besides, the strain of the epilayer is also related to the mismatch between the epilayer and the substrate. Therefore, it is worthwhile to study the strain in the $YSi_{1.7}$ layer on a Si (111) substrate which has a nearly zero mismatch at room temperature. To date, the only reported strain data is $e^{\parallel} = 0.4\%$ and $e^{\perp} = -0.1\%$ for a nonchanneled ion beam synthesized $YSi_{1.7}$ layer formed with 330 keV Y ions to a dose of $2\times10^{17}/cm^2$ at 450 °C, and $e^{\parallel} = 0.0\%$ for a thicker layer formed at a higher dose ($3.6\times10^{17}/cm^2$).[7] In this article, we present and explain the strain results of our CIBS-formed $YSi_{1.7}$ layers, and also the strain results of $ErSi_{1.7}$ and $GdSi_{1.7}$ layers.

II. EXPERIMENT

Yttrium chloride was used in a Nielsen plasma ion source to produce the Y ions. Si (111) substrates were implanted with 60 keV ^{89}Y atoms to a dose of $1.0\times10^{17}/cm^2$ at 450 °C. During implantation, the ion beam was directed within ±1.5° to the [111] axis of the Si substrate to keep the condition of channeled implantation. The as-implanted sample was subsequently annealed at 600 °C for 1 h and 1000 °C for 0.5 h in a vacuum furnace (10^{-7} Torr). A 1.57 MeV $^4He^+$ beam, produced by a 5SDH-2 pelletron, was used for the RBS/channeling measurements. Glancing exit geometry was used to improve the depth resolution. Off-normal axial angular scans were used to measure the angle, Φ_{epi}, between the [0001] and $\langle 1\bar{2}13\rangle$ axes which is related to the strain of the $YSi_{1.7}$ epilayer. A Rigaku double-crystal diffractometer with monochromatic Cu $K\alpha$ 1 x rays ($\lambda = 0.154\,05$ nm) was used to perform the symmetric and asymmetric x-ray $\theta-2\theta$ scans.

III. RESULTS AND DISCUSSION

A. Formation of $YSi_{1.7}$ layers

Figure 1 shows the random and aligned RBS spectra of the as-implanted sample. The arrows (labeled Y and Si) indicate the energy for backscattering from these elements at the surface. A simulation of the random spectrum revealed that the profile of Y atoms has a Gaussian-like shape with a total dose of $1.0\times10^{17}/cm^2$. The aligned spectrum in Fig. 1 shows a χ_{min} value of 16% indicating that the defect density in this sample is much lower than the value (≥75%) of the samples prepared by nonchanneled IBS.[6] The peak in the aligned spectrum at channel ~800 originates from the different orientation of the precipitates while the peak at channel ~830 is the damage peak. The x-ray symmetric θ-2θ scan ranging from 10° to 110° (not shown) of the as-implanted sample reveals that the hexagonal phase of Y silicide has already formed, and that it is the only phase in the as-implanted sample. This is quite different from the x-ray diffraction (XRD) result obtained from the samples prepared

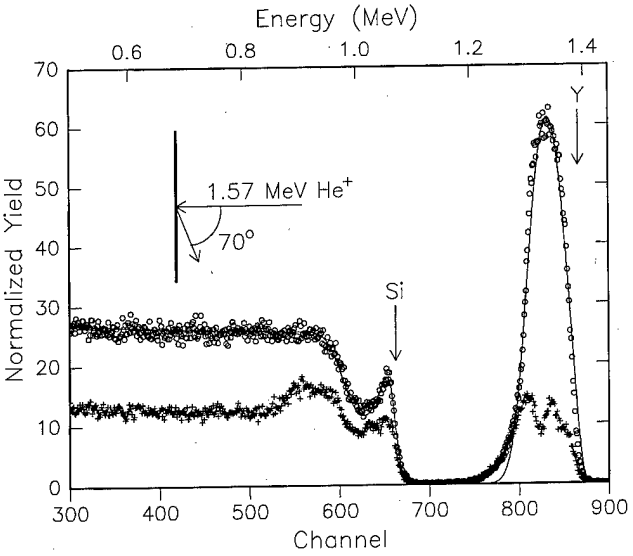


FIG. 1. Random (O), aligned (+) and simulated (solid line) spectra of an as-implanted sample implanted with 60 keV Y atoms to a dose of $1.0\times10^{17}/cm^2$ at 450 °C using channeled implantation. The inset shows the geometry used in the RBS measurements.

by Alford *et al.*[7] which showed no reflection peaks related to $YSi_{1.7}$, probably due to the very high χ_{min} value of their as-implanted samples.

The as-implanted sample was subsequently annealed at 600 °C for 1 h and 1000 °C for 0.5 h in a vacuum furnace. Figure 2 shows the RBS spectra of the annealed sample. The simulation (solid line) of the random spectrum reveals that a 48 nm buried $YSi_{1.7}$ layer with abrupt interfaces is formed underneath a 13 nm top Si layer. However, the side tails of the Y peak indicate that there are some yttrium silicide precipitates, mostly below the buried layer, due to the low diffusivity of the Y atoms. Similar phenomena also existed in $YSi_{1.7}$ samples prepared by nonchanneled IBS.[6,7] The aligned spectrum in Fig. 2 shows that the damage peak at

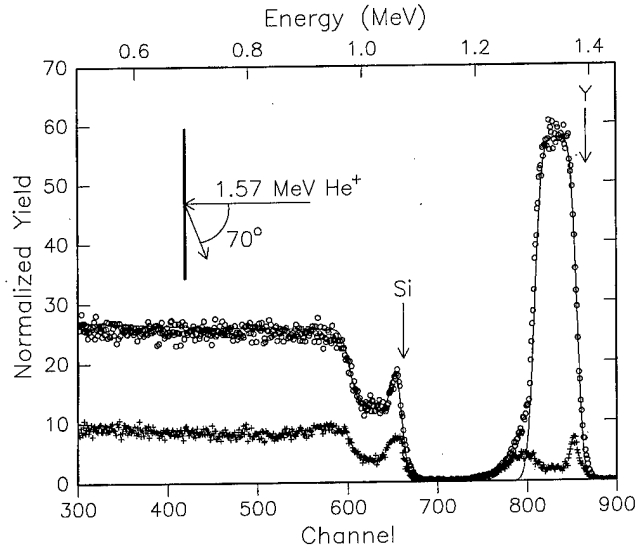


FIG. 2. Random (O), aligned (+) and simulated (solid line) spectra of the sample from Fig. 1 after annealing at 600 °C for 1 h + 1000 °C for 0.5 h.

channel ~830 is removed by annealing. The peak at channel ~800 which originates from the different orientation of the precipitates below the buried layer still exists, while the peak at channel ~850 is the interface peak which originates from misfit dislocations near the upper interface. Figure 2 further reveals that the χ_{min} value of the Y peak is only 3.5%, implying that the crystalline quality of this CIBS-formed $YSi_{1.7}$ layer is very good, which is obviously better than the best result ($\chi_{min}=7.1\%$) previously reported for the nonchanneled IBS-formed $YSi_{1.7}$ layers.[7] The reason for that is: the much lower close encounter probability of channeled implantation, as compared to the random implantation, results in a much lower defect density in the implanted layer.

B. Rotational angular scans around the Si [111] and $YSi_{1.7}$ [0001] axes

By tilting the sample 5.5° from the Si [111] axis and fixing three integration windows (channel numbers 300–530, 607–646 and 757–880) corresponding to the Si atoms in the substrate and the Si and Y atoms in the epilayer, respectively, we made an angular scan by rotating the sample in 80 steps with a step size of 1.5° around the normal of the sample, i.e., the scan is across many major planes which are perpendicular to the sample surface. Figure 3 shows three rotational angular scans corresponding to the Y and Si atoms of the epilayer and the Si atoms of the substrate. Combining the result from the angular scans around the Si [111] and $YSi_{1.7}$ [0001] axes (not shown) and x-ray symmetric $\theta-2\theta$ scan (see below) which confirms that the [0001] direction of the $YSi_{1.7}$ epilayer is aligned with the [111] direction of the Si substrate, we found that the result is quite similar to that of the $ErSi_{1.7}$/Si (111) system,[21] i.e., the azimuthal orientation of the $YSi_{1.7}$ epilayer is $YSi_{1.7}$ [0001] // Si [111] and $YSi_{1.7}\{11\bar{2}0\}$//Si {110}; $YSi_{1.7}\{10\bar{1}0\}$//Si {112} which can be further confirmed by the appearance of the $\langle 1\bar{2}13\rangle$ axis and the $\langle 210\rangle$ axis in the $YSi_{1.7}\{10\bar{1}0\}$ plane and Si{112} plane, respectively, as we will show below. Moreover, Fig. 3 also shows that the dips of the $\{10\bar{1}0\}$ family are missing for the Si atoms in the epilayer but do exist for the Y atoms in the same epilayer. This is not surprising because the lattice structure of $YSi_{1.7}$ and the location of Si and Y atoms in $YSi_{1.7}$ are identical to those in $ErSi_{1.7}$. Therefore, the reason of this phenomenon is the same as we explained in Ref. 21.

C. Angular scan along the $YSi_{1.7}\{10\bar{1}0\}$ and Si {112} plane

Figure 4 shows the angular scan along the $YSi_{1.7}\{10\bar{1}0\}$ and Si {112} planes, which are parallel with each other as shown in Fig. 3, by choosing two windows corresponding to the Y atoms in the epilayer and the Si atoms in the substrate. The angular range for this scan is −45° to −37° from the Si [111] direction. A $\langle 210\rangle$ dip of the scan in the Si {112} plane is expected to appear at −39.23° for a cubic structure and can be used for calibration. For the scan in the $YSi_{1.7}\{10\bar{1}0\}$ plane, a dip corresponding to the $\langle 1\bar{2}13\rangle$ axis is expected to appear at ~ 43°. Because the angle measured between the

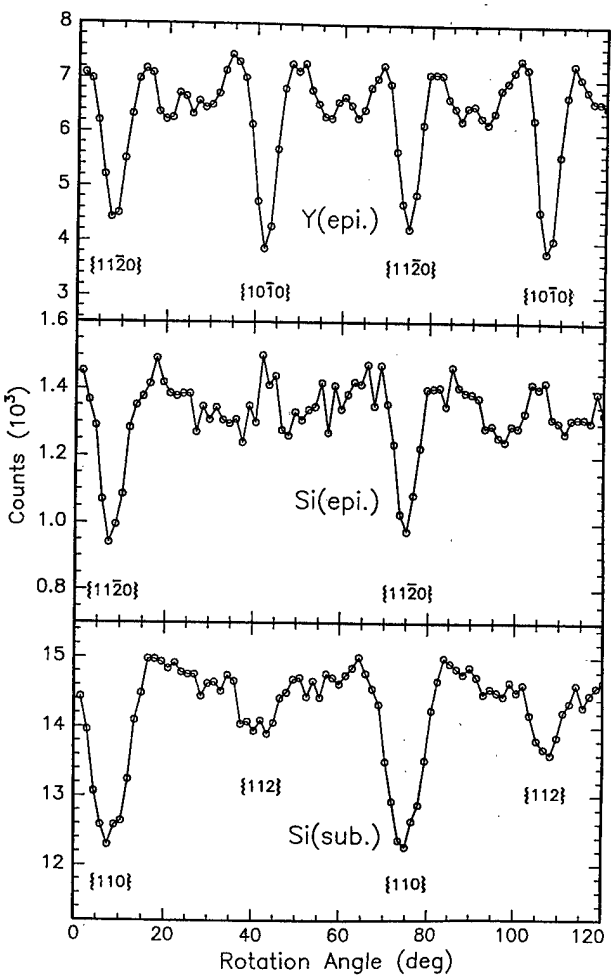


FIG. 3. Rotational angular scan of the annealed sample around the Si [111] axis showing the planar dips for the Y and Si atoms in the epilayer, and the Si atoms in the substrate. The $\{10\bar{1}0\}$ planar dips are missing for the Si atoms in the epilayer.

$YSi_{1.7}$ $\langle 1\bar{2}13\rangle$ dip and Si $\langle 210\rangle$ dip is 4.10°±0.05°, the measured angle Φ_{epi} between $YSi_{1.7}$ [0001] and $YSi_{1.7}$ $\langle 1\bar{2}13\rangle$ axis is 43.33°±0.05°. However, for bulk $YSi_{1.7}$ the angle Φ_{bulk} between [0001] and $\langle 1\bar{2}13\rangle$ axis is equal to $\tan^{-1}(a_{bulk}/c_{bulk})$, which should be 42.83°, using the lattice parameters $a_{bulk}=0.3842$ nm and $c_{bulk}=0.4144$ nm.[2] The $\Delta\Phi=\Phi_{epi}-\Phi_{bulk}=0.50\pm 0.05°>0$, indicating that the epilayer is compressively strained in the perpendicular direction. The tetragonal distortion deduced from this RBS/channeling study is $e_{T(RBS)}=\Delta\Phi/\sin\Phi\cos\Phi=1.75\%\pm0.17\%$.

D. X-ray symmetric $\theta-2\theta$ scan

Figure 5 shows the symmetric $\theta-2\theta$ scan of the annealed sample, with 2θ ranging from 10° to 110°, and indicates that the hexagonal $YSi_{1.7}$ phase is the only phase in the epilayer and that the $YSi_{1.7}$ [0001] axis is parallel to the Si [111] axis. To measure the perpendicular elastic strain, a more detailed symmetric $\theta-2\theta$ scan ranging from 94° to 100° was carried out (not shown). From the position of the Si (333) and

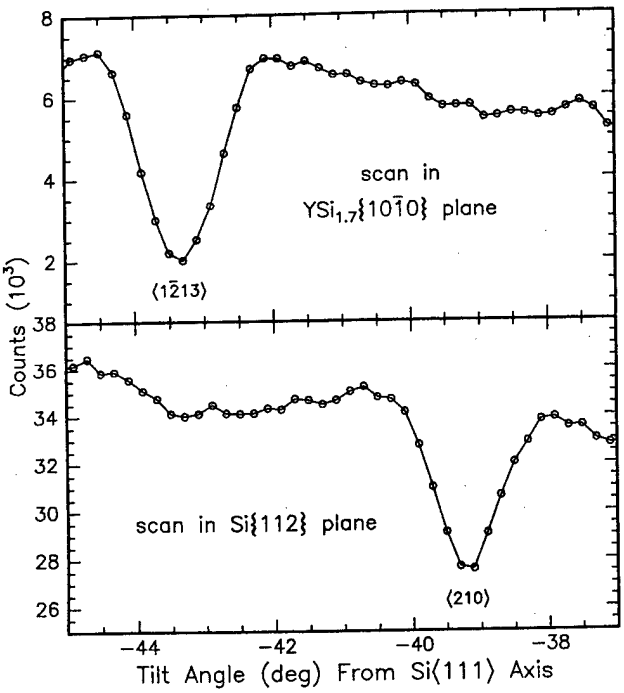


FIG. 4. Angular scan of the annealed sample along the Si {112} plane showing the ⟨210⟩ dip in the Si {112} plane and the ⟨1$\bar{2}$13⟩ dip in the $YSi_{1.7}$ {10$\bar{1}$0} plane.

$YSi_{1.7}$ (0004) peaks in this scan, a value of $\Delta\theta = 0.99°$ is obtained leading to a value of $c_{epi} = 0.4116$ nm. Therefore the deduced perpendicular elastic strain including the estimated error is $e^{\perp} = (c_{epi} - c_{bulk})/c_{bulk} = -0.67\% \pm 0.02\%$.

The c_{epi} is smaller than c_{bulk}, or $e^{\perp} < 0$ indicating that the epilayer is compressively strained in the perpendicular direction, the same conclusion as obtained from the RBS/channeling study in Sec. III C.

E. Asymmetric $\theta-2\theta$ scan

The parallel strain can be measured by the asymmetric $\theta-2\theta$ scan combined with the result from the symmetric $\theta-2\theta$ scan. As we discussed previously (see Ref. 22 for the details), two asymmetric $\theta-2\theta$ scans (one of them with the beam direction reversed) containing the Si (331) and $YSi_{1.7}$ (01$\bar{1}$3) reflection peaks as shown in Fig. 6 have to be carried out. From these two scans, $\Delta\theta^a = -1.076°$ and $\Delta\theta^b = -0.642°$ were obtained and from this the $\Delta\theta_d = (\Delta\theta^a + \Delta\theta^b)/2 = -0.859°$ was deduced. Finally, similar to the procedure in Ref. 22, $a_{epi} = 0.3882 \pm 0.0003$ nm, $e^{\parallel} = (a_{epi} - a_{bulk})/a_{bulk} = +1.04\% \pm 0.08\%$ and $e_{T(XRD)} = e^{\parallel} - e^{\perp} = 1.71\% \pm 0.10\%$ are deduced, which is consistent with $e_{T(RBS)} = 1.75\% \pm 0.17\%$.

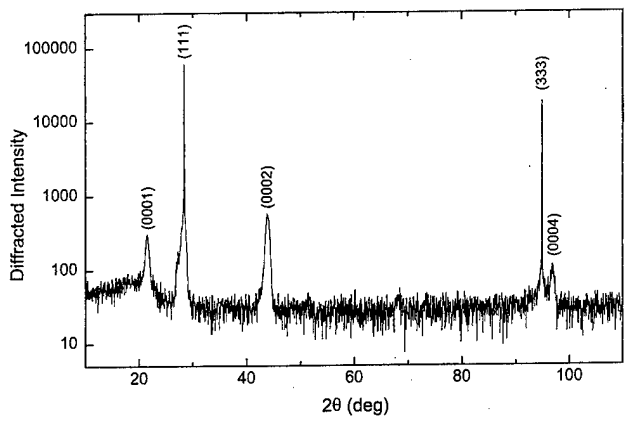


FIG. 5. Symmetric $\theta-2\theta$ scan of the annealed sample with 2θ ranging from 10° to 110°.

F. Discussion of the elastic strain

Table I lists the values of thickness, $e^{\perp}$, $e^{\parallel}$ and e_T from our results of $YSi_{1.7}$, $ErSi_{1.7}$ and $GdSi_{1.7}$ epilayers and the result from Ref. 7. The results listed in Table I show:

(1) $e^{\parallel}$ and $e^{\perp}$ values of $YSi_{1.7}$ layers are much larger than the values in Ref. 7. It is reported that the tetragonal distortion $e_T (= e^{\parallel} - e^{\perp})$ of the epilayer, which is related to the strain, decreases when the thickness of the epilayer increases and finally $e_T = 0$ when the thickness is larger than a certain value.[23,24] Because the thickness (48 nm) of our $YSi_{1.7}$ layer is smaller than that in Ref. 7, it is reasonable that the strain in our thinner epilayer is less relaxed and has a larger value than that in Ref. 7. Besides, the result in Ref. 7 also shows a thickness dependence of $e^{\parallel}$; and a strain-free epilayer was obtained at a thickness of 188 nm. In addition to the thickness dependence, other factors such as higher beam energy and a higher defect density in the implanted layer caused by the nonchanneled implantation in Ref. 7 may also influence the result of strain.

(2) The -1.1% lattice mismatch of the $ErSi_{1.7}$/Si (111) system resulted in the CIBS-formed quasipseudomorphic $ErSi_{1.7}$ layers ($\epsilon^{\parallel} = +0.13\%$),[22] while the $\sim 0\%$ lattice mismatch of $YSi_{1.7}$/Si (111) system, on the contrary, results in a nonpseudomorphic growth of $YSi_{1.7}$ layer with a rather large parallel strain relative to the Si substrate ($\epsilon^{\parallel} = +1.09\%$). It is clear that a simple pseudomorphic growth model cannot explain this. To explain the experimental strain result of RE silicide epilayers on Si substrate, three points should be emphasized: (i) a large difference of thermal expansion coefficient between the epilayer and the substrate plays an important role to the growth of the epilayer; (ii) the misfit dislocations formed at high temperature remain completely or partially during the cooling process, (iii) upon heating or cooling, the lattice constant $a_{epi}(T)$ of the epilayer at temperature T(°C) is controlled completely (for a pseudomorphic growth) or mostly by the lattice constant of the substrate $a_{Si}(T)$ which is related to the thermal expansion coefficient $\alpha(\mathrm{Si}) = 2.6 \times 10^{-6}$/K instead of the thermal expansion coefficient of the epilayer.

In the case of $YSi_{1.7}$, we assume that the thermal expansion coefficient $\alpha(YSi_{1.7})$ has a value in the order of 15×10^{-6}/K like several transition-metal disilicides of Ti, Ta, Mo and W.[25] Considering that the pseudomorphic growth is

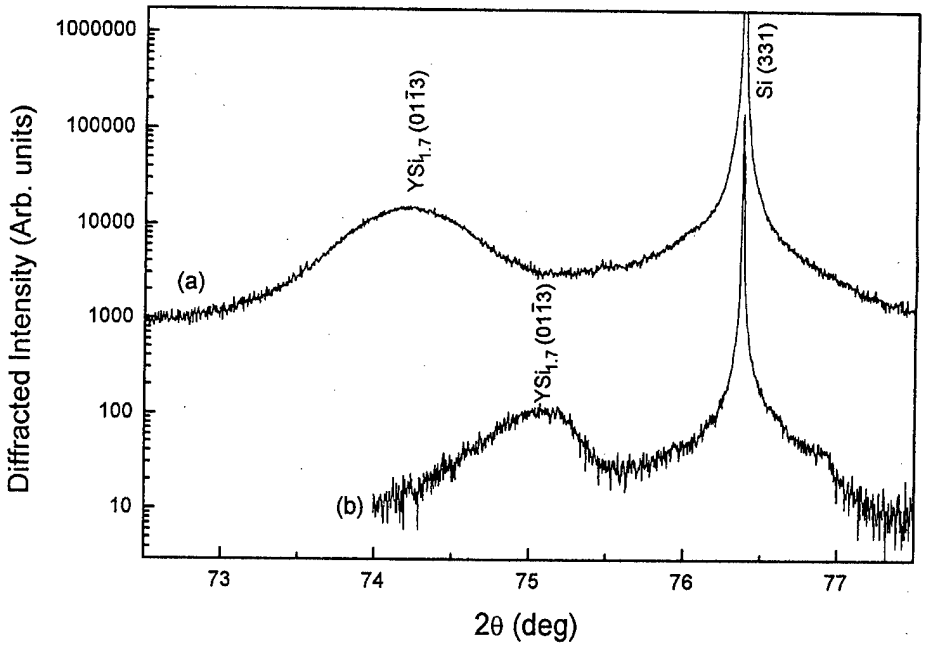


FIG. 6. Asymmetric $\theta-2\theta$ scan of the annealed sample showing the (331) and (01$\bar{1}$3) peaks: (a) glancing exit geometry, (b) glancing incident geometry, i.e., the beam direction is reversed with respect to the glancing exit geometry.

always retained upon heating, then the lattice constant of the epilayer $a_{epi}(T)$ as a function of temperature T is $a_{epi}(T) = a_{Si}(T) = 0.3840$ nm $[1-2.6\times10^{-6}\times(T-20)]$, and the parallel elastic strain $e^{\parallel}=[a_{epi}(T)-a_b(T)]/a_b(T)$, in which $a_b(T)=0.3842\times[1+15\times10^{6}(T-20)]$ is the lattice constant of $YSi_{1.7}$ bulk materials at temperature T, will be changed along the route A (20 °C)$\rightarrow B$(1000 °C) in Fig. 7 and a_{epi} will be changed from 0.3840 nm at point A to 0.384 98 nm, which results in $e^{\parallel}=-1.25\%$ at point B. Due to the large compressive parallel strain at point B, the strain is relieved and the misfit dislocations are formed leading to, for example, the point C. So that the compressive strain changes to a value of -0.17%, i.e., a_{epi}(1000 °C) changes from 0.384 98 nm to 0.38918 nm. Finally, if the misfit dislocations remain during the cooling process and the $a_{epi}(T)$ is controlled completely by $a_{Si}(T)$, then at room temperature, a_{epi}(20 °C) = 0.389 18 nm $[1-2.6\times10^{-6}\times980]$ = 0.388 19 nm and thus results in point D with $e^{\parallel}$ = (0.388 19–0.3842) nm/0.3842 nm = −1.04%. Although, the actual process might not be strictly the ABCD as shown in Fig. 7, it will clearly follow this trend and finally reach point D.

The −1.1% lattice mismatch of $ErSi_{1.7}$/Si (111) which results in a quasipseudomorphic growth with $a_{epi} \approx a_{Si}$(0.3840 nm), $\epsilon^{\parallel}=+0.13\%$ and $e^{\parallel}=+1.24\%$ (Ref. 22) can also be explained via a route like EFGH shown in Fig. 7 in a similar way.

For the $GdSi_{1.7}$/Si (111) system, which has the largest positive lattice mismatch of +0.94% among all RE $Si_{1.7}$/Si(111) systems. At first sight, the epitaxial growth of the $GdSi_{1.7}$ epilayers should result in a compressive strain in the (0001) plane and a tensile strain in the c-axis direction. However, both the results[26] of x-ray symmetric $\theta-2\theta$ scan ($e^{\perp}=-0.87\%$) and RBS/channeling ($\Delta\Phi=+0.50°$) studies confirm, on the contrary, that the epilayer actually is compressively strained perpendicularly. This experimental result can also be explained, via the route IJKL shown in Fig. 7, in a similar way with the model used above.

TABLE I. Comparison of the thickness, $e^{\perp}$, $e^{\parallel}$ and e_T values of $YSi_{1.7}$, $ErSi_{1.7}$ and $GdSi_{1.7}$ epilayers formed by IBS and CIBS.

Epilayer	E (keV)	Dose (10^{17}/cm^2)	Thickness (nm)	$e^{\perp}$ (%)	$e^{\parallel}$ (%)	e_T (%)	Ref.
$YSi_{1.7}$ (IBS)	330	2	100[a]	−0.1	+0.4	+0.5	7
$YSi_{1.7}$ (IBS)	330	3.6	188		+0.0		7
$YSi_{1.7}$ (CIBS)	60	1.0	48	−0.67	+1.04	+1.71	This work
$ErSi_{1.7}$ (CIBS)	90	1.6	80	−0.94	+1.24	+2.18	22
$GdSi_{1.7}$(CIBS)	90	1.3	70	−0.87	+0.88	+1.75	26

[a]Dose of 2×10^{17}/cm^2 should result in a thickness ~100 nm instead of 64 nm given in Ref. 7.

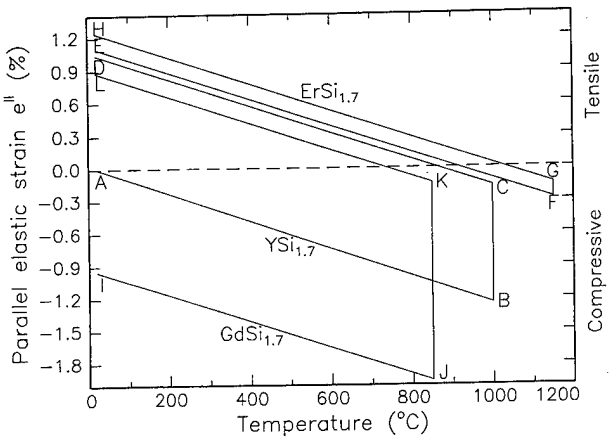


FIG. 7. Schematic plot of the parallel elastic strain $e^{\parallel}$ as a function of temperature for $YSi_{1.7}$/Si (111), $ErSi_{1.7}$/Si (111) and $GdSi_{1.7}$/Si (111) systems.

IV. CONCLUSIONS

(1) CIBS is not only indispensable for synthesis of RE silicides by implantation of RE elements with heavy mass, such as Er and Gd in a Si substrate, but it is also helpful in the synthesis of silicides with better quality by implantation of elements with medium mass such as Y.

(2) The strain value of our $YSi_{1.7}$ layer is markedly larger than published values. However, it is reasonable if we consider the strain-thickness dependence.

(3) To explain the strain of an epitaxial RE silicide layer on a Si (111) substrate, in addition to the lattice mismatch, the thermal mismatch, i.e., the difference of thermal expansion coefficient between the epilayer and the substrate should also be considered.

ACKNOWLEDGMENTS

The authors grateful to Yutian Wang for the XRD measurements and discussions. This work has been supported by the National Natural Science Foundation of China (No. 19775001), the National Integrated Optoelectronic (NIOE) Laboratory in China and the Belgian Fund for Scientific Research, Flanders (FWO) and Concerted Action (GOA) programs.

[1]J. A. Knapp and S. T. Picraux, Mater. Res. Soc. Symp. Proc. **54**, 261 (1986).
[2]J. A. Knapp and S. T. Picraux, Appl. Phys. Lett. **48**, 466 (1986).
[3]R. D. Thompson, B. Y. Tsaur, and K. N. Tu, Appl. Phys. Lett. **38**, 535 (1981).
[4]M. Gurvitch, L. Manchanda, and J. M. Gibson, Appl. Phys. Lett. **51**, 919 (1989).
[5]M. P. Siegal, F. H. Kaatz, W. R. Graham, J. J. Santiago, and J. Van der Spiegel, J. Appl. Phys. **66**, 2999 (1989).
[6]A. E. White, K. T. Short, R. C. Dynes, R. Hull, and J. M. Vanderberg, Nucl. Instrum. Methods Phys. Res. B **39**, 253 (1989).
[7]T. L. Alford and J. C. Barbour, Mater. Res. Soc. Symp. Proc. **157**, 137 (1990).
[8]R. Baptist, S. Ferrer, G. Grenet, and H. C. Poon, Phys. Rev. Lett. **64**, 311 (1990).
[9]T. L. Alford, P. Borgesen, J. W. Mayer, and D. A. Lilienfeld, J. Appl. Phys. **67**, 1288 (1990).
[10]T. L. Alford, P. Borgesen, J. W. Mayer, and D. A. Lilienfeld, Appl. Phys. Lett. **58**, 1848 (1991).
[11]T. L. Alford and J. W. Mayer, Appl. Phys. Lett. **59**, 2989 (1991).
[12]T. L. Lee, L. J. Chen, and F. R. Chen, J. Appl. Phys. **71**, 3307 (1992).
[13]S. Jin, J. H. Lin, L. J. Chen, W. D. Shi, and Z. Zhang, Nucl. Instrum. Methods Phys. Res. B **96**, 347 (1995).
[14]T. L. Alford, J. Appl. Phys. **77**, 1010 (1995).
[15]V. M. Koleshko, V. F. Belitsky, and A. A. Khodin, Thin Solid Films **141**, 277 (1986).
[16]M. F. Wu, A. Vantomme, G. Langouche, H. Vanderstraeten, and Y. Bruynseraede, Nucl. Instrum. Methods Phys. Res. B **54**, 444 (1991).
[17]M. F. Wu, A. Vantomme, H. Pattyn, G. Langouche, K. Maex, and J. Vanhellemont, Appl. Surf. Sci. **38**, 217 (1989).
[18]M. F. Wu, A. Vantomme, H. Pattyn, G. Langouche, and H. Bender, Appl. Surf. Sci. **102**, 184 (1996).
[19]M. F. Wu, A. Vantomme, H. Pattyn, and G. Langouche, Appl. Phys. Lett. **67**, 3886 (1995).
[20]M. F. Wu, A. Vantomme, H. Pattyn, G. Langouche, and H. Bender, Appl. Phys. Lett. **68**, 3260 (1996).
[21]M. F. Wu, A. Vantomme, J. De Wachter, S. Degroote, H. Pattyn, G. Langouche, and H. Bender, J. Appl. Phys. **79**, 6920 (1996).
[22]M. F. Wu, A. Vantomme, H. Pattyn, G. Langouche, Q. Yang, and Q. Wang, J. Appl. Phys. **80**, 5713 (1996).
[23]S. Hashimoto, J.-L. Peng, and W. M. Gibson, Appl. Phys. Lett. **47**, 1071 (1985).
[24]S. Hashimoto, L. J. Schowalter, R. W. Fathauer, and W. M. Gibson, Mater. Res. Soc. Symp. Proc. **56**, 247 (1986).
[25]G. Bai, M.-A. Nicolet, and T. Vreeland, Jr., J. Appl. Phys. **69**, 6451 (1991).
[26]Unpublished results related to the work in Ref. 20.

Implanted gallium-ion concentrations of focused-ion-beam prepared cross sections

T. Ishitani[a] and H. Koike
Instrument Division, Hitachi, Ltd., Hitachinaka 312-8504, Japan

T. Yaguchi and T. Kamino
Techno Research Laboratory, Hitachi Instruments Engineering Co., Ltd., Hitachinaka 312-8504, Japan

(Received 21 January 1998; accepted 3 April 1998)

A gallium (Ga) focused-ion-beam (FIB) has been popularly used to prepare cross-sectional samples for transmission electron microscopes (TEMs) and scanning electron or ion microscopes. However, characteristics of the FIB-prepared cross sections such as ion concentration and radiation damage have been little studied either in theory or in experiment. In the present study, cross sections prepared by 30 keV Ga FIB are modeled using a combination of analytical and Monte Carlo methods to calculate the implanted Ga concentration. It is found that the Si/W layered sample is cross sectioned at grazing angles $\beta \approx 2.5°$ and 6° for these layers, respectively. The implanted Ga ions for the cross-sectioned Si and W layers are concentrated very near their surfaces of <10 nm to yield the Ga concentrations C_{Ga} of about 4 and 9 at % for these layers, respectively. Although there is some differences in sample materials between the calculations and the experiments, the calculated C_{Ga} values for Si and W layers roughly agree with the experimental values for the magneto-optical disk TEM sample. This agreement firmly supports the present modeling of FIB-milled cross sections. © *1998 American Vacuum Society.* [S0734-211X(98)00104-8]

I. INTRODUCTION

In recent years, a gallium (Ga) focused-ion-beam (FIB) has been popularly used for cross-sectional sample preparation[1–3] for transmission electron microscopes (TEMs) and scanning electron or ion microscopes (SEMs or SIMs). The FIB cross sectioning has the following advantages: (1) cross sectioning at any specified position with high positional accuracy of 0.1 μm or better; (2) little restriction on sample materials; (3) stress-free milling; and (4) visual process from milling-area specification, milling monitor, to final-cut confirmation using SIM images. However, there is neither experimental nor theoretical reports on characteristics of the FIB-prepared cross sections such as ion concentration and radiation damage except for the two reports by Beag *et al.*[4] and Sakamoto *et al.*[5]

In the present study, assuming that the cross sections are formed under a steady state of ion-target interaction such as ion implantation, backscattering, and sputtering, the authors have analytically expressed the cross sections to calculate their ion concentrations. Monte Carlo (MC) method has been used to simulate both ion and recoiled target-atom trajectories to calculate the yields of ion implantation, backscattering, and sputtering. The calculated Ga concentrations depending on the sample material are compared with the experimental ones for the FIB cross-sectioned magneto-optical disk having a multilayered structure. Good agreements between them support the present model.

[a]Electronic mail: ishitani@cm.naka.hitachi.co.jp

II. CALCULATION

A. Steady-state cross-sectioning and surface concentrations

Let us consider Ga FIB cross sectioning of the target material A. Here, Ga is the most popular ion species of the cross sectioning FIB. At the final milling, the cross section is approximated to be in steady state, where as many implanted atoms are removed by sputtering as are replenished by implantation [see Fig. 1(a)]. Then, we have

$$S_{Ga} = Y_{imp} \tag{1}$$

and

$$S_{Ga}/S_A = R_{Ga\text{-}A}(C_{Ga}/C_A), \tag{2}$$

where S_{Ga} and S_A are the partial sputtering yields of components Ga and A, respectively, S is the total sputtering yield $(=S_{Ga}+S_A)$, Y_{imp} is the implantation yield, C_{Ga} and $C_A(=1-C_{Ga})$ are the atomic concentrations of their components near the target surface, and $R_{Ga\text{-}A}$ is the enhancement factor in the Ga preferential sputtering for the host matrix A (i.e., the ratio of probability for a Ga atom near the surface to be sputtered to that of an A atom to be sputtered). The $R_{Ga\text{-}A}$ value may not be unity because of differences in surface binding energies, sputter escape depths, and energy transfers within the collision cascade.[6]

From Eqs. (1) and (2), C_{Ga} is derived as

$$C_{Ga} = 1/[R_{Ga\text{-}A}(S/Y_{imp} - 1) + 1]. \tag{3}$$

At $R_{Ga\text{-}A} = 1$, it simply yields to

$$C_{Ga} = Y_{imp}/S. \tag{4}$$

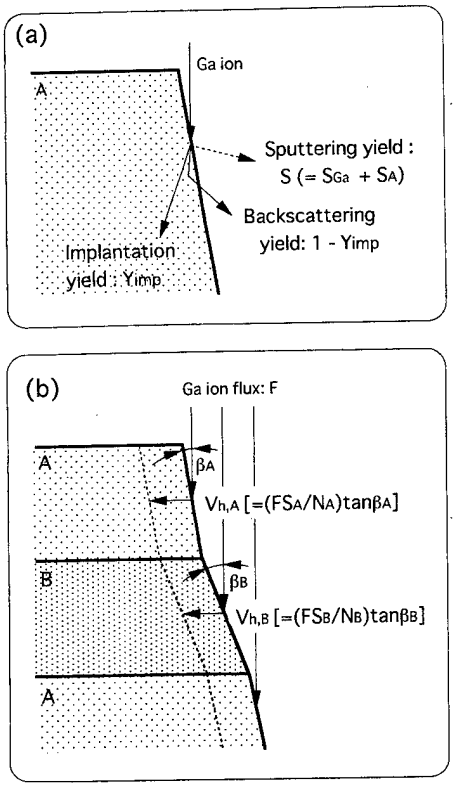


FIG. 1. (a) Basic parameters of Ga FIB cross sectioning of the specimen A, i.e., sputtering yield $S(=S_{Ga}+S_A)$, implantation yield Y_{imp} and backscattering yield. (b) Steady-state cross sectioning of A-B layered sample, where the equation of $V_{h,A}=V_{h,B}$ is satisfied.

When the beam of ion flux F (ions/m^2/s) incidents on the cross-section at grazing angle β as shown in Fig. 1(b), the cross-section is eroded at velocity V (m/s):

$$V=(F\ \sin\ \beta)(S/N), \tag{5}$$

where N is the atomic density of the sample. Then, the horizontal velocity of the cross section leads to

$$V_h=V/\cos\ \beta=(FS/N)\tan\ \beta \tag{6a}$$

$$=(\beta\cdot S/N)F \quad [\text{at } \beta(\text{rad})\ll 1]$$

$$\propto \beta S/N(\equiv k) \quad (\text{at } F=\text{constant}). \tag{6b}$$

As to the multilayered cross section composed of components A and B (with β_A and $\beta_B\ll 1$) as shown in Fig. 1(b), the steady-state milling satisfies a relationship of $V_{h,A}=V_{h,B}$, i.e.,

$$\beta_A S_A/N_A=\beta_B S_B/N_B \quad (\text{at } \beta\ll 1). \tag{7}$$

In a strict sense of the surface composition C_i (i=Ga or A) in Eqs. (2)–(4), it is referred to the properly weighted average composition of the layer from which sputtering occurs, i.e., the top two to three atomic layers (as discussed later),

$$C_i=(1/S_i)\int_0^\infty C_i(z)E_i(z)dz, \tag{8}$$

where $E_i(z)dz$ is the i-atom's contribution to S_i from a layer dz at depth z, and $C_i(z)$ is the depth-dependent i-atom concentration.

B. Depth-dependent Ga concentration

The modification of ion implantation profile due to sputtering can be calculated under a few simplifying assumptions: (1) the sputtering yield is constant, and is first the same for the substrate and the implanted ions; (2) there is no knock-on effect; and (3) the volume change due to radiation damage can be neglected. Multipling a normalized depth probability of the implanted ions, $P(z)$ (m^{-1}), by the ion flux on the cross-section, $F\ \sin\ \beta$ (ions/m^2/s), we get the density of ions implanted at depth z per unit time:

$$g(z,t)=(F\ \sin\ \beta)P(z),$$

$$=(VN/S)P(z) \quad (\text{ions/m}^3/\text{s}). \tag{9}$$

Then, the Ga implantation density profile (ions/m^3) at time t is calculated by integrating Eq. (9) as follows:

$$C(z,t)=\int_0^t g(z+Vt',t')dt',$$

$$=(VN/S)\int_0^t P(z+Vt')dt'. \tag{10}$$

The $C(z,t)$ change is schematically shown in Fig. 2. Using $z'=Vt'$, the steady-state implantation density profile $C(z)\{=C(z,t=\infty)\}$ [ions/m^3] is expressed as

$$C(z)=(N/S)Pa(z), \tag{11}$$

where

$$Pa(z)=\int_0^\infty P(z+z')dz'. \tag{12}$$

Here, the $C(z)$ and $Pa(z)$ values have maximums at the surface ($z=0$) and they decrease to zero at $z=$ maximum ion depth, as later shown in Fig. 5. Since $Pa(0)=Y_{imp}$, the Ga concentration at the surface, $C_{Ga}[=C(0)/N]$, equals to Y_{imp}/S, which is the same as Eq. (4).

The total implanted ions per unit area is obtained by integrating $C(z)$ with respect to z from 0 to ∞, and yields to

$$C_t=(N/S)Pb(z=\infty), \tag{13}$$

where

$$Pb(z)=\int_0^z Pa(z)dz. \tag{14}$$

In the TEM specimen preparation using FIB cross sectioning,[1,3,7] two trenches are FIB milled, leaving a narrow strip (approximately 0.1 μm in width), which is the TEM section of interest. The specimen thickness is usually larger

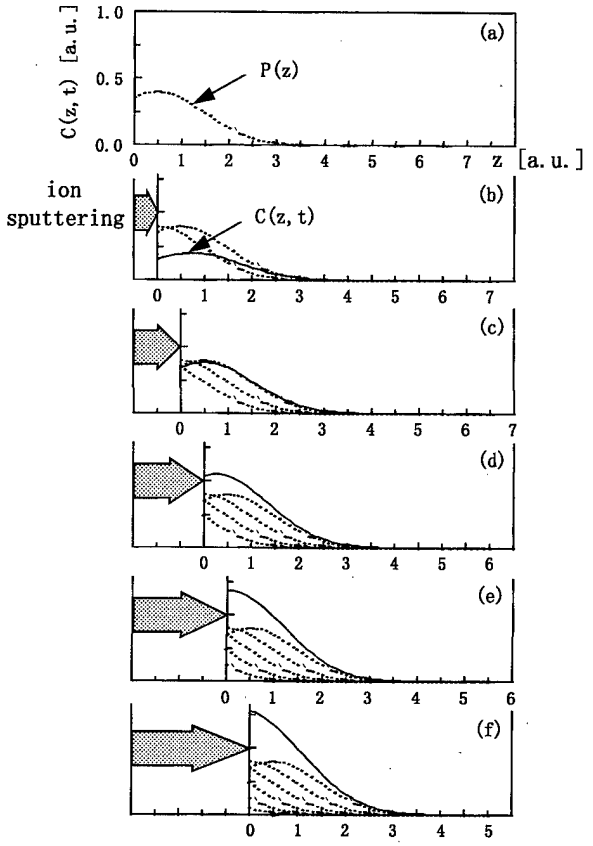


FIG. 2. Schematic modification of implanted ion depth profile due to ion sputtering.

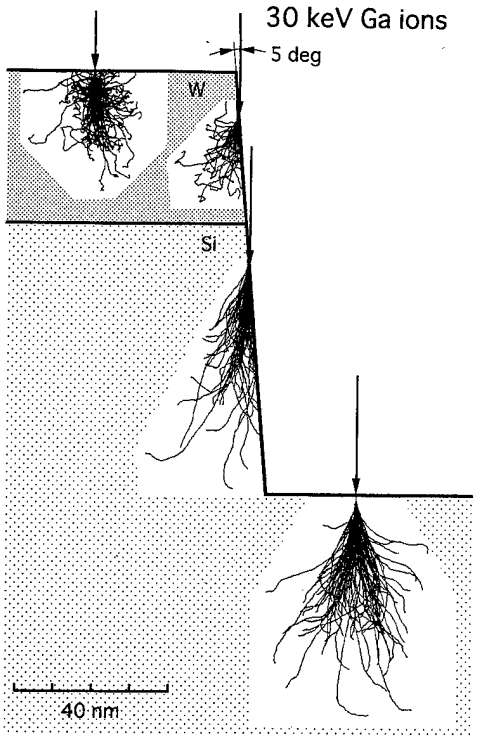


FIG. 3. Typical trajectories of 30 keV Ga ions bombarded on the W-Si sample. (The number of incident ions is 100.)

than the transverse depth of 30 keV Ga ions, typically about 15 nm for Si and 10 nm for W (as shown later in Fig. 3). Finally, the average Ga concentration in the t-thick specimen is calculated by

$$C_{t,\mathrm{av}}=2C_t/(2C_t+tN), \tag{15a}$$

$$\approx 2C_t/tN \quad \text{when } C_t \ll tN/2. \tag{15b}$$

Here, the factor 2 of $2C_t$ comes from both front- and back-side FIB cross sectioning of the TEM specimen. On energy-dispersive x-ray (EDX) analysis in TEM discussed in the later section, since the EDX information depth is much larger than the TEM sample thickness, the Ga EDX intensity does show the $C_{t,\mathrm{av}}$ value, not to the C_{Ga} value.

C. Monte Carlo simulation

Monte Carlo (MC) method has been applied to simulate both ion and recoiled-atom trajectories on the finishing stage of FIB cross sectioning. Basic properties such as S, Y_{imp}, and ion depth probability $P(z)$ are calculated under silicon (Si) and tungsten (W) targets bombarded by 30 keV Ga ions at various grazing angles β. Here, the Si and W have been chosen as typical components of Si devices. The Monte Carlo program used here is essentially the same as the previous one[2,8,9] based on a single scattering model. The following assumptions have been made. (1) The target has a random or amorphous-like structure. (2) Energetic particle is scattered by only one target atom at a time. (3) The ZBL interaction potential[10,11] can be used to describe its binary collision. (4) As to electronic energy loss at low energies, a velocity-proportional stopping power is used. (5) Accumulation of implanted ions is ignored. (6) Diffusion effect of implanted ions is also not considered. (7) The surface barrier for the sputtered atoms is a planar type and is taken to the surface binding energy U_s.

III. RESULTS AND DISCUSSION

A. Results with no preferential-sputtering correction

The atomic numbers, atomic masses, mass ratios, atomic densities, and binding energies of Ga, Si, and W are shown in Table I. The mass ratio is an important parameter in elastic two-body collision. Ga projectiles in the Ga-Si collision are forwardly scattered only in $\theta \leqslant 0.415$ $(=23.8°)$ $[=\sin^{-1}(M_{\mathrm{Si}}/M_{\mathrm{Ga}})]$ (in the laboratory system), in contrast with full-angle scattering in the Ga-W collision, i.e., $\theta \leqslant \pi$ $(=180°)$.

TABLE I. Atomic numbers Z, atomic masses M, mass ratios M/M_{Ga}, atomic weight densities ρ, atomic number density N, and binding energies E_{b} of Ga, Si, and W targets.

Target	z	M (amu)	M/M_{Ga}	ρ (kg/m^3)	N (atoms/m^3)	E_b (eV)
Ga	31	69.72	1.00	5.91E+3	5.11E+28	2.82
Si	14	28.086	0.403	2.42E+3	5.19E+28	4.70
W	74	183.85	2.64	1.93E+4	6.32E+28	8.68

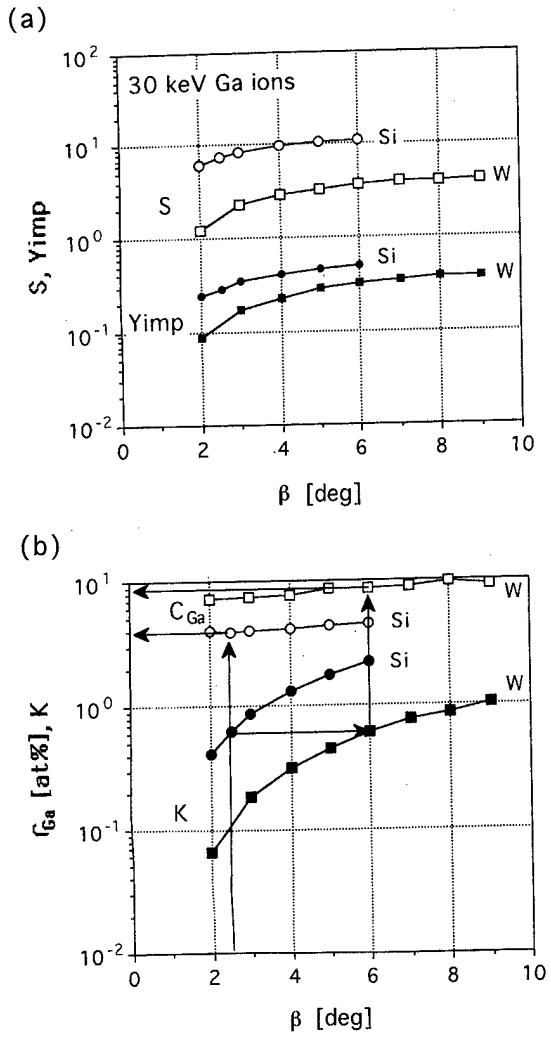


FIG. 4. (a) Calculated implantation yield Y_{imp} and sputtering yield S at various β for the 30 keV Ga ion — Si and W targets and (b) their corresponding values of $K[=(\beta S/N)\times 1E+29$; see Eq. (6b)] and surface Ga concentration C_{Ga} [defined as Eq. (4)] at various β.

Typical ion trajectories under the bombardment of only 100 Ga ions (at 30 keV) are plotted in Fig. 3 for W-Si layered cross section at $\beta=5°$. The ion trajectories at normal incidence ($\beta=90°$) are also plotted in the same figure to be compared with the cross-section's data. For the S and Y_{imp} calculations, a large number of trajectories of both 5000 Ga ions and their recoiled target atoms have been simulated to reduce their statistic fluctuations. The Y_{imp} and S values at various β are plotted in Fig. 4(a). Similarly, their corresponding values of $K[\equiv k\times 1E+29$; see Eq. (6b)] and C_{Ga} [defined in Eq. (4)] are plotted in Fig. 4(b). It is found that the W target is always smaller in both Y_{imp} and S than the Si target under the same β condition. Besides, the S and Y_{imp} curves gradually increase with β, but the $Y_{imp}/S(=C_{Ga})$ curve shows a very weak dependence on β.

Let us consider a W-Si layered cross section prepared in the steady-state milling, i.e., $V_{h,W}=V_{h,Si}$. Taking the steady-state β_{Si} value as 0.0436 ($=2.5°$) from the previous experiments,[1,7] we get $\beta_W\approx 0.105(=6°)$ and $C_{Ga,W}\approx 8.7$

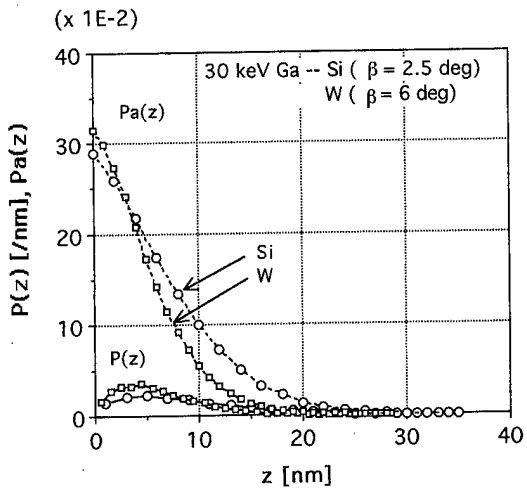


FIG. 5. Calculated Ga depth probability $P(z)$ and its accumulated probability $Pa(z)$: (a) Si ($\beta=2.5°$) target; (b) W ($\beta=6°$) target.

at %, in contrast with $C_{Ga,Si}\approx 3.9$ at % It is worthy of note that the K curves show a strong dependency on β, but the $C_{Ga}(=Y_{imp}/S\times 100)$ curve does not.

Calculated $P(z)$ and $Pa(z)$ profiles for the Si ($\beta=2.5°$) and W ($\beta=6°$) targets are plotted in Fig. 5. Energy distributions of the backscattered ions and sputtered atoms, $Yb(E)$ and $S(E)$, are shown in Fig. 6. [Only the $Yb(E)$ and $S(E)$ results have been obtained for the MC simulations of 10 000 ion bombardment.] The $Yb(E)$ peaks right under 30 keV observed at low β correspond to the ion reflection and they disappear with the increase in β as already predicted[8] The $S(E)$ distribution decreases proportional to E^n with $n\approx 1.7$ in the region of $E=1-10$ keV. This n value is in good agreement both with another MC simulation's value[12] and with a theoretical one,[13] which is derived from a mature collision cascade. The theoretical value expressed as $n=3-2m$ is 2.6 at $m=0.203$ for the ZBL potential.[11] The abrupt $S(E)$ decrease in the region of $E>10$ keV predicts that the sputtered atoms result from the small number of collisions, i.e., immature collision cascade. Backscattering yields ($=1-Y_{imp}$) for the Si ($\beta=2.5°$) and W ($\beta=6°$) targets are as large as 0.71 and 0.69, respectively, and their average energies are as high as 22 keV for Si and 20 keV for W. Since almost all the sputtered atoms originate in the surface layers of only 1nm thick (e.g., about 82 and 95% for the Si and W targets, respectively), the C_{Ga} values mean Ga concentrations in these surface layers as expressed in Eq. (8).

Table II shows the calculated $C_{t,av}$ values for typical

TABLE II. Calculated values of $Pa(z=0)$, $Pb(z=\infty)$, C_t, C_{Ga}, and $C_{t,av}(t=50$ and 100 nm) for Si ($\beta=2.5°$) and W ($\beta=6°$) targets.

Target	$Pa(0)$	$Pb(\infty)$	C_{Ga} (at %)	C_t (ions/nm^2)	$C_{t,av}$ [at %] $t=50$ nm	$t=100$ nm
Si ($\beta=2.5°$)	0.289	2.47	3.9	17.4	1.3	0.68
W ($\beta=6°$)	0.314	1.98	8.7	34.5	2.1	1.1

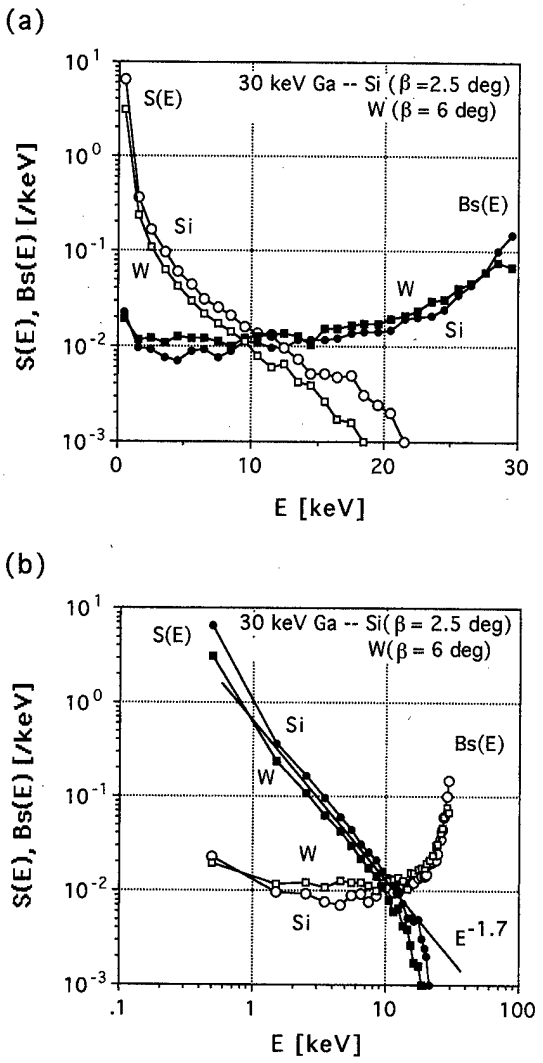


FIG. 6. Calculated energy distributions of the backscattered ions and sputtered atoms, $Yb(E)$ and $S(E)$: (a) $\log[Yb(E)$ or $S(E)]-E$ graph and (b) $\log[Yb(E)$ or $S(E)]-\log(E)$ graph.

TEM Si ($\beta=2.5°$) and W ($\beta=6°$) specimens of 50 and 100 nm in thickness as well as their basic values of $Pa(z=0)$, $Pb(z=\infty)$, C_{Ga}, and C_t. Here, no preferential sputtering is taken into consideration, i.e., $R_{Ga\text{-}A}=1$. Consequently, the calculation predicts that the C_{Ga}, C_t, and $C_{t,av}$ values for the W ($\beta=6°$) cross section are larger than those for the Si ($\beta=2.5°$) cross section by 2.2, 2.0, and 1.6 times, respectively. For the unknown thick layered structure TEM specimens, the $C_{t,av}$ ratios among the layers are more useful because of their independence of the specimen thickness as predicted from Eq. (15b).

TABLE IV. Calculated C_{Ga} values with and without a $R_{Ga\text{-}A}$ correction for the Si ($\beta=2.5°$) and W ($\beta=6°$) cross sections, where $R_{Ga\text{-}A}$ is the Ga preferential sputtering factor for the target A, i.e., Si or W.

	Correction factor	C_{Ga} (at %)	
	$R_{Ga\text{-}A}$	Without correction	With correction
Si ($\beta=2.5°$)	0.899	3.92	4.34
W ($\beta=6°$)	1.34	8.66	7.07

B. Preferential-sputtering correction

Preferential sputtering may be possible when there are differences both in atomic masses and potentials of ion-atom and atom-atom interactions and in surface binding energies of the constituents. To estimate these contribution to the $R_{Ga\text{-}A}[\equiv(S_{Ga}/C_{Ga})/(S_A/C_A)$; A=Si or W] value, two cases of $U_{Ga}=U_{Si/W}$ and $U_{Ga}=2.82$ eV ($\neq U_{Si/W}$) are considered in the MC calculations for Si ($\beta=2.5°$) and W ($\beta=6°$) cross sections containing 10 at % Ga. Tables III(a) and III(b) show the calculated preferential sputtering factors $R_{Ga\text{-}A}$ with the calculated Y_{imp}, S_{Ga}/C_{Ga} and $S_{Si/W}/C_{Si/W}$ values. It is found that the lighter component and/or the component with smaller U value is preferentially sputtered, as expected from a general rule. The $R_{Ga\text{-}A}$ corrected C_{Ga} values for the Si ($\beta=2.5°$) and W ($\beta=6°$) cross sections are shown in Table IV with the uncorrected values. Here, the $R_{Ga\text{-}A}$ values mentioned above have been approximately used for the Si and W cross sections of interest because of weak C_{Ga} dependence on $R_{Ga\text{-}A}$. Consequently, the C_{Ga} values are corrected from 3.9 and 8.7 at % to 4.3 and 7.1 at % for the Si and W cross sections, respectively. However, the $R_{Ga\text{-}A}$ correction brings about little change in the $C_{t,av}$ values because C_{Ga} modifies the $P(z)$ profile only near the surface.

TABLE III. (a) Calculated values containing 10 at % Ga. Y_{imp}, S_{Ga}/C_{Ga}, S_{Si}/C_{Si}, and $R_{Ga\text{-}Si}$ for the Si ($\beta=2.5°$).

Target	U_{Ga}, U_{Si} (eV)	Y_{imp}	S_{Ga}/C_{Ga}	S_{Si}/C_{Si}	R_{GaSi}
10% Ga-90% Si	4.70,4.70	0.271	0.626/0.1(=6.26)	6.34/0.9(=7.04)	0.889
	2.82,4.70	0.273	0.642/0.1(=6.42)	6.43/0.9(=7.14)	0.899
100% Si	···, 4.70	0.289	···	7.37	···

(b) Y_{imp}, S_{Ga}/C_{Ga}, S_W/C_W, and $R_{Ga\text{-}W}$ for the W ($\beta=6°$) containing 10 at % Ga.

Target	U_{Ga}, U_W (eV)	Y_{imp}	S_{Ga}/C_{Ga}	S_W/C_W	$R_{Ga\text{-}W}$
10% Ga-90% W	8.68,8.68	0.306	0.469/0.1(=4.69)	3.31/0.9(=3.68)	1.28
	2.82,8.68	0.315	0.512/0.1(=5.12)	3.44/0.9(=3.82	1.34
100% W	···, 8.68	0.314	···	3.63	—

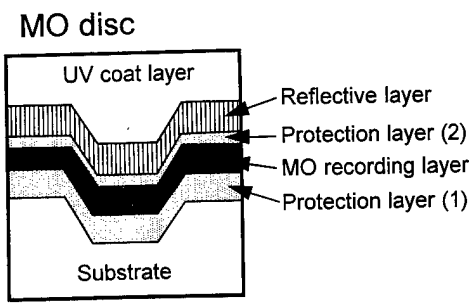


FIG. 7. Schematic diagram of cross-sectional magneto-optical disk showing multilayered structure.

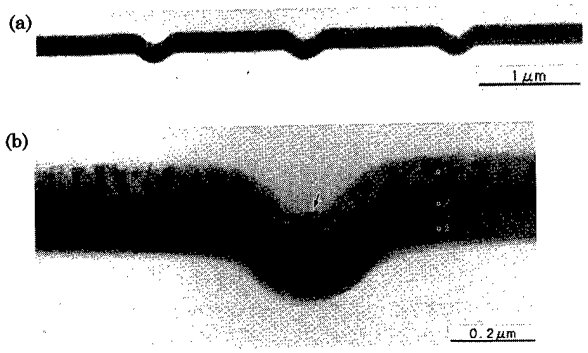


FIG. 8. Low and high magnification TEM images of cross-sectional magneto-optical disk. (The TEM system is a Hitachi model H-9000NAR of 300 kV accelerating voltage.) The disk has been cross-sectioned using 30 keV Ga FIB so as to be about 0.1 μm in thickness. EDX analysis data obtained from each layer indicates both the corresponding elements and the implanted Ga as shown in Table V.

IV. EXPERIMENTAL DATA AND DISCUSSION

A TEM specimen of magneto-optical (MO) disk has been prepared using Ga FIB cross sectioning to demonstrate the FIB's high validity.[2] The system used for FIB cross sectioning is Hitachi model FB-2000A with a 30 kV Ga ion source. Figure 7 schematically shows a layered structure of the MO disk, of which layers are composed of soft and hard materials as shown in Table V. Figures 8(a) and 8(b) show its low and high magnification TEM images, respectively. Here, the estimated specimen-thickness is about 0.1 μm. The layers have been EDX analyzed using Hitachi model HF-2000 TEM with a 200 kV field-emission type electron gun, which is capable of nm-level spatial resolution with minimum detectable mass fraction of <0.1 wt % on EDX.[14] Ga concentrations in the multilayers are tabulated in Table V. Only the Ga data in the magnetic recording layer are quantitatively calculated from its EDX full spectrum, while the other layers' data are relatively obtained from Ga-intensity ratios between the corresponding layers and the magnetic recording layer. Then, the experimental $C_{t,\mathrm{av}}$ values is as low as 1.8 at % for the magnetic recording layer (Fe-base alloy). Apart from some difference in the specimen's average atomic number, experimental $C_{t,\mathrm{av}}$ values are roughly equal to the calculated ones for the 50–100 nm thick TEM specimen composed of a W layer, i.e., 1.1~2.1 at %. However, the experimental C_t ratios (=4–6) between the magnetic and reflective/protective layers are different by factor of 2–3 from the calculated one (=2.0) between the W and Si layers.

As far as the authors know, only two experimental papers related on the C_{Ga} or $C_{t,\mathrm{av}}$ value have been published at present. One paper by Beag *et al.*[4] has shown $C_{t,\mathrm{av}}$ values of less than an EDX detection limit for the cross-sectional TEM sample of titanium nitride (TiN) layer on Si substrate. The other paper by Sakamoto *et al.*[5] has shown with secondary ion mass spectrometry (SIMS) using 30 keV Ga FIB that the C_{Ga} value is about 20 at % for a cobalt (Co) surface at $\beta = 80°$ (not the FIB-prepared TEM cross section). More experimental data are required to confirm firmly the present model's applicability.

V. SUMMARY

The Ga concentrations have been calculated for cross sectional specimens prepared using 30 keV Ga FIB. The cross sections are assumed to be formed under steady-state processes of ion implantation, backscattering, and sputtering. As to these process yields, Monte Carlo technique has been applied to simulate the trajectories of both ions and recoiled target atoms. The Ga concentration (C_{Ga}) value, which equals to the ratio of implantation yield (Y_{imp}) to sputtering yield (S), is valid with the specimen material. Then, each layer of the layered structure cross-section has its own C_{Ga} value. The calculation predicts that the Si and W layers are cross sectioned at grazing angles $\beta \approx 2.5°$ and 6°, respectively. The implanted Ga ions are concentrated very near the surface of <10 nm to yield the C_{Ga} values of about 4 at% for Si and 9 at% for W (after the preferential sputtering correction). Here, these $C_{\mathrm{Ga}}(=Y_{\mathrm{imp}}/S)$ values show weak β-dependence because both Y_{imp} and S decrease with the decrease of β. Although there is some difference in materials between the calculation and experiments, the calculated C_{Ga} values for Si and W layers are roughly equal to the experimental values for the magneto-optical TEM sample.

TABLE V. EDX data on the Ga concentration (corresponding to the $C_{t,\mathrm{av}}$ value) in each layer of the magneto-optical disk. The cross-sectional TEM images are shown in Fig. 8.

	UV coat layer	Reflective layer	Protection layer (2)	MO recording layer	Protection layer (1)	Substrate
Main element	C, O[a]	Al, Ti	Si, N	Fe, Co, Tb	Si, N	C, O[a]
Ga con. (at%)	0.2	0.5	0.5	1.8	0.3	0.4

[a]Polycarbonate.

[1]T. Ishitani and T. Yaguchi, Microsc. Res. Tech. **35**, 320 (1996).
[2]T. Ishitani, H. Tsuboi, H. Yaguchi, and T. Kamino, Scanning Microsc. (in press).
[3]T. Ishitani, Y. Taniguchi, S. Isakozawa, H. Koike, H. Yaguchi, H. Matsumoto, and T. Kamino, J. Vac. Sci. Technol. (in press).
[4]Y. W. Beag, M. Tarutani, K. Min, M. Kiuchi, and R. Shimizu, Jpn. J. Appl. Phys., Part 1 **33**, 2025 (1994).
[5]T. Sakamoto, M. Owari, and Y. Nihei, Jpn. J. Appl. Phys., Part 1 **36**, 1287 (1997).
[6]G. Betz and G. K. Wehner, in *Sputtering by Particle Bombardment II*, edited by R. Behrisch (Springer, Berlin, 1983), pp. 11–90.
[7]T. Ishitani, H. Tsuboi, H. Yaguchi, and H. Koike, J. Electron Microsc. **43**, 322 (1994).
[8]T. Ishitani, Jpn. J. Appl. Phys., Part 1 **34**, 3303 (1995).
[9]T. Ishitani and H. Tsuboi, Scanning **19**, 489 (1997).
[10]J. F. Ziegler, J. P. Biersack, and U. Littmark, *Stopping and Range in Matter* (Pergamon, New York, 1985).
[11]M. Nastasi, J. W. Mayer, and J. K. Hirvonen; *Ion-Solid Interactions: Fundamentals and Applications* (Cambridge University Press, Cambridge, 1996), pp. 80–83.
[12]J. P. Biersack and W. Eckstein, Appl. Phys. A: Solids Surf. **34**, 73 (1984).
[13]P. Sigmund, in *Sputtering by Particle Bombardment I*, edited by R. Behrisch (Springer, Berlin, 1981), pp. 9–71.
[14]D. B. Williams and C. B. Carter, *Transmission Electron Microscopy IV, Spectrometry* (Plenum, New York, 1996), pp. 631–634.

Widely changing probability of surface damage creation induced by a single ion in the MeV ion energy range

Hisato Ogiso[a] and Hiroshi Tokumoto
Joint Research Center for Atom Technology, National Institute for Advanced Interdisciplinary Research, 1-1-4 Higashi, Tsukuba, Ibaraki 305, Japan

Shizuka Nakano and Kazushi Yamanaka[b]
Mechanical Engineering Laboratory, 1-2 Namiki, Tsukuba, Ibaraki 305, Japan

(Received 11 November 1997; accepted 1 May 1998)

We report widely changing probability of surface damage creation induced by a single ion on highly oriented pyrolytic graphite specimens, each implanted with 3.1 MeV Si, Cu, As, Sr, Ag, or Au ions at a dose of 1.9×10^{11} cm^{-2}. By using a friction force microscope, we observed lattice disordered surface damage, and found the probability of surface damage creation varying from 0.02 to 0.54 depending on the variation in the ion species. To determine the reason for the larger dependence on ion species, we calculated the probabilities of knock-on atom generation by nuclear collision. The calculated probability of knock-on atom generation agreed well with the observed probability of surface damage creation. © *1998 American Vacuum Society.* [S0734-211X(98)03604-X]

I. INTRODUCTION

Ion beam processes such as ion implantation, ion beam enhanced deposition, and ion etching are widely applied in fields such as material modification, microtechnology, and electronic device fabrication. With ion beam processes, especially, single ion implantation proposed by Ohdomari and co-workers,[1] being increasingly applied to microtechnology and nanotechnology, an investigation of the effects of a single ion impact is expected to contribute significantly to knowledge about the actual conditions involved.

A scanning probe microscope (SPM), such as an atomic force microscope (AFM)[2] and a scanning tunneling microscope (STM),[3] provides an atomic resolution image, enabling defects and lattice disorders to be observed directly. In this respect SPM is useful for studying the effects on a single ion impact. Previous studies using SPM to detect tracks caused by a single ion have involved many ion species, such as B,[4] C,[5] N,[6] S,[6] Ar,[7,8] Kr,[9] Xe,[7,10] Au,[5,11] Pb,[9,10] and U.[9] The number density of surface damage has been reported to be almost identical to the number density of incident ions, and the clear dependence of the number density on ion species was not shown.

Against this background, we conducted a systematic investigation using a variety of ion species with an energy of 3.1 MeV, and observed a widely changing probability of surface damage creation on the graphite surface, ranging from 0.02 to 0.54, depending on the atomic number of ions, from 14 (Si) to 80 (Au). We used a friction force microscope (FFM).[12] At the same time, we carried out a simulation of knock-on atom generation probability by nuclear collision at the first and second atomic layer from the target surface, giving a remarkably good agreement to the measured probability of surface damage creation. Consequently, particular energy range of MeV was found to be responsible for the widely changing probability.

II. EXPERIMENTS

We used highly oriented pyrolytic graphite (HOPG) as the target material. The HOPG surface is suitable for observing change made by a single ion impact at atomic resolution, because STM or AFM lets the atomic image be observed readily. Ion irradiation was conducted on a tandem type accelerator (National Electrostatic Corp., 3SDH-4). A monochromatic beam was ensured using analyzing magnets and an electrostatic deflector to remove neutral particles. The beam was scanned two dimensionally: x and y directions, ensuring the dose uniformity in the irradiation area. The specimen was surrounded by negative bias electrodes (−300 V) to suppress secondary electron emission, making the current from specimen to ground equal to the ion beam current. Incident ion dose was measured by integrating the current. HOPG specimens 5×5 mm^2 were cleaved immediately before being placed in a vacuum target chamber at 10^{-5} Pa. The ion energy was 3.1 MeV, the current density was 8 nA/cm^2, and the dose was 1.9×10^{11} cm^{-2}. The specimens were subsequently observed in air using a SPM system (Seiko Instrument Inc., SPI3700). When the scanning direction was perpendicular to the cantilever, the lateral force between the tip and surface was detected by the torsion of the cantilever.

III. RESULTS AND DISCUSSIONS

Disordered lattice structure on Au ion irradiated HOPG surface in atomic resolution is shown in Fig. 1. The incident ion beam flux was very small, i.e., 5×10^{-4} ions/s/nm^2, so more than one ion rarely affected such a small area simulta-

[a]Electronic mail: hisato@nair.go.jp
[b]Present address: Department of Materials Processing, Graduate School of Engineering, Tohoku University, Aramaki, Aoba-ku, Sendai 980-77, Japan.

FIG. 1. FFM images (10 nm×10 nm) of 3.1 MeV Au ion-irradiated HOPG surface with two directions of scanning: (a) x direction (from the left to the right), (b) $-x$ direction (from right to the left).

neously. We therefore assumed that a single ion impact formed the disordered region of surface damage. We call the disordered region surface damage.

When the direction of scanning was reversed, the image contrast was inverted, as compared with Figs. 1(a) and 1(b). Figure 2 shows that the profiles of the cantilever-torsion signals at both scanning direction images. The torsion signal difference between the profiles on the surface damage was greater than that on intact surface. This indicates that the surface damage applied greater lateral force to the tip of the cantilever than the intact HOPG surface. Namely, a single ion impact can change the friction force in a nonometer-sized area. This finding suggests the possibility of applying a single ion beam process to nanoscale mechanical devices.

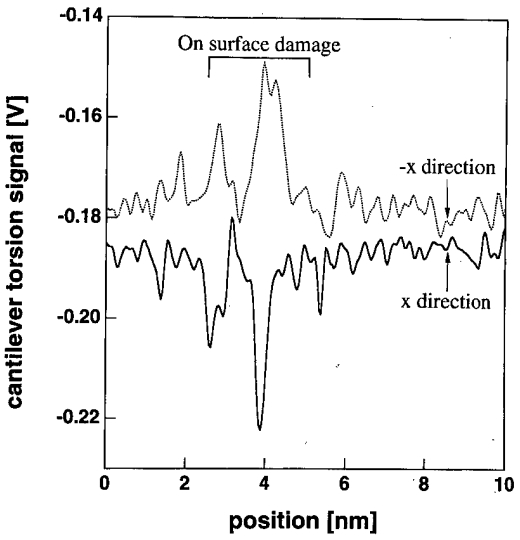


FIG. 2. Profiles of the cantilever-torsion signal with the x (solid line) and $-x$ (dashed line) directions.

Figure 3 shows the wide area images (500×500 nm^2), then the surface damage having the appearance of dots. Each surface of the images is implanted with different ion species. The number density of surface damage (Table I) apparently differ, even though the incident ion doses were all the same. We then obtained the probability of surface damage creation from the ratio of the number density to the incident ion dose, and plotted this as a function of the atomic numbers of incident ions (Fig. 4; large solid circles). We found that probability depends greatly on the ion species.

The total stopping process of ions in solids is divided into two parts; to the target nuclei (nuclear collision) and energy transfer from the ion to target electron (electronic collision). To determine which of the two was dominant in forming surface damage, we calculated the effect of the two numerically.

First, to estimate nuclear collision effect, we calculated the probability of knock-on atom generation by nuclear collision. For the calculation, we assumed that the nuclear col-

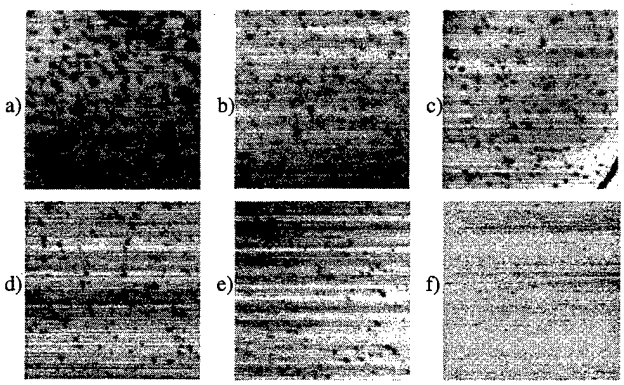


FIG. 3. FFM images (500×500 nm^2) of irradiated specimens showing the surface damage number dependence on ion species. Ion energy: 3.1 MeV, ion dose: 1.9×10^{11} cm^{-2}, ion species: (a) Au ion, (b) Ag ion, (c) Sr ion, (d) As ion, (e) Cu ion, (f) Si ion.

TABLE I. Observed number density of surface damage.

Ion species	Number density of surface damage (μm^{-2})	Surface damage creation probability
Si	44	0.023
Cu	380	0.20
As	528	0.28
Sr	824	0.43
Ag	876	0.46
Au	1052	0.55

lision is a two-body central-force scattering, and a knock-on atom is generated when the energy transfer E_t to the target atom is larger than displacement energy E_d required to produce a knock-on atom. E_t is a function of interatomic potential V, incident ion energy E_0, and impact parameter b, defined as the distance between the initial line of motion of the incident ion and the center of the target atom, i.e., $E_t = f(V, E_0, b)$. This function is calculated using the "magic formula" proposed by Biersack *et al.*[13] We then derive the region of b where $E_t > E_d$ is satisfied. The probability of knock-on atom generation is derived from the region of b and graphite lattice geometry (Fig. 5). We here give E_d to be 20 eV, and use universal potential [Eq. (1)][14] for the interatomic potential,

$$V(R) = \left(\frac{Z_1 Z_2 e^2}{R a_u}\right)\Phi(R),$$

$$\Phi(R) = 0.1818e^{-3.2R} + 0.5099e^{-0.9423R} + 0.2802e^{-0.4029R} + 0.02817e^{-0.2016R}, \quad (1)$$

$$R = \frac{r}{a_u}, \quad a_u = 0.8854\left(\frac{a_0}{Z_1^{0.23} + Z_2^{0.23}}\right),$$

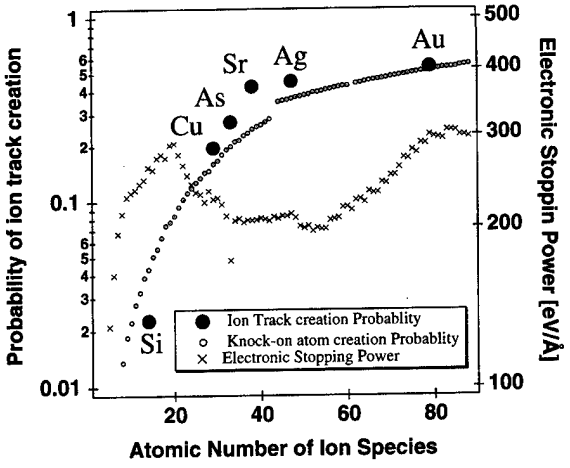


FIG. 4. Relationship between observed surface damage creation probability and atomic number of incident ions, showing calculated probabilities of knock-on atom generation by nuclear collision and electronic stopping power as a function of atomic number of incident ion atomic number, ● surface damage creation probability by nuclear collision (experiment), ○ knock-on atom generation probability by nuclear collision (calculated), × electronic stopping power (calculated).

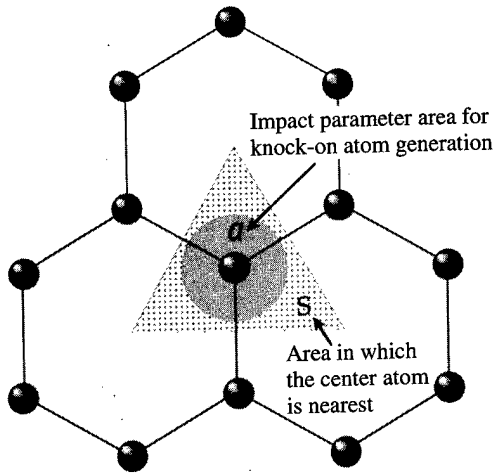


FIG. 5. Calculation of knock-on atom generation probability $P = a/S$.

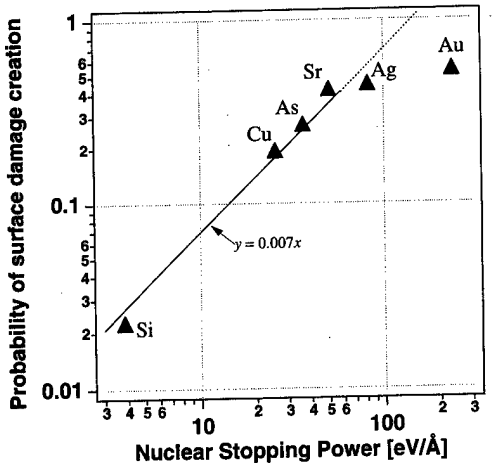


FIG. 6. Relationship between nuclear stopping power (calculated) and observed surface damage creation probability (experiment).

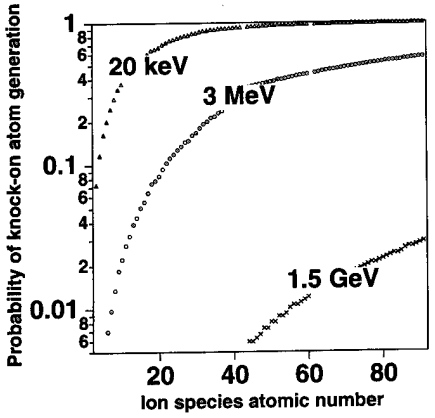


FIG. 7. Knock-on atom generation probability in 20 keV, 3 MeV, and 1.5 GeV ions as a function of ion species atomic number.

TABLE II. Calculated probability of knock-on atom generation, electronic stopping power, and experimental results of surface damage creation probability.

Energy range	Knock-on atom generation probability (calculated)	Electronic stopping power (eV/Å) (calculated)	Probability of surface damage creation (experimental results)	References
keV	~1	Small	~1	7,6,8
MeV	10^{-2}–10^{-1}	Small	10^{-2}–10^{-1}	This work
GeV	$\leqslant 10^{-2}$	Large	~1	9,10

where Z_1 and Z_2 are the atomic number, r is their separation, and a_0 is the Bohr radius (=0.529 Å). These assumptions are the same as those in TRIM,[13,14] the Monte Carlo simulation widely used to obtain the depth profile of implanted ions. Our numerical calculation was approximate, that is, only the primary knock-on atom was taken into account, and secondary knock-on atoms were not considered. Besides, we took the probability that was caused within the first and second atomic layers from the surface into account, because observed track images could be divided roughly into two clear and faint tracks. This calculated probability of knock-on atom generation was plotted as a function of the atomic number of ion species (Fig. 4), and agreed well with experimentally obtained surface damage creation.

Second, we examined electronic stopping power $-dE/dx$, electronic energy loss per unit ion path length, as the effect of electronic collision. Here we used the SRIM 98 (the latest version of TRIM) developed by Ziegler for calculation. We compared the calculated results (Fig. 4) with the experimental results of surface damage creation. Hence these comparisons indicate that the nuclear collision between an incident ion and target atoms is dominant in the surface damage creation. It was reported that electronic collisions created lattice defects in alkali halide crystals.[15] In our case, however, the electronic collision was not dominant.

To clarify the effect of the nuclear stopping on the surface damage creation, the relationship between the nuclear stopping power calculated by SRIM and the observed probability of surface damage creation is shown in Fig. 6. When the stopping power is small, the observed probability of surface damage creation is approximately proportional to the nuclear stopping power. When the stopping power is large, the probability tends to be saturated. This indicates that the dependence of number density of surface damage on ion species hardly appears, in case that nuclear stopping power is large. For comparison with other works, we calculated the knock-on atom generation with different ion energies using the same procedure (Fig. 7). In the case of 20 keV ions which corresponds to the work by Corteger *et al.*,[6] the probability of knock-on atom generation is nearly equal to 1 in almost all ion species. On the other hand, 1.5 GeV Xe ions, in case which Hagen *et al.* found surface damage under FFM,[10] should hardly leave the surface damage induced by nuclear collision, because they have a probability of less than 0.02. Nevertheless, the 1.5 GeV Xe ions have a large electronic stopping power of 1364 eV/Å—significantly larger than MeV and keV ions (3 MeV Xe: 198 eV/Å, 20 keV Xe: 30.3 eV/Å). The surface damage of GeV ions is therefore caused by electronic collision, while the surface damage of MeV and keV ions is caused by nuclear collision.

In the MeV range, the probability of knock-on atom generation varies over a wide range, depending on the ion species, and the probability of surface damage creation varies over a wide range, similar to the probability of knock-on atom generation (Table II). In the keV and GeV range, the effect of either nuclear or electronic collision is large enough that an ion always leaves a surface damage, respectively. In other words, number density does not depend on the ion species. The distinctive character of MeV ions is therefore responsible for the clear appearance of the dependence on ion species. This dependence finding is expected to contribute to the elucidation of the mechanism behind surface damage creation.

IV. CONCLUSION

We observed MeV-ion-irradiated HOPG using FFM, finding surface damage caused by single ion impact. The friction force between the surface on surface damage was caused by single ion impact. The friction force between the surface on surface damage and the tip of cantilever was larger than that on intact surface. The number density of surface damage and the tip of cantilever was larger than that on intact surface. The number density of surface damage apparently depended on the ion species. We calculated the probability of knock-on atom generation by nuclear collisions between an incident ion and target atom and the electronic stopping power to determine which collision process is responsible for the surface damage creation. The probability of knock-on atom generation varies over a wide range with the variation in the ion species, and shows a trend similar to the collision formed by the surface damage. We consider that the dependence on ion species appears only where the ion energy is at the MeV range.

ACKNOWLEDGMENT

This work was partly supported by the New Energy and Industrial Technology Development Organization (NEDO).

[1]M. Koh, K. Horita, B. Shigeta, K. Igarashi, T. Matsukawa, T. Tanii, S. Mori, and I. Ohdomari, Appl. Phys. Lett. **68**, 3467 (1996).
[2]G. Binnig, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

[3]G. Binnig, H. Rohrer, C. Gerber, and W. E., Phys. Rev. Lett. **49**, 57 (1982).
[4]H. Ogiso, S. Nakano, O. Kolosov, K. Yamanaka, and T. Koda, Proceedings of the Ninth Symposium on Surface Layer Modification by Ion Implantation, Tokyo, 1993 (unpublished), p. 57.
[5]H. Ogiso, S. Nakano, K. Yamanaka, and T. Koda, Proceedings of the First Meeting of the Ion Engineering Society of Japan, Tokyo, 1992 (unpublished), p. 153.
[6]R. Coratger, A. Claverie, A. Chahboun, V. Landry, F. Ajustron, and J. Beauvillain, Surf. Sci. **262**, 208 (1992).
[7] L. Porte, C. H. de Villeneuve, and M. Phaner, J. Vac. Sci. Technol. B **9**, 1064 (1991).
[8]T. Matsukawa, S. Suzuki, T. Fukai, and I. Ohdomari, Extended Abstracts of the 42nd Spring Meeting of the Japan Society of Applied Physics, 1995 (unpublished), p. 588.
[9]S. Bouffard, J. Cousty, Y. Pennec, and F. Thibaudau, Radiat. Eff. Defects Solids **126**, 225 (1993).
[10]T. Hagen, S. Grafström, J. Ackermann, R. Neumann, C. Trautmann, J. Vetter, and N. Angert, J. Vac. Sci. Technol. B **12**, 1555 (1994).
[11]J. Yan, Z. Li, C. Bai, W. S. Yang, Y. Wang, W. Zhao, Y. Kang, F. C. Yu, P. Zhai, and X. Tang, Nucl. Instrum. Methods Phys. Res. B. **75**, 1390 (1994).
[12]C. M. Mate, G. M. McCelland, R. Erlandsson, and S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987).
[13]J. P. Biersack and L. G. Haggmark, Nucl. Instrum. Methods Phys. Res. B **147**, 257 (1980).
[14]J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids* (Pergamon, New York, 1985).
[15]R. T. Williams, K. S. Song, W. L. Faust, and L. C. H., Phys. Rev. B **33**, 7232 (1986).

Evidence of depth and lateral diffusion of defects during focused ion beam implantation

C. Vieu,[a] J. Gierak, M. Schneider, G. Ben Assayag,[b] and J. Y. Marzin
L2MCNRS, 196 avenue Henri Ravéra BP 107, 92225 Bagneux, France

(Received 16 January 1998; accepted 15 May 1998)

We have investigated the evolution of the photoluminescence intensity of different GaAs/GaAlAs quantum wells to probe the defect penetration during focused ion beam implantation at various sample temperatures. By implanting homogeneous areas, it was possible to evaluate the depth extension of the ion-induced damage, while by localizing the implantation at a submicrometric length scale, it was also possible to quantify the lateral damage extension. Both the depth and lateral damage extensions are found to be reduced when the sample temperature is changed from 300 to 80 K. This result is used to demonstrate that a fast diffusion of nonequilibrium defects is taking place in the irradiated sample. It is found that this diffusion is highly anisotropic with a lateral diffusion length ten times higher than the depth one, suggesting that the defects easily diffuse in the GaAlAs barrier layers. The results are interpreted as a manifestation of a radiation enhanced diffusion mechanism. The advantage of using low temperature irradiations to better localize the injected defects by focused ion beam is highlighted, specially for quantum nanostructure fabrication. © *1998 American Vacuum Society.* [S0734-211X(98)07304-1]

I. INTRODUCTION

Dry etching is now a key process for micro- and nanofabrication of ultrasmall optoelectronic and electronic devices. However, nanometric etched structures are often limited by the damage induced by the energetic ions involved in the etching process. The damage issue has received increasing attention for several years specially for III–V semiconductor compounds. The numerous experimental results converge towards the demonstration that ion damage penetrates an order of magnitude deeper than what is expected from classical collision calculations in amorphous media like TRIM.[1] Both channeling effects taking place in the monocrystalline substrates used in the experiments and rapid defect diffusion during bombardment have been invoked to account for this discrepancy. Surprisingly, the same phenomenon has been observed for a large variety of irradiation conditions and ion incident energy range.

Laruelle *et al.*[2] have observed that ion implantation induced interdiffusion in GaAs/GaAlAs quantum wells (QWs) occurs down to a depth roughly ten times higher than the projected range of 150 keV Ga^+ ions implanted with a focused ion beam (FIB) system. The authors conclude that due to the channeling of some Ga^+ ions, defects are generated deeply in the structure and can promote the interdiffusion of remote QWs after annealing. In another energy range, Kasahara *et al.*[3] have observed that Ga^+ ions implanted at 1 keV with a low-energy FIB system generate defects far deeper than the ion range due to a possible diffusion of primary defects. Using a conventional Ar ion beam etching (IBE) system, Germann *et al.*[4] have shown that the defects induced by Ar^+ ions of 250 eV incident energy also penetrate around six times deeper than the projected range of the implanted ions. They also give some evidence that channelling of very low-energy ions can play a significant role in the large depth damage penetration. This latter point was confirmed by Stofell[5] through molecular dynamics simulations. More recently, Hu *et al.*[6–9] have demonstrated that for III–V semiconductor compounds both channeling effects and rapid defect diffusion must be considered to explain their whole experimental results obtained for Ar^+ ion beam bombardment at energies lower than 1 keV. They suggest that radiation enhanced diffusion (RED) could be the mechanism responsible for a defect diffusivity as high as 3×10^{-15} cm^2/s deduced from their experiments. This mechanism was also proposed by Nakaniski *et al.*[10] who have investigated the deactivation of Si donors in GaAs by Ar plasma induced defects through photoreflectance measurements.

All these studies were dedicated to the in-depth penetration of defects, but very little is known about the lateral extension of the damage profile which is the most interesting feature for nanofabrication. We do not know if the lateral damage extension is of the same order of magnitude as the in-depth one. Moreover, it would also be very interesting to investigate if the solutions proposed to limit the in-depth damage penetration such as the introduction of a superlattice to getter the defects[6,7] are also a benefit with respect to the lateral defect extension. One of the possible reasons for the lack of experimental results for lateral damage is that when using a broad beam etching system like in IBE experiments, lateral investigations need a sophisticated patterning of etch masks of different sizes on the sample surface. This kind of study can be considerably alleviated by using FIB systems where defect injection can be spatially localized with a nanometric accuracy. The lateral characterization of damage in etched structures is also nontrivial because when sidewalls are created, the free surfaces themselves strongly affect the optical and transport properties of the nanostructures due to

[a]Electronic mail: christophe.vieu@l2m.cnrs.fr
[b]Present address: CEMES/CNRS, 29 rue Jeanne Marvig BP 4347, 31055 Toulouse, France.

surface nonradiative centers and depletion layers.[11,12] The exact role of the damaged layers in the degradation of the properties of the etched structures can thus be difficult to evaluate. In this context, FIB implantation can be used to locally inject defects with a controlled dose without significantly etching the surface. The damage extension can then be followed in the absence of lateral free surfaces and all the degradation of the structure can be unambiguously attributed to the ion-induced damage. Previous work with FIB systems[2,13] has concentrated on the lateral extension of damage after an annealing treatment and the interdiffusion of QWs was used to detect the defect concentration. However, in these experiments the post-implantation damage profile is convoluted with the defect diffusion taking place during annealing.

The purpose of this work is to take advantage of the versatility of the FIB approach to fully investigate the spatial damage extension after implantation both in the depth and lateral directions.[14] The influence of temperature will be highlighted by comparing the effects of liquid nitrogen temperature (LNT) irradiations with room temperature (RT) irradiations. The experimental procedure to assess the damage extension is very classical and was earlier proposed by Wong *et al.*[15] to study RIE-induced damage in GaAs/GaAlAs heterostructures. The photoluminescence (PL) intensities of several QWs located at different depths in the sample are used as local probes for the defect generation. The decay of the PL intensity of each QW is assumed to be proportional to the defect concentration in the vicinity of the QW due to the presence of nonradiative centers. QWs are very sensitive probes since we can estimate that PL is sensitive to defect concentrations as low as 10^{15} cm^{-3},[16] but the exact nature of the defects itself remains unknown, since all the perturbations of the crystal lattice can affect the radiative process at the specific energy of the confined electronic levels of the QW. In the following, we will thus speak about damage in a very general way without more precision about the exact configuration of the defects involved in the trapping process.

After the presentation of the experimental setup, the investigation of the depth penetration of damage will be presented for different ion species, ion energies and irradiation temperatures. The transverse characterization of the ion-induced defect profile at different temperatures will then be studied by defining irradiation-free stripes of different widths. We will show that surprisingly the lateral extent of damage is much higher than the in-depth one. The main results will be discussed in terms of the rapid anisotropic diffusion of nonequilibrium defects during the irradiation.

II. EXPERIMENTAL SETUP

The FIB instrument used for the experiments described in this work was designed and fabricated in our laboratory.[14] In this system, the ion column is mechanically attached to the chamber of a JEOL 840 scanning electron microscope (SEM). The two beams (electron and ion beams) are crossing with an angle of 83° on the sample. This dual beam combination allows three complementary working modes: (i) Electron scanning microscopy for nondestructive observation of the area of interest and for positioning the sample under the ion patterning field with an excellent accuracy; (ii) scanning ion microscopy for adjusting the ion column parameters such as focus and astigmatism in a region close to the structure to be fabricated; (iii) computer controlled electrostatic motion of the ion spot for irradiating the sample with a given dose across an arbitrary pattern. For these experiments the SEM electron beam was systematically switched off during FIB irradiations, in order to avoid surface contamination and a possible enhancement of defect diffusion induced by electron bombardment.

To carry out the experiments described in the present article we have first used a standard gallium liquid metal ion source (LMIS) of the very well known needle type, see, e.g., Ref. 17. In order to investigate an ionic species with a lower atomic weight, compared to gallium, we have also developed an aluminum LMIS. Aluminum was chosen, because its atomic weight is half of gallium, also for his sufficiently low vapor pressure, and finally because it is not a dopant for GaAs/GaAlAs heterostructures. On the other hand, an important difficulty arises from the fact that aluminum is highly corrosive or reactive when melted, so usually refractory metals used for LMIS fabrication must be excluded. The new patented ion source geometry we have designed and realized is graphite based.[18] This source is able to deliver an aluminum beam with emission quality standards very close to those of the gallium type, and of course, is compatible with our FIB system.

A home-built cryogenic cooled stage was recently developed and added to our focused ion beam system for low temperature FIB microfabrication experiments and rapid "*in situ*" characterization.[14] This equipment allows a wide variation of the sample temperature. The minimum temperature achievable is about 80 K when liquid nitrogen is used.

A $GaAs/Ga_{0.67}Al_{0.33}As$ MQW structure grown by MBE was used throughout this study. The sample structure illustrated in Fig 1, includes three GaAs quantum wells with different thickness located respectively at 30, 65 and 105 nm from the surface. That way we have three independent sensors for defect generation, located at different depth exhibiting each one a specific PL energy peak, respectively 1.59, 1.69 and 1.8 eV. This structure was prepared with growth interruptions at each interface. In this case, the PL lines are split into doublets or triplets, related to monolayer thickness variations of the QWs.

Samples are processed using conventional photolithography and chemical etching. The etching solution used here was $H_3PO_4/H_2O_2/H_2O$ 3:1:40 with a measured etching rate about 100 nm/min. This process was carried out to define a matrix of independent mesa structures. The size of each mesa is about 80×80 μm^2 for 2 μm depth, the distance between two adjacent mesa structures is 800 μm. This process allows specific irradiation conditions to be performed independently on each mesa, and limits the influence of adjacent mesas during PL characterization.

Finally, the ion-induced damaged region profile was mea-

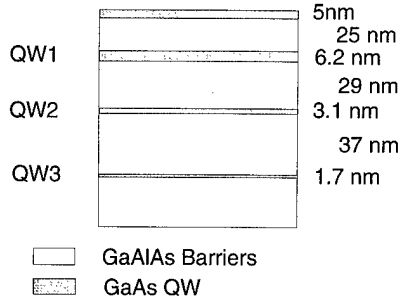


FIG. 1. Description of the multiquantum well sample structure, the aluminium concentration in the barriers is 0.33.

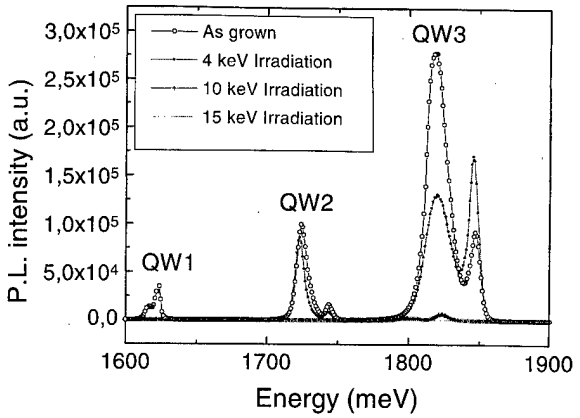


FIG. 2. Influence of a FIB irradiation on the PL emission intensity (1.7 K) of the MQW sample (gallium irradiation, room temperature, ion dose 10^{15} ions/cm^2).

sured directly after FIB irradiation without thermal annealing. Two different PL setups were employed. In a first experiment carried out at low temperature (1.7 K), PL characterization was employed to measure the degree of integrity of the different QWs. The laser beam (Ar) was focused in a 40 μm spot trough a transparent window on the cooled sample and we can consider that most of each mesa is probed during the analysis. In a second step, we have used spatially resolved PL. The focused laser beam (2 μm in diameter) is scanned across the sample with a piezoelectric driven stage, and PL signal is collected in relation with the focused laser probe position.

III. DEPTH CHARACTERIZATION OF THE ION-INDUCED DEFECT PROFILE

A. Influence of incident energy

Because ion energy is a very easily adjustable parameter in case of FIB microfabrication experiments, we performed a systematic evaluation of the longitudinal ion-induced defect extension for several conditions. We first investigated the effect of a FIB Ga^+ focused ion beam irradiation with variable beam energy on the PL signal emitted by the multiquantum well structure (MQW) illustrated in Fig. 1.

In this experiment the ion dose was fixed to a value of 10^{15} ions/cm^2 for which sputtering effects are negligible. The gallium FIB beam energy was varied from 2 to 15 keV, from mesa to mesa. As described above, each mesa was treated in an independent way and received a uniform irradiation at a fixed ion energy. For each energy, a precise measurement of the size of the patterning field and the ion spot current was conducted in order to adjust accurately the dwell time which in turn controls the total ion dose. A special care was also taken for positioning the ion irradiation area all over the mesa structure for ensuring that the ion irradiation process was properly done. This was greatly facilitated by using the *in situ* SEM imaging capability before and after FIB irradiation. Typical low temperature (1.7 K) PL spectra are reported in Fig. 2. FIB irradiated samples without annealing exhibit a strong loss of PL intensity that we attribute to non radiative centers generated by the ion bombardment in the vicinity of the quantum wells. When the incident ion energy is increased, the PL properties of the QWs are progressively affected in relation to their depth position. The damaged region increases with the ion incident energy. The typical evolution of PL signals can be illustrated through three representative ion energies. At 4 keV, only QW1 which is located 30 nm below the surface becomes not detectable, while the integrated PL intensity of QW2 and QW3 remains practically unchanged. We can say that the critical energy for the quenching of the PL intensity of QW1 is 4 keV for Ga^+ ions. For a given condition of laser illumination and a given signal detection system, the complete quenching of the PL signal occurs for a given concentration of defects. This critical concentration of defects is not precisely known but can be compared from one mesa to another if the illumination and detection conditions are identical. In our experiments special attention has been paid to illuminate and collect the PL intensity of the different QWs in exactly the same conditions, and we can thus assume that the quenching of the PL intensity is always representative of the same critical concentration of defects. For a 10 keV irradiation, QW1 and QW2 do not exhibit any PL signal while the PL intensity of QW3 is significantly affected. We can thus conclude that the critical concentration of defects responsible for the quenching of the PL has reached QW2 located 65 nm below the surface for an incident Ga^+ energy of 10 keV. Finally, an energy of 15 keV is needed for completely quenching the PL intensity of QW3 located 105 nm below the surface.

The damage profile induced by irradiation can be described using the concept of isodamage contours which are the lines connecting all the points of the sample which have received the same amount of damage.[19] Our experiment suggests that the isodamage contour corresponding to the critical concentration of defects required to quench the PL signal penetrates deeper and deeper with the ion incident energy. This finding is rather intuitive but what is much more surprising is that if we calculate the isodamage contours using conventional collisional calculations in amorphous target,[1,19] the isodamage corresponding to the quenching of the PL intensity should correspond to a defect concentration of 10^{12} cm^{-3}. This result has no physical meaning because it is three order of magnitude lower than the maximum sensitivity of PL measurements and also because this value is compa-

TABLE I. Critical ion energy for Ga^+ ions required to completely quench the PL signal of different QWs with an ion dose of 10^{15} ions/cm^2. The projected range was calculated with TRIM at the corresponding critical ion energy. Also indicated are the depth locations of the different QWs.

Quantum well No.	Depth (nm)	Critical ion energy (keV)	Projected range (nm)
1	30	4	4.8
2	65	6	6.2
3	105	15	11.5

rable with the equilibrium defect concentration at RT.[20] This shows that ion-induced damage extends much deeper that what is expected from conventional calculations which neglects both channeling and diffusion effects. However, one must keep in mind that the final extension of damage remains correlated to the ion incident energy which means that the primary collisional events have an influence on the final damage extension. These results are summarized in Table I, where the depth corresponding to the critical concentration of defects is compared with the projected range of the implanted Ga^+ ions as calculated by TRIM. We can see that whatever the incident energy the damage extension is roughly ten times higher than the projected range.

Samples irradiated with ion doses corresponding to the quenching of the PL signal have been observed in cross-sectional transmission electron microscopy. No extended defects could be observed inside the heterostructure. This result indicates that the defects responsible for the quenching of the PL signal are point defects.

B. Influence of the ionic species

In this part, we try to evidence the effect of ion mass in relation to the depth extension of the created damage. We have thus performed Ga^+ and Al^+ focused ion beam irradiations on the multiquantum well structure. The ion dose was kept constant at 10^{15} ions/cm^2, and the incident ion energy was varied from 2 to 15 keV, for the two ionic species. The PL characterization conditions are the same as previously described.

One typical result is presented in Fig. 3. It is shown that for the aluminum irradiation, the quenching of QW3 PL signal (105 nm depth) occurs for an ion energy about 6 keV, while an energy of 15 keV was required for Ga^+ implantations (Fig. 2).

The critical isodamage contour reaches the position of QW3 (105 nm) for an incident energy of 15 keV for Ga^+ irradiation, ($Rp = 11.5$ nm), and 6 keV in the case of Al^+ irradiation ($Rp = 10.4$ nm). Once again for both ion species we found that the depth extension of damage is ten times higher than the projected range. Moreover, for two different ion species with comparable projected range, we found the same depth extension of damage. This point confirms again that the primary collisional events have a strong influence on the final damage profile, even if they cannot account for the long range defect penetration.

According to previous works,[2–10] the long depth extension of ion-induced damage could be explained by channeling effects or diffusion mechanisms which are not taken into account in conventional ion range calculations. For our experimental conditions, it appears difficult to attribute the long range damage to primary channeled ions because of the reproducibility of our results. Indeed, though in our experiments the ion beam direction is oriented around the (001) direction of the substrate, the critical angle for dechannelling at 10 keV can be estimated in the range of 2 °–3 ° which is much less than for very low energy (<1 keV) ions. Our sample was glued on a conventional SEM sample holder without specific attention for the angular accuracy. If primary channeling was the main cause of damage depth penetration, we should have observed nonreproducible results. Moreover, the sample surface is covered with an amorphous oxide layer of around 2–3 nm which also limits primary channelling. The occurrence of channelling into ⟨011⟩ directions by large angle scattering events, as suggested by Stoffel[5] cannot be completely excluded. However, we think that this effect should not be very important in our experiments. For our typical energies around 10 keV, the effectiveness of large angle scattering into channels is reduced compared to sub keV experiments. Moreover, the ion dose used in our experiments is very low (1×10^{15} ions/cm^2) compared to other experiments with IBE or plasma systems where the typical ion doses were close to 10^{17} ions/cm^2. If we consider that 0.1% of the ions are injected into channels after large angle scattering, which is probably largely overestimated, we obtain an effective ion dose at the position of the QW around 10^{12} ions/cm^2. In previous works dedicated to ion implantation in QWs,[21] we have shown that after broad beam Ga^+ ion implantation at a dose of 1×10^{12} ions/cm^2 for energy conditions where the maximum of damage as calculated through conventional methods is centered on the QW depth position, the PL signal is not significantly affected. Finally, we have performed a control experiment where the sample was covered with a 50-nm-thick amorphous AlF_3 layer acting as an efficient dechanneling layer. After FIB irradiation with Ga^+ ions of 25 keV incident energy at a dose of 1×10^{15} ions/cm^2, we observe a quasi complete quenching of QW1

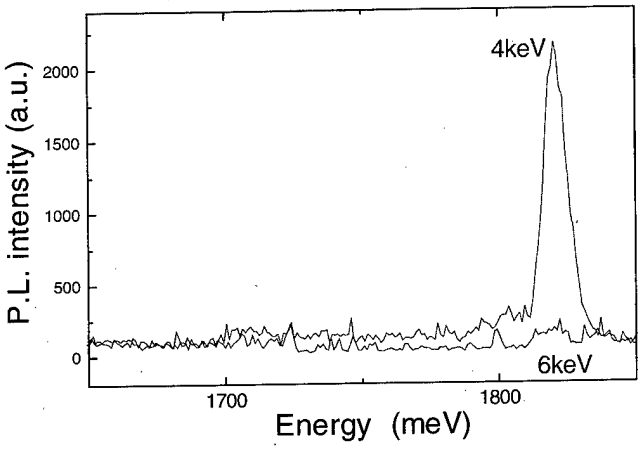


FIG. 3. Low temperature (1.7 K) PL signal emitted by QW3 for aluminium irradiations at room temperature for two incident energies.

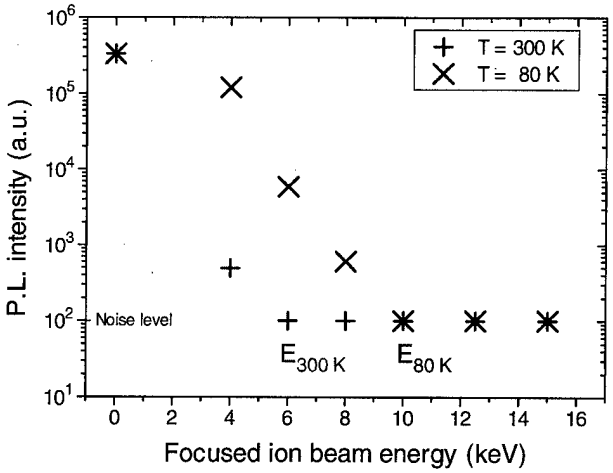


FIG. 4. Low temperature PL intensity of QW2 (65 nm depth), as a function of Ga^+ ion energy for different sample temperatures.

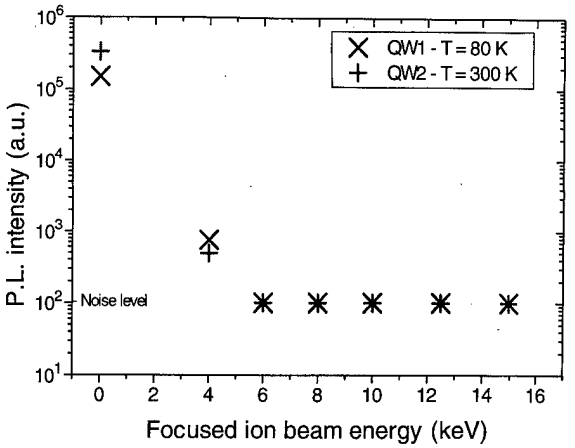


FIG. 5. Evolution of the low temperature (1.7 K) PL intensity as a function of Ga^+ ion incident energy. For QW2, located 65 nm below the surface, the irradiation was carried out at RT while for QW1, located 36 nm below the surface the sample was cooled at 80 K.

located 75 nm below the surface of the sample. For this irradiation condition, the projected range is 23 nm with a straggling of 8 nm which means that most of the ions are implanted in the AlF_3 layer. We thus conclude that even in an experimental condition where channelling is negligible the damage depth extension is much greater than what can be calculated with range calculations in amorphous media.

We thus believe that the main mechanism leading to the deep penetration of defects in our experimental conditions is an ion-induced collisional mechanism followed by rapid defect diffusion. In order to clarify this point, we have performed LNT irradiations to demonstrate that the defect penetration is strongly reduced at low temperature. If channeling was the main reason of long range disorder, we would have expected an increased defect penetration at low temperature because channeling can be affected by the lattice thermal vibrations.

C. Influence of the temperature of irradiation

In order to evidence a diffusion mechanism in the resulting defect profile, we have compared the effect of a Ga^+ focused ion beam irradiation on the MQW structure, which was kept either at room temperature, or near the liquid nitrogen temperature (80 K) during the bombardment. The surface ion dose was kept constant at 10^{15} ions/cm^2, and several energy steps were investigated in the range from 2 to 15 keV.

Figure 4 presents the typical evolution of the luminescence intensity for the quantum well located at 65 nm below the surface (QW2) as a function of the incident ion energy and for the two temperatures investigated. It appears clearly that the critical ion energy for the quenching of the PL emission, varies with the sample temperature. As seen previously the critical isodamage contour reaches the position of QW2 (65 nm) for an ion energy of 6 keV at room temperature. For a sample kept at low temperature (80 K), the extinction of the PL signal occurs for an incident energy of 10 keV. These results clearly show that the depth defect extension is lowered when the ion beam irradiation is operated on a sample kept at 80 K due to the reduced diffusion of the defects, even if the sample is subsequently warmed at RT. The defects involved in this diffusion are nonequilibrium defects mobile during the irradiation time which should be related to the dwell time of 10 μs per point for an ion dose of 1×10^{15} ions/cm^2.

By comparing the behaviors of QW1 and QW2, it is now possible to deduce the relative depth extension of the defect profile between 300 and 80 K by using the photoluminescence signal emitted both by QW1 and QW2 when irradiated at different temperatures. Figure 5 presents the evolution of the luminescence intensity for the quantum well located at 65 nm below the surface (QW2) as a function of the incident ion energy and for a RT implantation together with the PL evolution of QW1 located at a depth of 36 nm when irradiated at 80 K. In both cases the quenching of the PL signal occurs for an ion energy of 6 keV despite the fact that the quantum wells are separated by 29 nm. It is then possible to conclude that the critical isodamage contour corresponding to the quenching of the PL reaches 65 nm for a room temperature irradiation while it extends only to 36 nm for a LNT irradiation. The relative depth extension of damage between 300 and 80 K is then simply given by the distance between the two QWs of the structure: 29 nm.

IV. LATERAL CHARACTERIZATION OF THE ION-INDUCED DEFECT PROFILE

The original experimental FIB strategy we have used to evaluate the lateral extension of the ion-induced defect profile consists of defining irradiation-free stripes of different widths, surrounded by uniformly irradiated regions. The irradiations were performed at room temperature and at low temperature (80 K), using the same 10 keV Ga^+ focused ion beam probe and an identical surface deposited ion dose of 10^{15} ions/cm^2. It is important to keep in mind that for this energy QW2 is severely damaged for both room temperature and low temperature irradiation conditions. The width of the

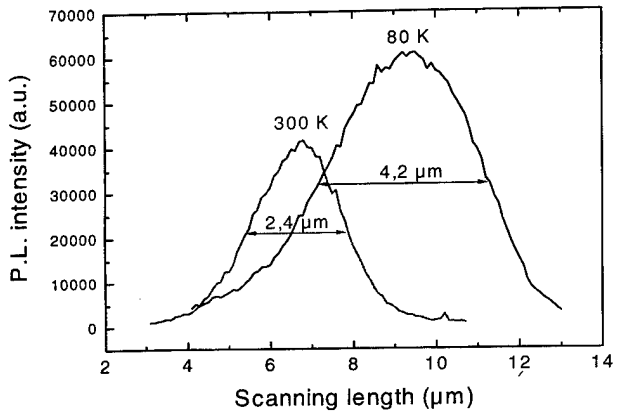


FIG. 6. Comparison of the effective emitting width of two stripes obtained for different irradiation temperatures (QW2 65 nm depth, 10 keV Ga^+, 10^{15} ions/cm^2). The curves represent the evolution of the low temperature (4 K) PL signal of QW2 when the exciting laser probe is scanned across an unimplanted window of 5 μm width.

virgin stripe W determined by the focused ion beam position, is varied from mesa to mesa from 5 to 0.05 μm. The effective width of this defect-free stripe is reduced compared to W due to different phenomena. First, the current profile in the ion spot, which is assumed to be Gaussian has a diameter which can be estimated to 70 nm for our 10 keV Ga^+ beam. In comparison, the lateral dispersion of the defects calculated from collisional considerations can be neglected in our energy range, because it is less than 5 nm. On the other hand, the effective width of the stripe can be reduced due to the possible diffusion of the defects in the lateral direction during the bombardment. This latter effect is the object of interest of this experiment. Of course, it is very difficult to measure the exact current distribution within the ion spot and the first factor of reduction of the defect-free stripe is not precisely known. However, by comparing RT and LNT irradiations performed in exactly the same conditions, we can deduce the relative lateral extent of the damage profile between RT and LNT, which is not dependent on the ion optics. In order to perform this comparison two samples have been irradiated using the same FIB operating conditions and in a very short time interval (less than 2 h), the only varying parameter in this experiment being the lateral diffusion length of the ion-induced defects. The effective width of the FIB patterned stripes was then measured with a spatially resolved low temperature (4 K) PL experiment. A finely focused laser beam of 2 μm diameter is scanned across the stripe using a piezoelectric driven stage, and photoluminescence signal is collected in relation with the focused laser probe position.

Figure 6 shows two different scans of the laser probe across an unimplanted 5 μm computer-defined wide stripe, for both RT and LNT irradiation temperatures.

The FWHM of the emitting area, measured on the PL signal appears significantly larger in case of a low temperature irradiation (4.2 μm) compared to the room temperature irradiation conditions (2.4 μm), and is much closer to the computer defined width W (5 μm). This result seems to indicate that lateral defect diffusion is higher at RT than at 80 K. The PL measurements we have obtained at RT and LNT for different widths of the unimplanted stripe: W are summarized in Table II.

TABLE II. Effective emitting widths of the unimplanted stripes for different irradiation temperatures and computer predefined widths W.

	Experimental emitting width $T=300$ K	Experimental emitting width $T=80$ K
$W=5$ μm	2.4 μm	4.2 μm
$W=2$ μm	0	2.7 μm
$W=1$ μm	0	1.8 μm
$W=0.8$ μm	0	1.8 μm
$W=0.6$ μm	0	1.8 μm

For room temperature, no emitting region can be detected as soon as W is equal to 2 μm or less. At low temperature (80 K), down to $W=0.6$ μm, an emitting region is preserved between the implanted areas. The experimental width of the emitting area as given, for example, in Fig. 6, is the convolution of the profile of the emitting region of the sample with the laser spot profile. For W below 1 μm, the real width of the emitting region is much smaller than the laser spot size used in the PL characterization and thus the effective width of the emitting region appears to be constant. This allows us to deduce the diameter of our focused laser spot, which can be estimated around 1.8 μm. These results clearly indicate that when the irradiation is carried out at 80 K, the lateral extent of the damage profile is considerably reduced. For an unimplanted width defined by the scanning of the FIB spot of 2 μm, the isodamage corresponding to the quenching of the PL reaches QW2 in the depth direction and completely overlaps in-between the irradiated sections of the sample when the bombardment is realized at RT. In contrast, for a LNT irradiation this isodamage also reaches QW2 in the depth direction but does not overlap laterally, and a region with a lower concentration of defects is left between the implanted areas giving rise to a detectable PL signal.

In order to extract quantitative information from our results, we have developed a simple model of damage extension. First, we suppose that the spatial profile of the damage created by a single ion incoming on the surface can be described by a Gaussian expression as

$$C(z,r)=\frac{C_p(z)}{2\pi\sigma_t^2}e^{-r^2/2\sigma_t^2}, \tag{1}$$

where z denotes the depth and r the radial distance. C_p represents the projected damage profile which takes into account the long range in-depth penetration of the defects. σ_t is the lateral dispersion of the damage profile including the effects of the FIB ion spot, the straggling due to collisions and the lateral diffusion of the defects:

$$\sigma_t^2=\sigma_{\text{fib}}^2+\sigma_{\text{def}}^2+4\delta^2. \tag{2}$$

Here, the current profile within the ion spot is assumed to be Gaussian with a standard deviation σ_{fib}, σ_{def} is the lateral

straggling of the defect distribution as conventionally calculated in amorphous targets without any diffusion and δ is the lateral diffusion length of the defects during the irradiation which is the quantity of interest. After implantation of broad pads with an ion dose D separated by a width W, the damage profile can be obtained by the convolution of the profile due to a single ion given in Eq. (1) with an injection function describing the pattern scanned with the ion beam.[22] We obtain

$$C(z,r)=D\cdot C_p(z)\left[1+\frac{1}{2}\mathrm{erf}\left(\frac{r-w/2}{\sigma_t}\right)-\frac{1}{2}\mathrm{erf}\left(\frac{r+w/2}{\sigma_t}\right)\right], \tag{3}$$

where the origin of the radial distance is at the center of the unimplanted window. Following Germann *et al.*[4] we can relate the ratio between the PL intensity of a damaged QW located at a depth z_{qw} with the intensity of a virgin reference part of this QW as follows:

$$\frac{I_{QW}(r)}{I_{ref}}=1/[1+\alpha C(z=z_{QW},r)], \tag{4}$$

where α is a constant. In our experiments we have measured the PL of QW2 located 65 nm below the surface after an implantation of Ga^+ ions of 10 keV at a dose of 1×10^{15} ions/cm^2. For a homogeneous implant, the PL signal is quenched which corresponds roughly to a ration $I_{qw}/I_{ref}=10^{-4}$. The defect concentration after a homogeneous irradiation is expressed by

$$C(z)=D\cdot C_p(z). \tag{5}$$

Combining Eqs. (4) and (5) we find

$$1+\alpha D C_p(z_{QW2})=10^4. \tag{6}$$

The lateral profile of the PL intensity of QW2 through the unimplanted window of width w is then roughly given by

$$\frac{I_{QW2}(r)}{I_{ref}}=\frac{1}{1+10^4\left[1+\frac{1}{2}\mathrm{erf}\left(\frac{r-w/2}{\sigma_t}\right)-\frac{1}{2}\mathrm{erf}\left(\frac{r+w/2}{\sigma_t}\right)\right]}. \tag{7}$$

Finally, the experimental profile obtained by scanning the laser probe across the unimplanted window is given by convolution with the profile of the laser exciting spot which is assumed to be Gaussian with a standard deviation ϕ:

$$\frac{I^{exp}_{QW2}(r)}{I_{ref}}=\frac{I_{QW2}(r)}{I_{ref}}\otimes\frac{1}{\sqrt{2\pi}\phi}e^{-r^2/2\phi^2}. \tag{8}$$

ϕ is experimentally obtained when a very small emitting area (much smaller than ϕ) is probed in spatially resolved PL. As seen in Table II this case is verified for LNT irradiations with $W<1$ μm. We obtain a reasonable value of $\phi=800$ nm which corresponds roughly to a laser spot of 2 μm in diameter.

We have used Eqs. (7) and (8) in order to extract σ_t, the lateral extension of damage as a function of temperature. The best fit has been obtained for $\sigma_t(\mathrm{LNT})=150$ nm$\pm$25 nm and $\sigma_t(\mathrm{RT})=540$ nm$\pm$30 nm.

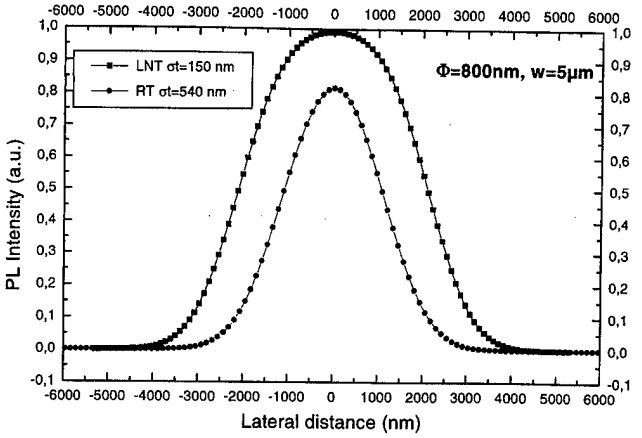


FIG. 7. Calculated effective emitting widths of two stripes obtained for different irradiation temperatures (QW2 65 nm depth, 10 keV Ga^+, 10^{15} ions/cm^2). The curves represent the calculated evolution of the PL signal of QW2 when the exciting laser probe is scanned across an unimplanted window of 5 μm width. The lateral extension of damage at RT is $\sigma_t(\mathrm{RT})=540$ nm and $\sigma_t(\mathrm{LNT})=150$ nm at LNT.

Figure 7 presents the calculation of the PL scanning profile for a window width $W=5$ μm. The experimental features of Fig. 6 are perfectly reproduced by the calculation. Figure 8 presents the calculation for $W=2$ μm. As expected experimentally, the calculation predicts negligible intensity for the RT implantation while intense PL emission is still recorded for the LNT case. The whole experimental cases as those summarized in Table II can be very well described by our calculation using this set of values for σ_t. The large margins on the determination of σ_t are due to the dispersion observed on the numerous experimental profiles recorded for each value of W.

Using Eq. (2) it is now possible to have a direct access to the difference in the lateral diffusion length of the defects for the two investigated temperatures. Indeed, it appears that

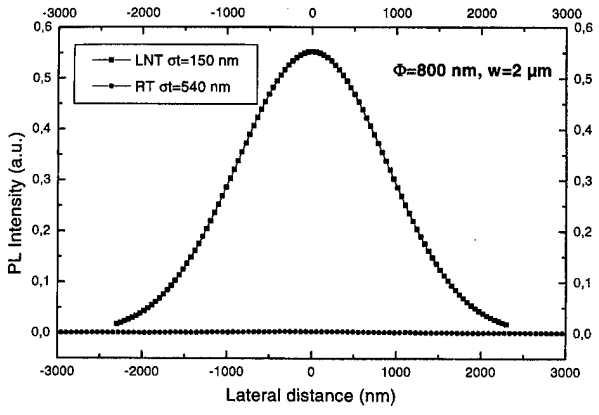


FIG. 8. Calculated effective emitting widths of two stripes obtained for different irradiation temperatures (QW2 65 nm depth, 10 keV Ga^+, 10^{15} ions/cm^2). The curves represent the calculated evolution of the PL signal of QW2 when the exciting laser probe is scanned across an unimplanted window of 2 μm width. The lateral extension of damage at RT is $\sigma_t(\mathrm{RT})=540$ nm and $\sigma_t(\mathrm{LNT})=150$ nm at LNT. There is no more emissive region for a RT implantation.

$$\sigma_t(\mathrm{RT})^2-\sigma_t(\mathrm{LNT})^2=4[\delta(\mathrm{RT})^2-\delta(\mathrm{LNT})^2]. \quad (9)$$

We thus obtain: $[\delta(\mathrm{RT})^2-\delta(\mathrm{LNT})^2]^{1/2}=260$ nm. Thus the relative lateral extension of damage between 80 and 300 K is 260 nm. This value does not depend on the spot size of the FIB system nor on the lateral straggling of defects due to collisions, since in the difference expressed by Eq. (9) the quantities σ_{fib} and σ_{def} disappear. Comparing with our results obtained for the depth extension of damage (Sec. III), it also appears that the relative lateral extension of damage is ten times higher than the relative depth extension. The fast diffusion of nonequilibrium defects during irradiation is highly anisotropic with enhanced diffusion in a plane parallel to the surface.

A good estimation of σ_{fib} can be obtained by etching experiments with the ion probe, we found a value around $\sigma_{\mathrm{fib}}=35$ nm for an incident energy of 10 keV. On the other hand, σ_{def} can be estimated through conventional Monte Carlo simulations to be $\sigma_{\mathrm{def}}=5$ nm. Using these values, we obtain: $\delta(\mathrm{RT})=270$ nm and $\delta(\mathrm{LNT})=70$ nm. The lateral defect diffusion length is thus strongly reduced at 80 K compared to 300 K but it still remains non-negligible. Either this value really reflects the low temperature diffusion length of nonequilibrium defects during irradiation or it can also be due to a possible diffusion of equilibrium defects when the sample is warmed up from 80 to 300 K after the irradiation.

V. DISCUSSION

We have shown that the temperature of the sample has a strong influence on the damage extension induced by FIB irradiation. These results demonstrate that the main mechanism responsible for the poor localization of the injected defects by FIB both in the depth and lateral directions is a fast diffusion of nonequilibrium defects during the irradiation time. This fast diffusion can be possibly explained through a RED mechanism. Many electron-hole pairs are generated during FIB implantation and their nonradiative recombination can provide an additional energy to localized defects for diffusing. Furthermore, our experimental results indicate that defect diffusion is highly anisotropic with an enhanced diffusion length in the plane of the QW layers. This anisotropy suggests that the defects are diffusing rapidly in the GaAlAs barrier layers while their diffusion is reduced in GaAs QW layers. This is in good agreement with the results of Green *et al.*[6] who have shown that GaAs/GaAlAs superlattices could be used to getter the defects in order to reduce their depth penetration. This is also a good support for a RED mechanism since nonequilibrium defects are found to diffuse easily in the higher band gap material. However, these new results on the lateral damage extension point out that the use of a superlattice to reduce the depth penetration of defects can be deleterious for the lateral selectivity of the defect injection.

From our results it clearly appears that the main limiting phenomenon when attempting to realize FIB lateral patterning on III–V heterostructures is the fast diffusion of ion-induced defects. This diffusion can be drastically reduced if we use low irradiation temperatures. The difference between the damage extension at 300 and 80 K is as high as 29 nm in the depth direction and 260 nm in the lateral direction. This very important unexpected finding clearly evidences the advantage of low temperature implantations, when damage injection must be highly localized for advanced nanostructure fabrication for example.

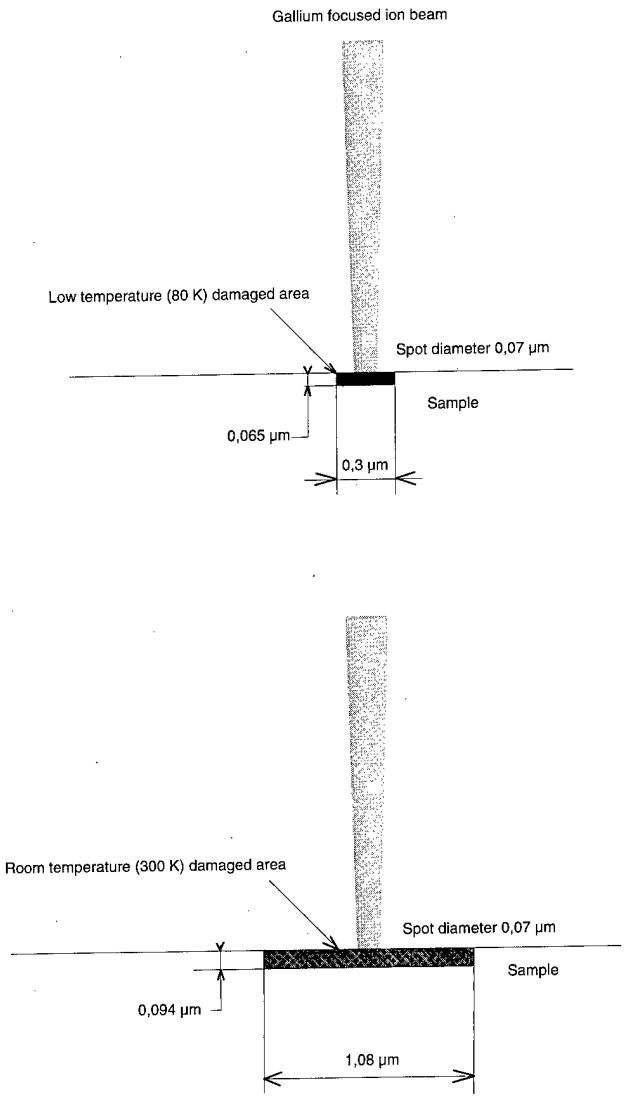


FIG. 9. Schematic variation of a 10 keV gallium FIB induced defect profile, vs sample temperature during ion irradiation.

Figure 9 is a schematic representation of the difference in the defect localization under FIB irradiation for the two sample temperatures. The LNT diffusion length can be estimated to be around 70 nm. This latter value is not negligible and it is still permitted to think that a further decrease in sample temperature from 80 to some 20 K would continue to improve defect localization. New experiments are currently in progress at a sample temperature of 25 K.[23] Another experiment is being implemented to check that a RED mechanism is really taking place in our samples. Using our coupled FIB/SEM system it is indeed very easy to bombard the irradiated surface with an electron beam during the FIB irradia-

tion. According to the RED mechanism this would lead to an enhancement of the fast diffusion of nonequilibrium defects.

VI. CONCLUSION

Experimental evidence of the fast diffusion of defect during FIB implantation has been given. From a careful analysis of PL results it was possible to quantify the damage extension due to diffusion at room temperature both in-depth and in the lateral directions. Values as large as 29 and 260 nm for the depth and lateral directions, respectively, have been obtained. It clearly appears that defect diffusion in III–V heterostructures occurs much more efficiently in the lateral direction than in depth. The ratio of diffusion anisotropy is as large as 10 which suggests that diffusion is enhanced along the heterointerfaces. This property can be explained by a radiation enhanced diffusion mechanism where the nonradiative recombinations of electron-hole pairs created by the irradiation can assist defect diffusion. This fast diffusion of nonequilibrium defects can be drastically reduced by lowering the sample temperature during the bombardment. This very important unexpected finding clearly evidences the advantage of low temperature implantations, when damage injection must be highly localized as for advanced nanostructure fabrication for example.

ACKNOWLEDGMENT

One of the authors (J.G.) would like to acknowledge Pr. F. Miserey from the Conservatoire des Arts et Métiers for his encouragement and continuous help throughout this study.

[1]J. P. Biersack and L. G. Haggmark, Nucl. Instrum. Methods Phys. Res. B **182/183**, 199 (1981).
[2]F. Laruelle, A. Bagchi, M. Tsuchiya, J. Merz, and P. M. Petroff, Appl. Phys. Lett. **56**, 1561 (1990).
[3]H. Kasahara, H. Sawaragi, R. Aihara, K. Gamo, S. Namba, and M. HasselShearer, J. Vac. Sci. Technol. B **6**, 974 (1988).
[4]R. Germann, A. Forchel, M. Bresch, and H. P. Meier, J. Vac. Sci. Technol. B **7**, 1475 (1989).
[5]N. G. Stoffel, J. Vac. Sci. Technol. B **10**, 651 (1992).
[6]D. L. Green, E. L. Hu, and N. G. Stoffel, J. Vac. Sci. Technol. B **12**, 3311 (1994).
[7]C.-H. Chen, D. L. Green, and E. L. Hu, J. Vac. Sci. Technol. B **13**, 2355 (1995).
[8]C.-H. Chen, D. L. Green, E. L. Hu, J. P. Ibbetson, and P. M. Petroff, Appl. Phys. Lett. **69**, 58 (1996).
[9]C.-H. Chen, E. L. Hu, U. K. Mishra, J. P. Ibbetson, X. Wu, and J. S. Speck, Appl. Phys. Lett. **69**, 1728 (1996).
[10]H. Nakanishi, K. Wada, and W. Walukiewicz, J. Appl. Phys. **78**, 5103 (1995).
[11]A. Forchel, H. Leier, B. E. Maile, and R. Germann, in *Festkörperprobleme, Advances in Solid State Physics*, edited by U. Rössler (Vieweg, Braunschweig, 1988), Vol. 28, p. 99.
[12]H. Linke, I. Maximov, D. Hessman, P. Emanuelsson, W. Qin, L. Samuelson, and K. B. Meyer, Appl. Phys. Lett. **66**, 1403 (1995).
[13]Y. Hirayama, S. Tarucha, Y. Suzuki, and H. Okamoto, Phys. Rev. B **37**, 2774 (1988).
[14]J. Gierak, G. Ben Assayag, M. Schneider, C. Vieu, and J. Y. Marzin, Microelectron. Eng. **30**, 253 (1996).
[15]H. F. Wong, D. L. Green, T. Y. Liu, D. G. Lishan, M. Bellis, E. L. Hu, P. M. Petroff, P. O. Holtz, and J. L. Merz, J. Vac. Sci. Technol. B **6**, 1906 (1988).
[16]R. Dingle, C. Weisbuch, H. L. Störmer, H. Morkoç, and A. Y. Cho, Appl. Phys. Lett. **40**, 507 (1982).
[17]L. W. Swanson, Nucl. Instrum. Methods Phys. Res. **218**, 347 (1983).
[18]J. Gierak and G. Ben Assayag, Microelectron. Eng. **30**, 261 (1996).
[19]C. Vieu, A. Claverie, J. Fauré, and J. Beauvillain, Mater. Sci. Eng., B **2**, 27 (1989).
[20]S. Y. Chiang and G. L. Pearson, J. Appl. Phys. **46**, 2986 (1975).
[21]C. Vieu, Defect Diffus. Forum **119–120**, 127 (1995).
[22]C. Vieu, G. Ben Assayag, and J. Gierak, Nucl. Instrum. Methods Phys. Res. B **93**, 439 (1994).
[23]M. Schneider, J. Gierak, and C. Vieu (unpublished).

On the improvement in thermal quenching of luminescence in SiGe/Si structures grown by molecular beam epitaxy

W. M. Chen,[a] I. A. Buyanova, G. Pozina, B. Monemar, W. X. Ni, and G. V. Hansson
Department of Physics and Measurement Technology, Linköping University, S-581 83 Linköping, Sweden

(Received 7 September 1997; accepted 18 May 1998)

Thermal quenching of photoluminescence (PL) from SiGe/Si quantum well (QW) structures grown by molecular beam epitaxy is shown to be more severe when grown at a lower temperature. The mechanism responsible for the thermal quenching of PL is discussed as being due to thermally activated nonradiative recombination channels, related to defects in both Si barriers and SiGe QW. Nonradiative defects in Si can be rather efficiently deactivated by post-growth treatments such as hydrogenation and thermal annealing, leading to a significant improvement in the thermal quenching behavior of PL from single QW structures. Nonradiative defects in SiGe are found to be thermally stable, on the other hand, evident from the experimentally observed minor role played by post-growth thermal annealing in the thermal quenching of PL from multiple QW structures.
© *1998 American Vacuum Society.* [S0734-211X(98)06704-3]

I. INTRODUCTION

There have been long-standing efforts in searching for efficient optoelectronic devices based on Si-related materials systems, stimulated by their potential compatibility with mature Si technology and new possibilities for various layered and quantum devices. Many difficulties associated with Si-based materials systems, such as low radiative efficiency and a severe thermal quenching of luminescence, have hindered progress in realistic device applications. The low radiative efficiency is known to be inherently due to the indirect energy band gap. Many approaches have been attempted to overcome the low radiative efficiency, including band-gap engineering by atomic layer superlattices and defect engineering by incorporating light emitting impurities in the material. The dominant mechanisms for the severe thermal quenching have, on the other hand, been discussed in terms of strong surface recombination[1] and the presence of nonradiative channels[2,3] that are thermally activated. While the surface recombination can be suppressed by proper surface treatments, a complete removal of nonradiative defects is still far beyond reach. This is largely due to a lack of definite experimental evidence and knowledge as to what the defects are and where they are located. This has thus deterred any deliberate elimination of these nonradiative channels and has undermined efforts in improving the radiative efficiency of the material by band-gap and defect engineering.

In this work we reveal that the dominant mechanism responsible for the rapid thermal quenching of SiGe/Si heterostructures grown by molecular beam epitaxy (MBE) is due to the presence of nonradiative defects introduced during the growth. We shall demonstrate effects of growth temperature and post-growth treatments on the nonradiative defects. We shall also provide experimental evidence on the origin and location of the defects.

[a]Electronic mail: wmc@ifm.liu.se

II. EXPERIMENT

Three types of samples were selected in this study, as listed in Table I, all grown on (100) Si substrate by MBE with a Balzers UMS 630 Si-MBE system.

The first type of the samples is used to study effects of growth temperature and post-growth treatments. A comparison of results between the first two types of the samples will shed light on the contribution from the SiGe layers. The third type of the sample is studied to single out the contribution from the Si barriers. All of the samples were started with a 1000 Å undoped Si buffer layer, and were finally capped by a 1000 Å undoped Si layer. The growth rate was typically 1–2 Å/s.

Post-growth hydrogen treatment was done at around 200 °C for 60 min, inside a quartz reactor with a remote dc H plasma at a pressure of 2.0 mTorr. The post-growth thermal annealing was carried out at 500 °C for 15 min in an argon gas environment.

The photoluminescence (PL) experiments were carried out with an Oxford variable temperature cryostat, where sample temperature could be varied between 2 and 300 K. The 514.5 and 1090 nm lines of an argon-ion laser were used for above and below band-gap optical excitation of the Si barriers. PL emissions were first dispersed by a Jobin–Yvon 0.5 m double grating monochromator and then collected by a cooled North-coast Ge detector. The optically detected magnetic resonance (ODMR) experiments were performed at the *X* band (9.23 GHz) using a modified Bruker ER-200D ESR spectrometer, equipped with a TE_{011} microwave cavity with optical access in all directions. PL emissions from the samples, under illumination of the UV multilines (333.6–363.8 nm) of an Ar^+ laser, were monitored by a cooled Ge detector. The ODMR signal was obtained by detecting a synchronous change in the PL with field modulation of the magnetic field. A derivative lineshape of the ODMR is observed in this case when the magnetic field is modulated on and off the spin resonance conditions.

TABLE I. List of samples studied in this work.

Types of samples	Growth temperature (°C)	Thickness (Å)
Undoped $Si_{0.8}Ge_{0.2}$/Si single quantum well (SQW)	420 and 620 °C	QW width: 32 Å
Undoped $Si_{078}Ge_{0.22}$/Si multiple QWs (MQWs)	420 °C	QW width: 50 Å Spacer thickness: 200 Å Number of period: 20
Undoped thin Si epilayers	420 °C	2000 Å

III. RESULTS AND DISCUSSION

In Fig. 1 we show PL spectra and their dependence on the measuring temperature from the single-quantum-well (SQW) structures with optical excitation above the band gap of the Si barriers. The PL emission arising from the excitonic recombination confined within the QWs is denoted as X^{NP}, X^{TA}, and X^{TO} for no-phonon, transverse acoustic, and optical phonon assisted transitions, respectively. The higher energy part of the PL spectra originates from the Si buffer layer and the substrate.

As it can clearly be seen, the SiGe QW emission from the sample grown at 420 °C quenches rapidly with increasing measuring temperature, and falls below the detection limit already around 30 K. A detailed analysis of the Arrhenius plot of the integrated PL intensity as a function of the reciprocal temperature (Fig. 2) reveals that the thermal quenching is governed by two thermally activated processes with rather low activation energies, i.e., $E_a^1 = 1.7$ meV and $E_a^2 = 5.1$ meV, respectively. These activation energies were determined by a curve fitting to the expression describing the PL thermal quenching process: $I_{PL} \propto [1 + C_1 \exp(-E_a^1/kT) + C_2 \exp(-E_a^2/kT)]^{-1}$, where C_1 and C_2 are prefactors. These activation energies are much too low to be attributed to the intrinsic thermal activation of holes from the QW, which is about 100 meV determined from the energy difference between the exciton from the SiGe QW and the free exciton from Si and assuming that the band offset between Si and SiGe occurs predominantly in the valence band. The dominant process governing the thermal quenching in this structure is therefore extrinsic, due to nonradiative defects to be discussed below.

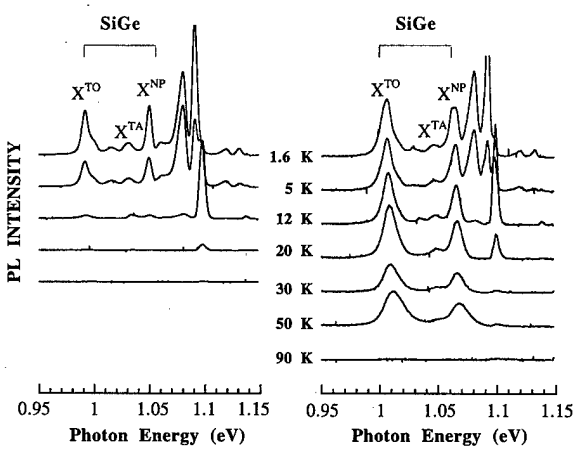


FIG. 1. PL spectra from the SiGe SQW structures grown at 420 and 620 °C, as a function of temperature. The optical excitation wavelength is 514.5 nm.

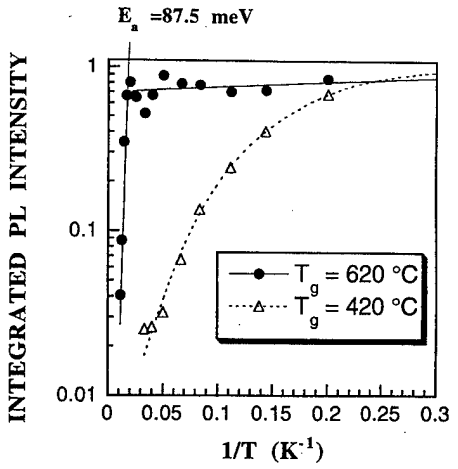


FIG. 2. Arrhenius plots of the integrated PL intensity measured from the sample structures shown in Fig. 1. The optical excitation wavelength is 514.5 nm.

A. Effect of growth temperature

The thermal quenching behavior of PL from the SQW structures can be significantly improved by raising growth temperature. As it is evident from Fig. 1 and Fig. 2, the thermal quenching of PL from the SQW grown at 620 °C is found to be much less severe as compared to that grown at 420 °C. The thermal activation is now dominated by a process with an activation energy of about 87.5 meV, close to the QW depth estimated from the energy difference of the excitons in this structure. This shows that the thermal quenching is dominated by the intrinsic thermal activation of the holes from the QW, and the extrinsic effects play a less important role in structures grown at a higher temperature.

B. Effects of post-growth treatments

Effects of post-growth treatments by hydrogen and thermal annealing were studied to add new information on properties of the nonradiative process causing the thermal quenching of PL from the SQW grown at 420 °C. It is found that these post-growth treatments lead to a significant improvement in the thermal quenching behavior of the structure, see Fig. 3, making the intrinsic thermal activation step visible at the high temperature side. There is, however, still a sizeable contribution from a nonradiative process after the hydrogen treatment, leading to a premature PL quenching of a low activation energy at 4 meV [seen by an additional shoulder in the Arrhenius plot (Fig. 3)]. In contrast, the post-growth thermal annealing appears to be very efficient, resulting in a weak thermal quenching until the intrinsic activation from the QWs becomes dominant.

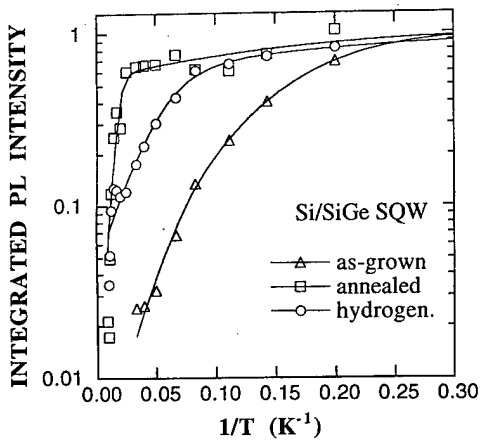


FIG. 3. Arrhenius plots of the integrated PL intensity measured from the SiGe SQW grown at 420 °C, showing the effects of the post-growth treatments. The optical excitation wavelength is 514.5 nm.

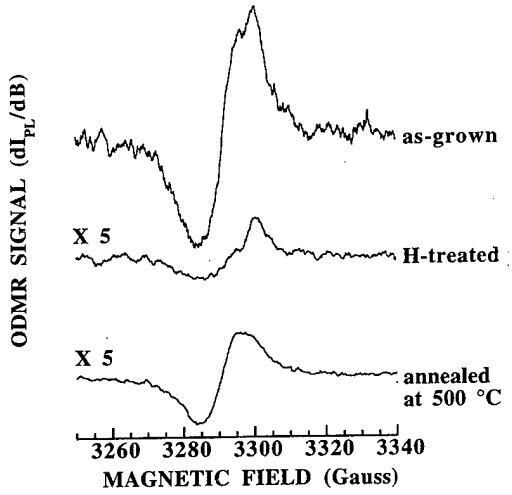


FIG. 4. ODMR spectra from the as-grown and post-growth treated SQW grown at 420 °C, measured at 5 K.

C. Role of nonradiative defects

Two main mechanisms have so far been suggested in the literature to account for thermal quenching of PL from SiGe QW structures: surface recombination[1] and nonradiative defects.[2,3] In our study the 514.5 nm laser line was used as the excitation source with a penetration depth of about 1 μm and can be regarded as bulk excitation. This has minimized any contribution from surface recombination, leaving nonradiative defects as the dominant cause for the thermal quenching of PL.

In order to obtain information on the chemical identity and microscopic structure of the nonradiative defects, the ODMR experimental technique has been employed. The key to the experimental approach is to utilize the competing carrier recombination processes between radiative and nonradiative defects.[4] In this case the nonradiative defects are monitored by magnetic resonance, while the radiative channels are detected by photoluminescence. A magnetic-resonance enhanced recombination via nonradiative channels, provided that they are among the dominant recombination channels, leads to a corresponding reduction in free carrier concentration available for recombination via radiative channels. This results in a decrease in PL intensity, or equivalently a negative ODMR signal.

In Fig. 4 ODMR spectra from the SQW structure are shown, which originate from nonradiative defects present in the structure. A decrease in the ODMR signal strength after the post-growth treatments indicates a corresponding reduction of the defect concentration, which is consistent with the improvement in the thermal quenching behavior of PL discussed above. Due to a broad linewidth, no structure could be resolved from the ODMR spectra, unfortunately. Such a broadening is likely due to an strong overlap of ODMR signals from both the SiGe QW and the Si barriers, and also due to a possible potential distribution (induced by alloy fluctuation and a variation in the QW width and the strain field) which the defects experience.

D. Nonradiative defects in the Si barriers

To separate the contribution by the Si barriers, a parallel study of the Si epilayers grown under the identical conditions but without the SiGe layer(s) is done. The resulting ODMR spectra from both as-grown and post-growth treated Si epilayers are illustrated in Fig. 5. Here a narrower ODMR linewidth due to the absence of the strained SiGe enables a positive identification of each ODMR line to its corresponding defect.[5]

One of the dominant nonradiative defects, for example, is the vacancy-oxygen (V-O) complex which gives rise to the ODMR signal denoted as "2:"[5] This complex has previously been shown to be the predominant defect introduced only after bombardment of CZ Si by high energy particles (e.g., electrons). We showed that such defect can also be

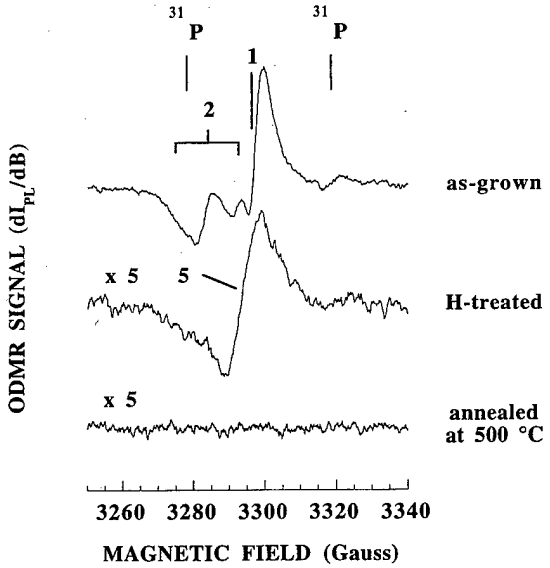


FIG. 5. ODMR spectra from the as-grown and post-growth treated Si layer grown at 420 °C, measured at 5 K.

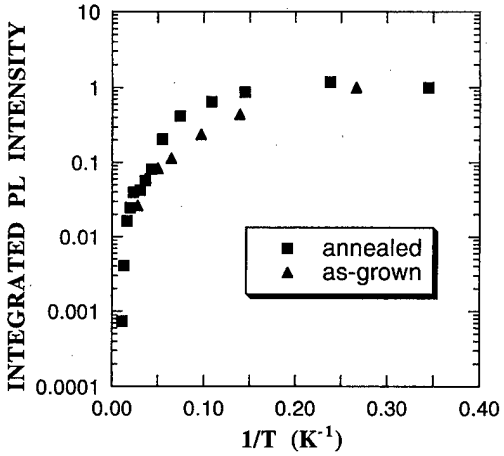


FIG. 6. Arrhenius plots of the integrated PL intensity measured from the SiGe MQW grown at 420 °C, showing the effects of the post-growth treatments. The optical excitation wavelength is 514.5 nm.

introduced in the as-grown MBE Si and related structures grown at low temperature, as a result of a low surface adatom mobility and a low oxygen desorption rate during the low temperature growth.[5] The formation of the V-O complex can particularly be promoted by the occurrence of ion bombardment, e.g., during potential-enhanced doping.

It is clearly shown that the V-O defect can be nearly completely removed by either hydrogenation or by thermal annealing at 500 °C, in agreement with previously reported annealing behavior for this defect in bulk CZ Si. It can also be noticed that there is still a defect (denoted as "5") remaining in the hydrogen treated sample. The origin of this remaining defect is still unknown, argued to be (a) either already present in the as-grown material but obscured by the much stronger ODMR signals "2" and "1" from the V-O complex and another possibly vacancy-related defect (b) or introduced by the hydrogen treatment. The effect of the thermal annealing on removing the nonradiative defects is proven to be strongest. The effect of the post-growth treatments on the nonradiative defects (Fig. 5) and on the thermal quenching of the QW PL (Fig. 3) correlates remarkably well, establishing a direct link between these defects and the PL thermal quenching. It is therefore believed that the nonradiative defects in the Si barriers are to a great extent responsible for the rapid thermal quenching of PL from the SiGe SQW grown at 420 °C via strong competing carrier capture and recombination processes in the barriers so that less nonequilibrium carriers can be trapped and recombine in the SiGe QW.

E. Nonradiative defects in SiGe

Multiple QW structures were chosen to study contributions of nonradiative defects in the SiGe layer to the PL thermal quenching, as the relative volume of SiGe increases. Surprisingly, post-growth annealing has little effect on the thermal quenching behavior of PL from the MQW structure (Fig. 6) in contrast to the observation in the SQW structure. This means that the dominant nonradiative defects in this structure is thermally stable up to 500 °C. To show that the dominant defects are situated in the SiGe QWs, the below Si band-gap excitation by the 1090 nm laser line was employed to avoid any contribution from the Si barriers since they no longer participate in the carrier generation-recombination processes. No noticeable difference can be observed between the above and below barrier excitation, indicating the thermal quenching of PL is caused by nonradiative defects situated in the SiGe QW layers.

IV. SUMMARY

In this work, we have shown that grown-in nonradiative defects are largely responsible for a rapid thermal quenching of luminescence from SiGe quantum wells, impairing practical applications of these structures at room temperature. We have provided experimental evidence that the dominant mechanism responsible for a strong thermal quenching of PL from the SiGe SQW grown at low temperatures (<500 °C) is due to thermally activated, efficient nonradiative recombination channels situated in the Si barriers, with rather low activation energies of about 1–5 meV. This leads to a complete quenching of luminescence at a temperature as low as 30 K. A post-growth hydrogen treatment is demonstrated to be only moderately effective in passivating the nonradiative defects and in reducing thermal quenching. Post-growth thermal annealing at high temperature (e.g., >500 °C) is, on the other hand, shown to be highly effective and leads to a nearly complete removal of the nonradiative defects in Si monitored in the ODMR experiments. By removing these nonradiative defects and thus the shunt path for carrier recombination, a significant improvement in thermal quenching behavior of luminescence from the SiGe SQW structures has been achieved. The thermal quenching, in this case, is dominated by the intrinsic thermal activation of the holes from the QWs with an activation energy of about 90 meV in the structures studied. These results correlate very well with effects of hydrogenation and thermal annealing of the nonradiative defects monitored in the ODMR experiments. Selective optical excitation above and below the band gap of the Si barriers has been used to determine relative contributions of nonradiative recombination channels present in the SiGe MQW and the Si barriers. The dominant nonradiative defects in the SiGe MQW structure have been shown to be thermally stable, in sharp contrast to the situation in the SQW structure, and therefore cannot be removed by post-growth annealing. These defects are situated in the SiGe layers, which play an increasingly important role in the thermal quenching of PL from the SiGe QW structures when the relative volume of SiGe increases. This explains why these thermal stable defects in SiGe dominate the thermal quenching of PL in the SiGe MQW structure but not in the SiGe SQW structure. Further studies are required to seek for the way to efficiently eliminate these nonradiative defects and to improve the radiative efficiency of the material. At present, high temperature growth seems to provide better optical properties of MBE SiGe QW in particular when PL thermal quenching is concerned.

Presented at the Silicon Heterostructure Conference, Barga, Italy, 15–19 September 1997.

[1]A. St. Amour, J. C. Sturm, Y. Lacroix, and M. L. W. Thewalt, Appl. Phys. Lett. **65**, 3344 (1994).
[2]S. Fukatsu and Y. Shriraki, J. Cryst. Growth **150**, 1025 (1995).
[3]M. Wachter, K. Thonke, R. Sauer, F. Schäffler, H. J. Herzog, and E. Kasper, Thin Solid Films **222**, 10 (1992).
[4]W. M. Chen and B. Monemar, Appl. Phys. A: Solids Surf. **A53**, 130 (1991).
[5]W. M. Chen, I. A. Buyanova, W.-X. Ni, G. V. Hansson, and B. Monemar, Phys. Rev. Lett. **77**, 4214 (1996).

Changes in morphology using atomic hydrogen during $Si/Si_{1-x}Ge_x$ molecular beam epitaxy growth on Si (100)

C. Silvestre, G. G. Jernigan, M. E. Twigg, and P. E. Thompson[a)]
Code 6812, Naval Research Laboratory, Washington, DC 20375

(Received 25 March 1998; accepted 29 May 1998)

Si and $Si_{1-x}Ge_x$ molecular beam epitaxial growths have been conducted with and without 10^{-3} Pa atomic hydrogen (AH) overpressure to observe the effect of AH on surface morphology. $Si_{1-x}Ge_x$ ($x=0.1$ and 0.2) layers 3 nm thick grown at 710 °C without AH showed flat epitaxy by transmission electron microscopy but identical layers grown in the presence of AH showed interface stress undulations. Stress undulations were also observed for a 20% Ge alloy grown with AH at 600 °C. For 30% Ge alloy layers grown at 710 °C, undulations were seen both with and without AH with a $\approx$250 nm period, however the amplitudes of the undulations were greater with AH. Low energy electron diffraction investigation of Si homoepitaxy revealed, at growth temperatures between 600 and 800 °C, improved epitaxy with 100 s of 10^{-3} Pa AH. In combination, these results are evidence that AH has increased the adatom mobility.[S0734-211X(98)13004-4]

I. INTRODUCTION

There is current interest in the use of atomic hydrogen (AH) with Si molecular beam epitaxy (MBE) to improve the epitaxial growth.[1–9] The effects of AH on a static Si (100) surface have been studied[10] and are known to result in a roughening of the surface and a conversion of the (2×1) surface reconstruction to a (1×1) surface reconstruction. However, the effect of AH on a growing Si (100) surface may not necessarily be the same as on a static Si (100) surface. It is clear AH affects growths because of previous reports that AH can affect surface order,[10] act as a surfactant,[1] and change dopant activation.[9] However, a complete understanding of the effect AH has on a growing Si (100) surface is needed prior to insertion in a device fabrication process.

In this article, we examine changes in surface morphology caused by AH during $Si_{1-x}Ge_x$ MBE growth. During $Si_{1-x}Ge_x$ MBE growth without AH, surface morphology changes occur due to stress relief mechanisms associated with the lattice mismatch between the SiGe alloy and the Si substrate.[11,12] The surface morphology of $Si_{1-x}Ge_x$ layers is determined by minimizing the sum of the surface free energy (due to the surface reconstruction) and the bulk free energy (due to the elastic strain). In sufficiently thick Ge alloy films, strain is relieved through the formation of dislocations. Studies have shown that alloy films with Ge concentration >20%, the formation of surface undulations (or ripples) reduces bulk elastic strain energy.[12–16] An undulated $Si_{1-x}Ge_x$ layer is illustrated in Fig. 1. An undulation can be characterized by the distance between crests, λ, and the amplitude of the crests, A. This misfit strain energy originates from the 4% larger lattice spacing of Ge compared to Si. Undulation formation can relieve bulk stress. During growth, the crest region will expand laterally, relieving elastic stress. The reduction in elastic free energy must more than compensate for the energy cost of the expanded surface area due to rippling. While the undulated surface may be the thermodynamically favored low-energy configuration, undulation may not occur during growth due to kinetic limitations. The formation of undulations requires significant lateral adatom movement. At rapid growth rates, sufficient time may not be available for this movement to occur. At low Ge concentrations, the strain energy, which drives the lateral adatom movement, may be insufficient for the formation of undulation even for modest growth rates.

In our experiment, we use cross-sectional transmission electron microscopy (XTEM) to examine $Si_{1-x}Ge_x$ multiple quantum well structures exposed to AH during growth as a function of Ge concentration, substrate temperature, and growth rate. From the XTEM images, we observe surface undulations between the SiGe quantum wells in the Si spacer layers. We will demonstrate that the use of AH induces $Si_{1-x}Ge_x$ to form surface undulations under conditions where undulations would not normally occur.

II. EXPERIMENT

Growths were performed with and without AH in a MBE system equipped with elemental Si and Ge *e*-beam evaporators and a hot filament gas doser. The Si and Ge growth rates available ranged from 0.001 to 0.2 nm/s and 0.001 to 0.1 nm/s, respectively. The system base pressure was 5×10^{-9} Pa, and a typical pressure during growth was 5×10^{-7} Pa. Substrate temperature was monitored by a pyrometer that was calibrated based on the eutectic temperatures of Au and Al on an equivalent Si wafer. The substrate temperature uncertainty is estimated to be ±10 °C. Si and Ge growth rates were calibrated using surface profilometry. Molecular hydrogen was introduced into the MBE system through the hot filament gas doser. The doser cracked the hydrogen on a 2150 °C tungsten filament and collimated the AH flux. Cracking efficiency is estimated to be 30%.[17] Prior

a)Electronic mail: thompson@estd.nrl.navy.mil

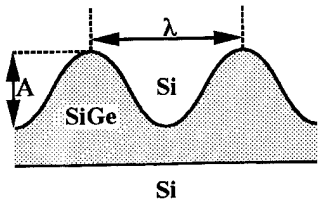


FIG. 1. Strained epitaxial layers may elastically relieve stress by forming an undulated surface. The undulations are characterized by two parameters: λ, the distance between adjacent crests; and *A*, the amplitude of the undulations. The increase in the surface free energy is more than compensated for by the decrease in the strain energy. Schematic per Cullis *et al.* (Refs. 12–14).

to growth, 75 mm, B-doped, 1–20 Ω cm, Si (100) wafers were chemically cleaned *ex situ*,[18] which resulted in a hydrogen terminated surface.

The Si deposition rate was kept constant at 0.050 nm/s, and the Ge growth rates were 0.006, 0.012, and 0.021 nm/s for the 10%, 20%, and 30% Ge alloy layers, respectively. The SiGe layer growth time was kept constant at 48 s, resulting in increasing alloy thicknesses of 2.7, 3.0, and 3.6 nm. The alloy layer thicknesses were selected to be less than the Mathews–Blakeslee critical thickness,[19] so dislocations would never occur. The constant alloy-layer growth time fixed the duration of adatom diffusion during alloy layer growth. The application of 10^{-3} Pa AH occurred 100 s prior to, during, and 100 s after the Ge alloy layers. Alloy layers were separated by 30 nm Si spacer layers. For Ge alloy growth at 600 °C, a 166 nm buffer layer was grown at 650 °C prior to lowering the wafer temperature for alloy growth. For Ge alloy growth at 710 °C, the 166 nm buffer layer was grown at 710 °C. Following completion of the Ge alloy layers, a 225 nm Si capping layer was grown at the Ge alloy growth temperature. All samples imaged in XTEM were thinned by mechanical lapping and ion milling at ambient temperature. These XTEM samples were imaged in dark field using a (400) reflection.

III. RESULTS

XTEM images of the 10%, 20%, and 30% Ge alloy layers grown at 710 °C without AH are shown in Fig. 2. XTEM images of equivalent alloy layers grown with AH are shown in Fig. 3. For the 10% and 20% Ge alloy layers grown without AH, no surface undulations were seen. For the 30% Ge alloy layers grown without AH, undulations are observed as periodic lateral oscillations. For all alloys grown with AH, (Fig. 3) undulations are present, even for Ge concentrations of 10% and 20%. Further, the oscillations in subsequent layers are observed to be in registry with oscillations in the initial layer. These are the lowest Ge concentrations for which undulations have been observed. Comparing the XTEM images of the 30% Ge alloy layers grown with and without AH, it is observed that the wavelengths of the oscillations are approximately the same, ≈250 nm. The XTEM image of the 30% Ge growth with AH shows evidence of Asby–Brown contrast[20] indicating a larger undulation amplitude than for the 30% Ge film grown without AH.

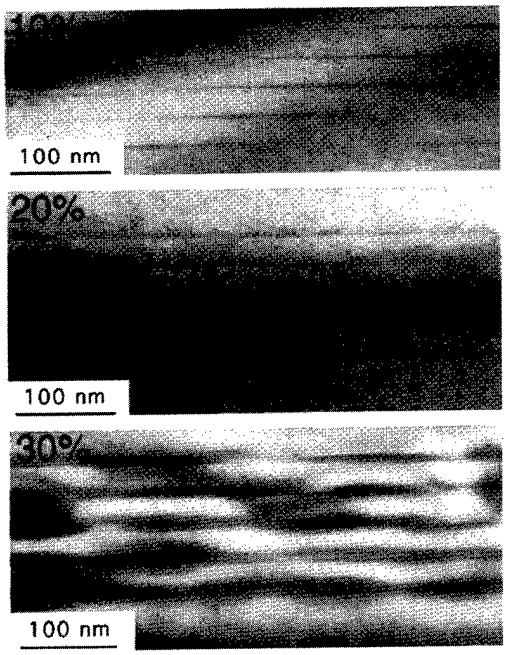


FIG. 2. XTEM images of five alloy layers with 10% (top), 20% (middle), and 30% (bottom) Ge grown on Si (100) at 710 °C without AH exposure.

The temperature dependence of the formation of the undulations was investigated by growing the 20% Ge structure with AH at 710 and 600 °C. The resulting XTEM images are shown in Fig. 4. At 710 °C and 600 °C, strong undulations

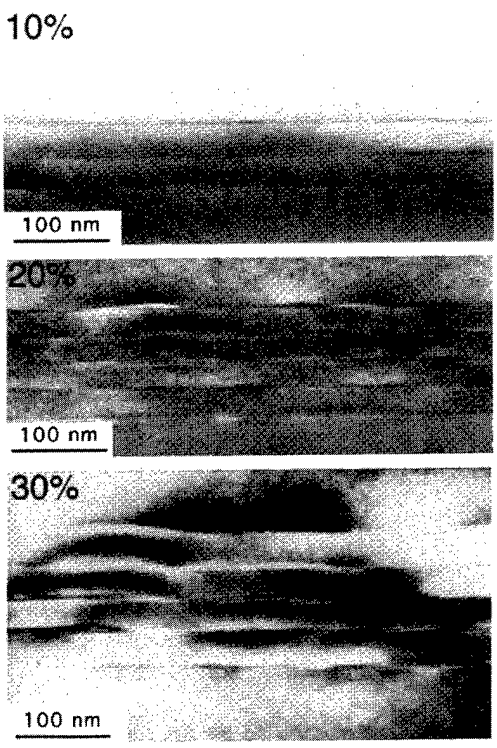


FIG. 3. XTEM images of five alloy layers with 10% (top), 20% (middle), and 30% (bottom) Ge grown on Si (100) at 710 °C with 10^{-3} Pa AH exposure for 100 s prior to, during, and for 100 s after the alloy layers.

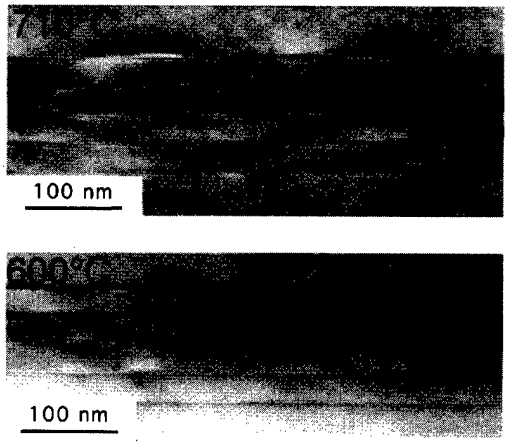


FIG. 4. XTEM image of five 3 nm thick $Si_{0.8}Ge_{0.2}$ layers grown on Si (100) at 710 (top) and 600 °C (bottom) with 10^{-3} Pa AH exposure for 100 s prior to, during, and for 100 s after the 20% Ge alloy layers. The stress oscillations are visible in both.

occurred as a result of AH exposure. It has been observed that the wavelength of the undulation has decreased significantly with growth temperature. The wavelength was 200 nm for the growth at 710 °C and 70 nm for the growth at 600 °C.

The growth rate for the 20% Ge structure, grown at 710 °C without AH, was lowered to see if undulations could be induced to form. The Si and Ge deposition rates were reduced by a factor of 12 to the minimum available, 0.004 nm/s and 0.001 nm/s, respectively. The growth conducted at these rates without AH still showed no evidence of undulations, thus indicating that AH exposure had a greater effect on surface morphology than a 12-fold increase in the time for surface adatom diffusion.

To verify that the effects of AH on morphology were not due to radiative heating or background gas effects, a 20% Ge alloy was grown at 710 °C under an exposure to Xe. The Xe passed through the doser under identical conditions to those used for AH. No undulations were visible in the XTEM images of this sample, but undulations were observed in an equivalent sample grown with AH (Fig. 3). Thus, the undulations are clearly the result of the interaction of AH and the growing surface.

To better understand the contribution of AH to the growth process, the effect of AH on Si homoepitaxy was evaluated by low energy electron diffraction (LEED). A qualitative LEED examination of Si epitaxial wafers grown with and without 100 s, 10^{-3} Pa AH at 600, 700, and 800 °C showed improved LEED pattern sharpness and decreased LEED background brightness with AH. The difference between LEED patterns with and without AH exposure was greatest at 600 °C. The Si deposition process leads to surface adatoms and we interpret the changes in LEED as due to the enhanced adatom migration and incorporation.[9] The changes in LEED images are due to a decrease in step density, which would lead to sharper spots, and point defects, which would lead to decreased background brightness.[21]

IV. DISCUSSION

Our experimental work has revealed two primary results: (1) AH causes undulations to occur in 10% and 20% Ge alloy thin films grown at 710 °C, where, without AH, the undulations would not occur; (2) AH causes the undulations in 30% alloy thin films to increase in amplitude. Both of these results may be explained by increased adatom mobility due to AH exposure.

For undulation formation to occur as a strain relief mechanism, adatoms must have sufficient energy to laterally diffuse on a surface. For a 3.6 nm, 30% Ge alloy film, enough strain energy is provided for undulation formation. However, for a 2.7 nm, 10% Ge alloy and a 3.0 nm, 20% Ge alloy, the strain energy alone is not enough to drive the movement of adatoms. The application of AH lowers the surface kinetic barriers to adatom diffusion and allows the formation of undulations for the case where sufficient strain energy did not previously exist. For the 30% Ge film, the application of AH lowers the barrier to adatom diffusion, resulting in an even larger effect in undulation amplitude, as was observed.

This interpretation of our experimental data is supported by previous work on Si and compound semiconductors. Boland has reported enhancement of Si adatom diffusion resulting from the presence of hydrogen on the surface in a review article.[10] Using scanning tunneling microscopy, Boland showed that an exposure to AH could induce adatom island formation on the Si (111)(7×7) surface. In an analysis of chemical vapor deposition growth from silane, hydrogen on the surface was postulated to reduce the magnitude of the barriers to site to site hopping.[22] The use of AH during GaAs (331) growth by MBE increases adatom mobility.[23] Further, the observation of improved LEED patterns that we report above may also be explained by enhanced adatom diffusion.

The undulation amplitude is a function of Ge concentration when the film is at equilibrium, but under kinetic limitations the amplitude is also dependent on adatom diffusivity. If surface adatom diffusivity increases, then for a constant growth time the amplitude of the oscillations can increase to the thermodynamic maximum. For the 30% Ge films we see that without AH the amplitude is limited by surface diffusion. The use of AH increases adatom diffusivity allowing the undulation amplitude to reach its maximum. Comparing the undulation amplitude as a function of Ge concentration dependence when AH was used, we see the strength of oscillation increases with Ge concentration. This is to be expected as the driving force for the formation of oscillations is misfit strain. The temperature dependence of the undulation wavelength, however, is not fully understood and is currently under investigation.

V. CONCLUSION

Our studies have shown that AH exposure during growth results in interface undulations in 10% and 20% Ge alloy layers, as seen by XTEM. Without AH exposure, there were no undulations. Previously, these undulations have only been seen for Ge alloy layers with concentrations greater than

20%, and are used as evidence to show that AH exposure increases the surface adatom mobility. AH-enhanced adatom mobility accelerates the development of $Si_{1-x}Ge_x$ undulations. At high Ge concentrations, where undulations occur without AH exposure, AH-enhanced adatom mobility increases the amplitude of the undulations.

ACKNOWLEDGMENTS

This work is supported by ONR and AFOSR. We acknowledge the assistance of L. Ardis for technical support.

[1]G. Ohta, F. Fukatsu, Y. Ebuchi, T. Hattori, N. Usami, and Y. Shiraki, Appl. Phys. Lett. **65**, 2975 (1994).
[2]A. Sakai and T. Tatsumi, Appl. Phys. Lett. **64**, 52 (1994).
[3]K. Nakagawa, A. Nishida, Y. Kimura, and T. Shimada, Jpn. J. Appl. Phys., Part 2 **33**, L1331 (1994).
[4]K. Nakagawa, A. Nishida, Y. Kimura, and T. Shimada, J. Cryst. Growth **150**, 939 (1995).
[5]M. Okada, T. Shimizu, H. Ikeda, S. Zaima, and Y. Yasuda, Appl. Surf. Sci. **113/114**, 349 (1997).
[6]S. J. Kahng, J. Y. Park, K. H. Both, J. Lee, Y. Khang, and Y. Kuk, J. Vac. Sci. Technol. A **15**, 927 (1997).
[7]S. J. Kahng and Y. Kuk, Sci. Rep. Res. Inst. Tohoku Univ. A **44**, 59 (1997).
[8]P. E. Thompson, C. Silvestre, M. Twigg, G. Jernigan, and D. S. Simons, Thin Solid Films (to be published).
[9]C. Silvestre, P. E. Thompson, G. Jernigan, and D. Simons, J. Vac. Sci. Technol. A (to be published).
[10]J. J. Boland, Adv. Phys. **42**, 129 (1993).
[11]M. Albrecht, S. Christiansen, J. Michler, W. Dorsch, and H. P. Srank, Appl. Phys. Lett. **67**, 1232 (1995).
[12]A. G. Cullis, MRS Bull. **21**, 21 (1996).
[13]A. G. Cullis, D. J. Robbins, A. J. Pidduck, and P. W. Smith, J. Cryst. Growth **123**, 333 (1992).
[14]A. G. Cullis, in *Evolution of Epitaxial Structure and Morphology,* edited by A. Zangwill, D. E. Jesson, D. Chambliss, and R. Clarke (Material Research Society, Pittsburgh, PA, 1996), p. 303.
[15]T. Walther, C. J. Humphreys, and A. G. Cullis, Appl. Phys. Lett. **71**, 809 (1997).
[16]A. J. Pidduck, D. J. Robbins, and A. G. Cullis, in *Microscopy of Semiconducting Materials 1993,* edited by A. G. Cullis, A. E. Staton-Bevan, and J. L. Hutchinson (Institute of Physics, Bristol, UK, 1993), p. 609.
[17]G. W. Wicks, E. R. Rueckwald, and M. W. Koch, J. Vac. Sci. Technol. B **14**, 2184 (1996).
[18]P. E. Thompson, M. E. Twigg, D. J. Godbey, K. D. Hobart, and D. S. Simons, J. Vac. Sci. Technol. B **11**, 1077 (1993).
[19]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[20]P. Hirsh, A. Howie, R. B. Nicholson, D. W. Pashley, and M. J. Whelan, *Electron Microscopy of Thin Crystals* (Krieger, Huntington, NY, 1977).
[21]M. Henzler, in *Dynamical Phenomena at Surfaces, Interfaces, and Superlattices,* edited by F. Nizzoli, K. H. Rieder, and R. F. Willis (Springer, Berlin, 1984), p. 14.
[22]M. Tao and L. P. Hunt, J. Electrochem. Soc. **144**, 2221 (1997).
[23]H. P. Schonherr, J. Fricke, Z. Niu, K. J. Friedland, R. Notzel, and K. H. Ploog, Appl. Phys. Lett. **72**, 566 (1998).

Stabilizing the surface morphology of $Si_{1-x-y}Ge_xC_y$/Si heterostructures grown by molecular beam epitaxy through the use of a silicon-carbide source

E. T. Croke,[a] J. J. Vajo, and A. T. Hunter
HRL Laboratories, LLC, 3011 Malibu Canyon Road, RL63, Malibu, California 90265

C. C. Ahn
California Institute of Technology, Pasadena, California 91125

D. Chandrasekhar, T. Laursen, David J. Smith, and J. W. Mayer
Center for Solid State Science, Arizona State University, Tempe, Arizona 85287-1704

(Received 9 September 1997; accepted 18 May 1998)

$Si_{1-x-y}Ge_xC_y$/Si superlattices were grown by solid-source molecular beam epitaxy using silicon carbide as a source of C. Samples consisting of alternating layers of nominally 25 nm $Si_{1-x-y}Ge_xC_y$ and 35 nm Si for 10 periods were characterized by high-resolution x-ray diffraction, transmission electron microscopy (TEM), and Rutherford backscattering spectrometry to determine strain, thickness, and composition. C resonance backscattering and secondary ion mass spectrometries were used to measure the total C concentration in the $Si_{1-x-y}Ge_xC_y$ layers, allowing for an accurate determination of the substitutional C fraction to be made as a function of growth rate for fixed Ge and substitutional C compositions. For C concentrations close to 1%, high-quality layers were obtained without the use of Sb-surfactant mediation. These samples were found to be structurally perfect to a level consistent with cross-sectional TEM ($<10^7$ defects/cm^2) and showed considerably improved homogeneity as compared with similar structures grown using graphite as the source for C. For higher Ge and C concentrations, Sb-surfactant mediation was found to be required to stabilize the surface morphology. The maximum value of substitutional C concentration, above which excessive generation of stacking fault defects caused polycrystalline and/or amorphous growth, was found to be approximately 2.4% in samples containing between 25 and 30% Ge. The fraction of substitutional C was found to decrease from roughly 60% by a factor of 0.86 as the $Si_{1-x-y}Ge_xC_y$ growth rate increased from 0.1 to 1.0 nm/s. © *1998 American Vacuum Society.*
[S0734-211X(98)06804-8]

I. INTRODUCTION

We report developments in the growth of $Si_{1-x-y}Ge_xC_y$/Si heterostructures that are now making high-quality material available for device fabrication. A promising picture is beginning to emerge that includes the possibility of achieving large band-gap differences between Si and $Si_{1-x-y}Ge_xC_y$ lattice matched to a Si substrate.[1–3] These developments, if realized, would allow for further advancements in heterojunction bipolar transistor (HBT) technologies[4,5] based on Group IV elements, since reliability issues relating to the increased strain in high-Ge content, SiGe HBTs can, in principle, be minimized or eliminated using $Si_{1-x-y}Ge_xC_y$.[6,7] Furthermore, the picture also includes the possibility of achieving significant conduction band offsets for $Si_{1-x-y}Ge_xC_y$ coherently strained or lattice matched to a Si substrate.[3] Together with $Si_{1-x}Ge_x$ or compressively strained $Si_{1-x-y}Ge_xC_y$, complementary metal–oxide–semiconductor (MOS) devices consisting of high-mobility field effect transistors involving both *n*- and *p*-type two-dimensional carrier gas (2DCG) structures are envisioned.[8,9] A $Si_{1-x-y}Ge_xC_y$-based approach might, therefore, allow for implementation of next-generation, high-speed CMOS at integration levels consistent with state-of-the-art Si device processing capabilities.

Fundamentally, we are interested in understanding the maximum amount of Ge and C that can be substitutionally incorporated into $Si_{1-x-y}Ge_xC_y$ in a manner consistent with the growth of device-quality material. We have shown that the use of an e-beam-sublimated graphite source in the molecular beam epitaxy (MBE) of $Si_{1-x-y}Ge_xC_y$ leads to a rough surface morphology and inhomogeneous incorporation for Ge and C concentrations as low as 9% and 1%, respectively.[10] This surface instability was observed to worsen for thicker films, depending on the Ge and C fractions, limiting either the concentrations or the thickness to very low (impractical) levels. Believing that the limitations were due to the presence of immobile molecules of C_2 and C_3 on the surface, we began experimenting with the use of Sb as a surfactant.[11] We discovered that much higher concentrations of both Ge and C (23% and 1.8%, respectively) could be reached in arbitrarily thick layers of defect-free material.[12]

In this article, we present new results which we have obtained through the use of an e-beam-sublimated silicon-carbide (SiC) source for C. The C source substantially

[a]Electronic mail: croke@hrl.com

TABLE I. $Si_{1-x-y}Ge_xC_y$/Si superlattice sample set (10 periods each).

Sample[a]	Sb Coverage	r_{SiGeC} (nm/s)	HRXRD/RBS/CRBS					SIMS		C_s/C_t[d] (%)
			T_{Si} (nm)	T_{SiGeC} (nm)	%Ge	$\%C_s$	$\%C_t$	%Ge	$\%C_t$	
HA96.017	0 ML	0.23	33.0	24.0	10.8	0.9	N/D	N/D	N/D	N/D
HA97.020	0 ML	0.24	35.0	25.0	9.6	1.3	1.9	N/D	1.3[b]	66±42
HA97.055	0.5 ML	0.16	38.5	27.5	23.4	2.2	3.6	N/D	N/D	61±21
HA97.106	0.5 ML	0.09	33.0	22.0	30.6	2.2	3.5	31[c]	3.4[c]	62±22
HA97.111	0.5 ML	0.48	35.0	23.5	29.3	2.4	4.2	28[c]	4.3[c]	56±17
HA97.115	0.5 ML	1.02	33.5	25.0	25.6	2.2	4.4	24[c]	4.1[c]	51±14

[a]Samples HA96.017 and HA97.020 were each grown at 450 °C. Samples HA97.055, .106, .111, and .115 were grown at 500 °C.
[b]SIMS calibration of $\%C_t$ in sample HA97.020 was determined using sensitivity factors derived from a comparison of HRXRD/RBS/CRBS and SIMS data taken from samples HA97.020 and HA97.055, adjusting for changes in matrix due to the different Ge concentrations.
[c]SIMS calibration of samples HA97.106, .111, and .115 was based on HRXRD/RBS/CRBS analysis of HA97.055.
[d]The ratio C_s/C_t was calculated from the values obtained from the HRXRD/RBS/CRBS analysis.

changes the nature of the molecular species reaching the growing $Si_{1-x-y}Ge_xC_y$ layers. We will show that for 9.6% Ge and 1.3% C, use of the SiC source results in homogeneous incorporation of Ge and C, a significant improvement over the results obtained previously with our graphite source. We will also present results from characterizations of $Si_{1-x-y}Ge_xC_y$/Si superlattices of varying Ge and C concentrations in excess of 10% and 1%, respectively, grown using our SiC source. We have used high-resolution x-ray diffraction (HRXRD) to measure the period and the average strain of the superlattices, Rutherford backscattering (RBS), C resonance backscattering (CRBS), and secondary ion mass spectrometries (SIMS) to measure the Ge and C concentrations, and transmission electron microscopy (TEM) to study the microstructural detail. We also present data showing the maximum amounts of Ge and C that can be grown defect-free using this source, including the results of experiments involving Sb-surfactant mediation, and compare these results with those previously obtained using the graphite source. Finally, we describe an analysis of HRXRD, RBS, and CRBS data that, taken together, provide an estimate of the fraction of substitutional C present in the samples as a function of $Si_{1-x-y}Ge_xC_y$ growth rate.

II. EXPERIMENT

Several $Si_{1-x-y}Ge_xC_y$/Si superlattices were grown for this study in a Perkin–Elmer (Model 430S) Si MBE system (base pressure$<1\times10^{-10}$ Torr) designed for solid-source deposition onto heated, 5 in. Si substrates. Source materials consisting of a shaped Si charge, Ge chunks, and a 0.25-in.-thick SiC wafer (obtained from CREE Research, Inc.) were loaded into each of three Temescal electron-beam evaporators. In the case of Ge, a graphite crucible liner was used to provide enhanced stability and more uniform heating than would otherwise be possible if the Ge were placed directly into the copper hearth. For the SiC, a special graphite holder was designed to minimize thermal contact with the hearth and provide a method for containing the pieces of SiC that would inevitably break off from the original wafer during use.

A. $Si_{1-x-y}Ge_xC_y$/Si sample set

(100) Si substrates were prepared *ex situ* by spinning them to 4000 rpm in a Laurell Technologies polypropylene spinner (Model WS-200-4NPP/RV) and pouring 5% HF [1:10 dilution in de-ionized (DI) water] over the polished surface until they became hydrophobic. After a short rinse (5 s in DI water), the wafers were loaded into a buffer chamber and then into the growth chamber. Prior to deposition, the substrates were heated to 850 °C and exposed to a 0.01 nm/s Si flux,[13] in order to remove any oxide contamination remaining from the HF etching process. Wafers prepared in this manner exhibited a clean (2×1)-reconstructed surface, as observed with reflection high-energy electron diffraction (RHEED). Following oxide desorption, the substrate temperatures were reduced to either 450 or 500 °C and a Si buffer layer was grown, typically to a thickness on the order of 50–100 nm. After growth of the buffer layer and prior to deposition of the superlattices, some samples received 1/2 monolayer (ML) predeposit of Sb, to reduce surface diffusion rates, allowing for more uniform incorporation and the preservation of an atomically smooth growth front. Finally, 10-period superlattices, consisting of nominally 25 nm $Si_{1-x-y}Ge_xC_y$/35 nm Si in each bilayer, were grown under various conditions in order to study crystalline quality and substitutional C fraction as a function of composition and growth rate. Growth parameters (temperature and Sb coverage prior to the start of superlattice growth) and results from HRXRD/RBS/CRBS and SIMS characterizations are shown in Table I.

B. SiC vs graphite sources for C deposition

SiC was chosen as a source for C due to the expectation that the nature of the depositing species would change considerably, as compared with use of a graphite source.[14] In a

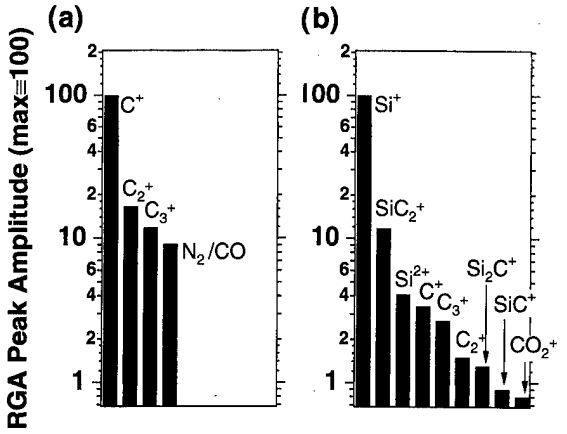


FIG. 1. Relative RGA peak amplitudes for (a) graphite vs (b) silicon-carbide sources. RGA peak amplitude for various constituents generated through e-beam sublimation of graphite and silicon carbide are presented. The data are shown relative the major component (set equal to 100) in each source.

previous paper,[12] we showed that high-quality, homogeneous, $Si_{1-x-y}Ge_xC_y$ layers could be grown using a graphite source for substitutional C concentrations approaching 2%, provided that the growth was Sb mediated. Without Sb mediation, even for layers in which the Ge and C concentrations were only 10% and 1%, respectively, we observed roughening of the $Si_{1-x-y}Ge_xC_y$ surface and evidence of inhomogeneous lateral incorporation. Through the use of the SiC source, we expected to deposit molecular species that contained both Si and C,[14] rather than just C, and hence improve the homogeneity and perhaps, stabilize the surface morphology in the process.

While the SiC source was hot, we scanned, using a residual gas analyzer (RGA), over masses ranging from amu 2 (H_2) to amu 75, in order to characterize the species that sublimated from the source under typical operating conditions. These scans were compared with scans obtained under base vacuum conditions (i.e., sources cold, LN_2 shrouds cold, and cryopump on and pumping) and with those obtained previously using a graphite source. The results are shown in Fig. 1. The data were normalized so that the amplitude of the RGA signal for the major constituent in each source was set equal to 100. For the graphite source, monomeric C (amu 12) appears as the peak with the largest signal amplitude but contributions from C_2 (amu 24) and C_3 (amu 36) were also significant. In comparison, the major constituent sublimating from the SiC source, was found to be Si, although we also observed SiC_2 and, to a lesser extent, C, C_2, and C_3. Closed-loop control of the SiC source (and approximately, the amount of C contained in the flux) was achieved by tuning the RGA to amu 52 (SiC_2) and automatically adjusting the power to the electron gun with a computer to maintain a predetermined setpoint value.

C. Method for computing sample composition from HRXRD, RBS, and TEM

Superlattice composition and layer thicknesses were calculated from data obtained from a combination of measurements involving HRXRD, RBS, CRBS, and TEM. Four pieces of information were required, in addition to the shutter opening and closing times, to uniquely determine the structure. We used HRXRD to provide the average strain and superlattice period, RBS to measure the average Ge concentration, and TEM lattice imaging to obtain the Si layer thickness. From these data, the four relevant fluxes, identified as (1) Si from the primary Si source, (2) Si from the SiC source, (3) Ge, and (4) C, were calculated assuming unity sticking coefficient for the various constituents in a manner consistent with Vegard's Law and the lattice constants for Si, Ge, and β-SiC. Finally, from the fluxes and the shutter opening and closing times, we calculated the layer thickness and composition of the $Si_{1-x-y}Ge_xC_y$ layers. CRBS was used to determine the total C concentration ($\%C_t$ in Table I), since only the substitutional C concentration, y ($\%C_s$ in Table I), could be accessed by the above method. The ratio, $\%C_s/\%C_t$, also shown in Table I, therefore represents the percentage of incorporated C contributing to the strain.

(004) HRXRD spectra were obtained by collimating Cu $K\alpha$ radiation, selected through the aid of a four-crystal Ge monochromator, onto each sample in a θ–2θ geometry. The spectra generally consisted of a Si substrate peak at 34.5644° and a series of superlattice peaks. Analysis of superlattice period and average strain followed the method of Speriosu and Vreeland.[15] Standard 2 MeV He^{2+} backscattering was used for probing Si and Ge, whereas 4.3 MeV He^{2+} C-resonance backscattering probed the C component. These ion backscattering techniques were used to measure the average Ge and C concentrations, since individual layers could not be resolved. The estimated uncertainties associated with the RBS and CRBS measurements were ±0.5 at. % for both the average Ge and average C concentrations.

SIMS characterization was performed immediately after each growth to provide a convenient method for verifying sample composition from run to run. The analysis system consisted of a Perkin–Elmer (Model 595) Auger Microprobe equipped with a 3500 SIMS II attachment. A 2 keV O_2^+ primary beam was sufficient to resolve each individual $Si_{1-x-y}Ge_xC_y$ layer. The sample HA97.055 was used as a calibration standard as it had been previously characterized by HRXRD/RBS/CRBS.

III. RESULTS AND DISCUSSION

A. TEM micrographs compare growth using graphite vs SiC sources

In Fig. 2, we compare cross-sectional TEM micrographs obtained from two $Si_{0.90}Ge_{0.10}C_{0.01}$/Si superlattice samples grown without the benefit of Sb mediation. The images were taken on a Philips EM430 operated at 200 keV using a (400) two-beam condition near a [011] zone axis. The sample on the left side of the figure (HA96.017) was grown using a graphite source, while the sample on the right (HA97.020) was grown using the SiC source. The micrograph of HA96.017 shows contrast *within* the $Si_{1-x-y}Ge_xC_y$ layers, suggesting that incorporation was laterally inhomogeneous. HA97.020 shows no such contrast and reveals the layers to

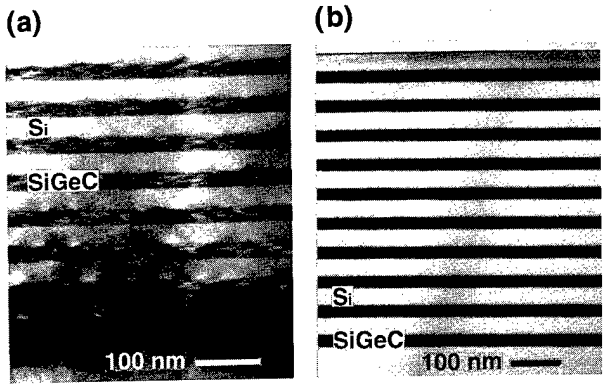


FIG. 2. Cross-sectional TEM micrographs compare a superlattice grown using the graphite source vs a superlattice grown using the silicon-carbide source. (a) Graphite source: HA96.017 240 Å $Si_{0.883}Ge_{0.108}C_{0.009}$/330 Å Si (no Sb-mediation); (b) silicon-carbide source: HA97.020 250 Å $Si_{0.891}Ge_{0.096}C_{0.013}$/350 Å Si (no Sb-mediation). Significant improvement of homogeneity and interface abruptness due to the use of a silicon-carbide source are demonstrated for two $Si_{0.89}Ge_{0.10}C_{0.01}Si$ superlattices. Contrast within the $Si_{1-x-y}Ge_xC_y$ layers and roughening of the interfaces for the case in which Si was grown on $Si_{1-x-y}Ge_xC_y$ are seen for sample HA96.017, but not for HA97.020.

be quite uniform and the interfaces, abrupt. The interfaces in HA96.017 are sharp only for the case where $Si_{1-x-y}Ge_xC_y$ was grown on Si. These observations were confirmed by RHEED studies of the growing $Si_{1-x-y}Ge_xC_y$ layers for which the patterns became spotted. No such spottiness was observed at these compositions for samples grown using the SiC source.

At higher C concentrations, even growth using the SiC source becomes problematic, although the results still represent an improvement over the graphite source. For the case of a 23% Ge, 2.5% C superlattice (composition estimated from analysis of a previous sample), the RHEED pattern at the end of the last period was spotted, indicating a rough surface. There were no peaks observed in HRXRD and so we concluded that the sample must have contained a high density of stacking defects (although the material was basically single crystal). Growth at these compositions using graphite, without the surfactant, had typically resulted in amorphous films.

B. Surfactant-mediated growth of high-Ge, high-C content superlattices using the SiC source

As we found with the graphite source, growing $Si_{1-x-y}Ge_xC_y$ at compositions in excess of 20% Ge and 2% C requires the use of a surfactant to stabilize the surface and enhance uniform incorporation of the constituents. For the following, all samples received 1/2 ML Sb predeposit prior to growth of the first $Si_{1-x-y}Ge_xC_y$ layer. Two series of samples were grown at 500 °C at nominal $Si_{1-x-y}Ge_xC_y$ deposition rates of 0.1, 0.5, and 1.0 nm/s in order to determine the dependence of growth rate on material quality and substitutional C fraction at higher Ge and C concentrations. Compositions were selected near the condition for perfect lattice match (approximately 9.4:1 Ge:C ratio) and about which we had previously found stacking fault defects began to appear. These nominal concentrations were 25% Ge/2.5% C for the first series and 25% Ge/2.0% C for the second series.

In the first series (samples not listed in Table I), we found that at a $Si_{1-x-y}Ge_xC_y$ growth rate of 0.1 nm/s, the RHEED pattern after the last Si layer was spotted, indicating that the layers were single crystal. Further analysis with HRXRD and TEM confirmed that a very high density of stacking defects was present. At 0.5 nm/s, the RHEED pattern disappeared entirely before growth of the superlattice was completed, indicating that the layers were amorphous. An attempt to grow this composition at 1.0 nm/s was deemed unnecessary since increasing the growth rate had been found to produce an amorphous structure.

In the second series, we reduced the C flux slightly and obtained the results shown in Table I for samples HA97.106, .111, and .115. In Fig. 3(a), we show a cross-sectional TEM micrograph from sample HA97.106. HRXRD/RBS analysis determined the Ge and C concentrations in the $Si_{1-x-y}Ge_xC_y$ layers to be 30.6±1.3% and 2.17±0.14%, respectively. The uncertainties reflect those associated with the four inputs to the calculation, propagated independently and then summed in quadrature. The analysis also determined the $Si_{1-x-y}Ge_xC_y$ growth rate to be 0.090±0.003 nm/s. The sample is free of extended defects and inhomogeneities over an interface length scale of approximately 100 μm, resulting in an upper limit on the defect density of 10^7 cm^{-2}.[16]

In Fig. 3(b), we show SIMS profiles for Si, Ge, and C as a function of depth into sample HA97.106. As mentioned earlier, HA97.055 was used as a calibration standard for the purpose of converting the raw SIMS data (counts) into concentrations. By averaging the data shown in Fig. 3(b) over the first three superlattice layers, we calculated a value of 3.42±1.16 at. % for the total C concentration. This compares to a value of 3.62±1.21 for HRXRD/CRBS analysis of the sample directly. The percentage of substitutional C was then determined by taking the ratio of the substitutional C concentration (measured by HRXRD/RBS) and the total C concentration (measured by HRXRD/CRBS), yielding 62±22%.

C. Results from substitutional C fraction vs growth rate experiments

Two additional samples were grown in the second series and received approximately the same substitutional C concentration as HA97.106. The growth rate was increased to 0.5 nm/s in sample HA97.111 and finally, to 1.0 nm/s in sample HA97.115. As mentioned previously, TEM micrographs revealed HA97.106 to be free from defects. However, samples .111 and .115 showed signs of a breakdown in epitaxial quality as increasing numbers of stacking faults and threading dislocations were found as the growth rate was increased. HA97.111 and .115 were analyzed for composition, in the same manner as for HA97.106. The results are shown in Table I. In Fig. 4(a), we have plotted the substitutional C fraction as a function of growth rate for samples HA97.055, .106, .111, and .115. The large uncertainties as-

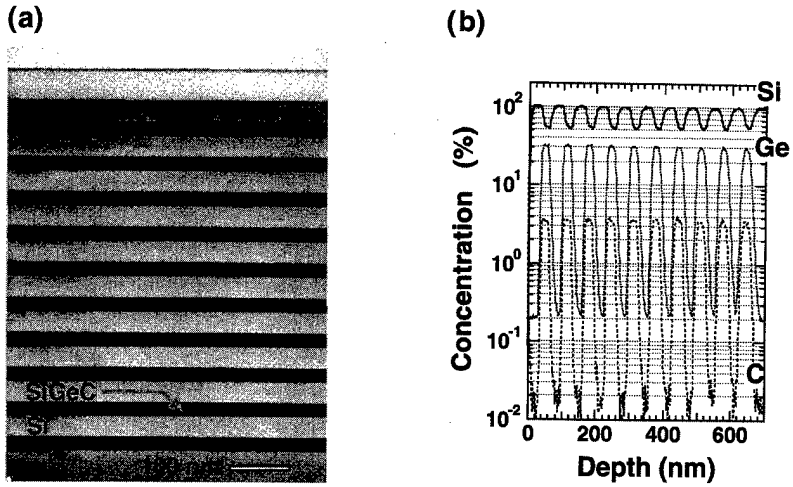


FIG. 3. Cross-sectional TEM micrograph and SIMS profile from sample HA97.106, a 33 nm Si/22 nm $Si_{0.672}Ge_{0.306}C_{0.022}$ superlattice grown using the graphite source. In (a), a TEM micrograph taken from sample HA97.106 demonstrates that perfect structural quality was obtained at a composition of 31% Ge and 2.2% C using the silicon-carbide source. 1/2 ML Sb was deposited on a Si buffer layer prior to growth of the first $Si_{1-x-y}Ge_xC_y$ layer. (b) SIMS data showing Si, Ge, and C concentrations vs depth for the sample in (a) are plotted. The data were calibrated using HA97.055 as a standard, characterized for composition by HRXRD/RBS/CRBS.

sociated with the data arise primarily from the difficulty in obtaining an accurate measurement of the total C concentration by CRBS. We found that we could eliminate much of the uncertainty and obtain a relative measure of the substitutional C fraction by plotting the substitutional C concentration divided by the relative count rate for C in the $Si_{1-x-y}Ge_xC_y$ layers to Si in the Si layers $[\%C_s/(N_C/N_{Si})]$ as measured by SIMS [Fig. 4(b)]. The data show that the substitutional C fraction decreases to a level of approximately 86% of the starting value (from roughly 60% to about 50%) with increasing $Si_{1-x-y}Ge_xC_y$ growth rate.

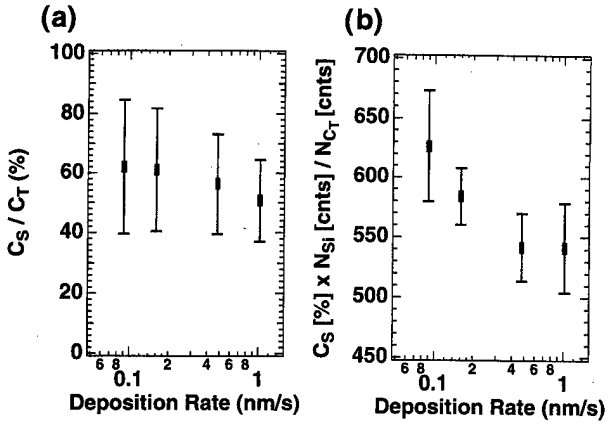


FIG. 4. Plot of substitutional carbon fraction for several samples vs $Si_{1-x-y}Ge_xC_y$ growth rate. (a) The substitutional C fraction, obtained by dividing the substitutional C concentration (measured by HRXRD/RBS) by the total C concentration (measured by CRBS) is plotted vs $Si_{1-x-y}Ge_xC_y$ deposition rate for several samples. The error bars on the data arise primarily from uncertainty in the CRBS measurement of the average C concentration. In (b), we obtained a relative measure of the substitutional C fraction using the raw data from SIMS depth profiles performed on the four samples. The data represented are now only proportional to the substitutional C fraction, yet the uncertainty has been reduced to such an extent that a trend toward lower substitutional fractions at higher growth rates can be identified.

IV. CONCLUSION

We have studied the growth of $Si_{1-x-y}Ge_xC_y$/Si superlattices by solid-source MBE through the use of a SiC source for C. Sample composition and thickness information obtained by HRXRD, RBS, CRBS, and SIMS were used to develop an understanding of the role of growth rate on crystalline quality and substitutional C fraction, with and without the use of Sb surfactant mediation. At 10% Ge/1% C (substitutional), $Si_{1-x-y}Ge_xC_y$/Si superlattices were found by TEM to grow defect-free, without the need for Sb-surfactant mediation. The constituent Ge and C atoms incorporated in a more uniform manner, as compared with similar samples grown using a graphite source. For Ge and C concentrations approaching 25% Ge and 2.5% C, growth without surfactant mediation resulted in amorphous films. At slightly lower C concentrations, we found that defect-free material could be grown using 1/2 ML predeposit of Sb prior to growth of the superlattices. For Sb-mediated samples, we found that at higher growth rates, crystalline quality became degraded, and the substitutional C fraction decreased slightly. The SIMS data provided a measurement of the relative decrease in substitutional C: ~86% as much at 1.0 nm/s than at 0.1 nm/s. The absolute substitutional fraction is less accurately known, due to the larger uncertainty associated with the CRBS measurement. We measured a C substitutional fraction of 62±22% at 0.09 nm/s using the SiC source.

ACKNOWLEDGMENT

Partial support for this work was provided by the Defense Advanced Research Projects Agency (DARPA) monitored

by Lt. Col. Gernot Pomrenke under Contract No. MDA972-95-3-0047.

Presented at the Silicon Heterostructures Conference, Barga, Italy, 15–19 September 1997.

[1]S. Furakawa, H. Etoh, A. Ishizaka, and T. Shimada, U.S. Patent No. 4 885 614 (1989).
[2]R. A. Soref, J. Appl. Phys. **70**, 2470 (1991).
[3]B. L. Stein, E. T. Yu, E. T. Croke, A. T. Hunter, T. Laursen, A. E. Bair, J. W. Mayer, and C. C. Ahn, Appl. Phys. Lett. **70**, 3413 (1997).
[4]D. L. Harame, J. M. C. Stork, B. S. Meyerson, K. Y.-J. Hsu, J. Cotte, K. A. Jenkins, J. D. Cressler, P. Restle, E. F. Crabbé, S. Subbanna, T. E. Tice, B. W. Scharf, and J. A. Yasaitis, Tech. Dig. Int. Electron Devices Meet., 71 (1993).
[5]E. Kasper, A. Gruhle, and H. Kibbel, Tech. Dig. Int. Electron Devices Meet., 79 (1993).
[6]K. Eberl, S. S. Iyer, S. Zollner, J. C. Tsang, and F. K. LeGoues, Appl. Phys. Lett. **60**, 3033 (1992).
[7]K. Eberl, S. S. Iyer, and F. K. LeGoues, Appl. Phys. Lett. **64**, 739 (1994).
[8]K. Rim, J. Welser, and J. L. Hoyt, Tech. Dig. Int. Electron Devices Meet., 517 (1995).
[9]K. Ismail, Tech. Dig. Int. Electron Devices Meet., 509 (1995).
[10]E. T. Croke, A. T. Hunter, C. C. Ahn, T. Laursen, D. Chandrasekhar, A. E. Bair, David J. Smith, and J. W. Mayer, J. Cryst. Growth **175/176**, 486 (1997).
[11]H. J. Osten, E. Bugiel, and P. Zaumseil, J. Cryst. Growth **142**, 322 (1994).
[12]E. T. Croke, A. T. Hunter, P. O. Pettersson, C. C. Ahn, and T. C. McGill, Thin Solid Films **294**, 105 (1997).
[13]D. C. Streit and F. G. Allen, J. Appl. Phys. **61**, 2894 (1987).
[14]A. R. Verma and P. Krishna, *Polymorphism and Polytypism in Crystals* (Wiley, New York, 1966), p. 99.
[15]J. S. Speriosu and T. Vreeland, J. Appl. Phys. **56**, 1591 (1984).
[16]The defect density of $<1\times10^7$ cm^{-2} was calculated based on the assumption that the TEM sample was 100 nm thick and that we had observed defect-free material over approximately 100 μm of interface.

Low temperature scanning tunneling microscope-induced luminescence of GaN

S. Evoy,[a] C. K. Harnett, and H. G. Craighead
School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853

T. J. Eustis
Department of Materials Science and Engineering, Cornell University, Ithaca, New York 14853

W. A. Davis, M. J. Murphy, W. J. Schaff, and L. F. Eastman
School of Electrical Engineering, Cornell University, Ithaca, New York 14853

(Received 3 February 1998; accepted 22 May 1998)

We report the low temperature scanning tunneling microscope-induced luminescence of molecular beam epitaxy grown α-GaN. Semiquantitative spectroscopic analysis suggests near band edge emission, as well as emission covering the rest of the visible range. The relative intensity of band edge emission increases by one order of magnitude under liquid helium cooling. We also report the first photon emission images of GaN obtained with this technique. These images reveal stronger band edge emission at the center of crystallites. This study is complemented with a scanning electron microscope-induced cathodoluminescence analysis. Cathodoluminescence is dominated by the hexagonal (D°, X) transition and reveals evidence of small quantities of the cubic phase.
© *1998 American Vacuum Society.* [S0734-211X(98)10104-X]

I. INTRODUCTION

The GaN system is of great interest for optoelectronic applications in the green, blue, and UV spectral emission regions. The last few years have seen such accomplishments as the production of highly efficient GaN-based blue and green light-emitting diodes,[1] as well the realization of GaN-based laser diodes.[2–4] Such advances have been made in spite of various issues such as high dislocation densities[5] and the defect induced "yellow luminescence."

These issues prompted an ongoing interest in spatially resolved cathodoluminescence (SRCL) studies of this material.[6–10] For example, SRCL analysis suggests that yellow luminescence is associated with the presence of extended defects and specifically with the existence of low angle grain boundaries.[9] Another study documents the luminescence from both hexagonal and cubic phase crystallites in a given sample, as well as the spatial dependence of specific radiative transitions in the cubic phase.[8]

Such studies are traditionally performed by fitting collection optics in a scanning electron microscope (SEM). While the resolution of the technique is limited by the extent of the generation volume,[11] other methods can be used for carrier injection. The scanning tunneling microscope (STM) offers nanometer scale resolution[12] and an accurate control of the injection bias. These additional advantages offer great potential for nanoscale luminescence studies of optoelectronic materials. The STM-induced luminescence (STL) of GaAs heterostructures has first been reported by Abraham *et al.*[12] and Renaud and Alvarado.[13] The technique was then applied to the imaging of quantum wires[14–16] and dots.[17–19] However, there have been few reports of STL of the nitride system.[20,21]

The STL technique differs fundamentally from traditional SEM-induced cathodoluminescence (SEM-CL). In the latter, the electron beam acts as a primary source of power, generating carriers through impact ionization in a roughly isotropic manner. The STM, however, introduces minority carriers directly into one of the energy bands. These carriers will then recombine with the opposite type already available in the material.[12] Given that surface states represent efficient traps, it is suggested that these carriers must reach the bulk to generate radiative recombination events. Considering band bending effects at the surface (Fig. 1), this condition leads to a parabolic dependence of luminescence efficiency on tip bias, and also restricts the effective injection angle leading to photon emission.[13]

We have performed both the STM-induced and SEM-induced luminescence of GaN grown by molecular beam epitaxy (MBE). The STL of metalorganic chemical vapor deposition (MOCVD)-grown GaN has recently been reported.[20,21] These studies were performed at room temperature and did not include images of the photon emission. This is the first report of low temperature STL of GaN, and the first report of STL emission images of this material. Furthermore, while some previous authors operated the STM tip as a field emitter,[21] our work was completely performed with tunneling injection. These previous reports will be included in the discussion of our results. Finally, the STL study is complemented with low temperature SEM-induced cathodoluminescence (SEM-CL).

II. EXPERIMENT

The STL experiments were carried through in a JEOL JSTM-4500 microscope equipped with a cold finger for liquid helium (LHe) cooling. Actual sample temperature is expected to range between 20 and 30 K.[22] Pressure in the observation chamber lies in the upper 10^{-11} Torr range. A 1.5

[a]Electronic mail: se20@cornell.edu

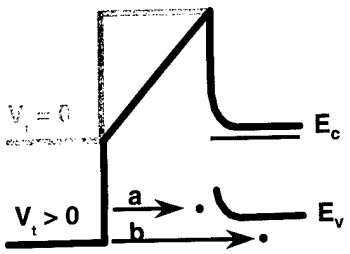


FIG. 1. Energy diagram of the STM-induced luminescence of a highly doped *n*-type semiconductor. Under no external bias, Fermi levels of tip and surface align. Under sufficient bias, tip Fermi level is aligned with valence band, and holes are injected in semiconductor (electrons are tunneled out). Injected at low biases, carriers (a) are trapped at surface. Injected at higher biases, carriers (b) may reach bulk and recombine radiatively.

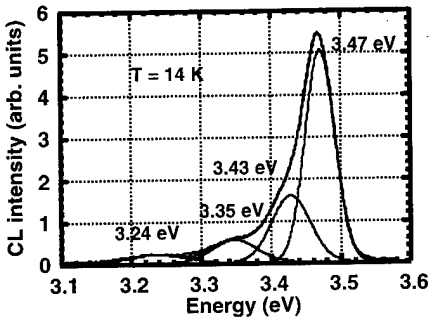


FIG. 2. SEM-CL spectrum of a 30 μm×30 μm area (5 kV, 100 pA, T = 14 K). Spectrum is deconvolved into four Gaussian peaks.

in. lens is located 1.5 in. away from the tunnel region at an angle of 55° with the normal of the surface. Collected photons are forwarded through the UHV window and refocused onto the GaAs cathode of a Hamamatsu R943-02 photomultiplier tube (PMT). Considering the solid angle covered by the lens, the transmission factor of the various components, and the efficiency of the PMT, the overall quantum efficiency of the detection system is evaluated at 1%. Bandpass interference filters are used to isolate specific wavelength ranges. Tips are prepared from tungsten wires using a NaOH (2 M) dc electrochemical etch.

The SEM-CL is performed in a Zeiss 960 microscope modified with a custom built cooling stage. Luminescence is collected with a parabolic mirror, forwarded through a fused silica window, focused onto the entrance slit of a 0.5 m monochromator, and detected with a R943-02 PMT. All SEM-CL results are obtained using a 5 kV acceleration voltage, corresponding to penetration depth of 0.3 μm in GaN.[9]

Experiments are performed on Si-doped (*n*-type) α-GaN films grown on sapphire by MBE. An Oxford Car25 rf plasma source is used as the nitrogen source. A 85 nm GaN buffer layer is first grown at 850 °C, followed by a 250 nm undoped layer grown at 970 °C and a 2.8 μm thick doped layer also grown at 970 °C. Doping concentration of $n = 2.5\times10^{-18}$ cm^{-3} and mobility of $\mu = 113$ cm^2/V s are measured with the four-point Hall technique.

Samples are cleaned in a series of sonicated solvents, dipped in 30% HF for 1 min,[23] rinsed in distilled water, and dried under nitrogen flow. Prior to STM analysis, the sample is also degassed in the preparation chamber (mid-10^{-10} Torr range) using indirect heating at 500 °C for 1 h.

III. RESULTS AND DISCUSSION

A. SEM-CL

A cathodoluminescence spectrum obtained at $T = 14$ K from an area of 30 μm×30 μm is shown in Fig. 2. Deconvolution reveals the dominance of a 3.47 eV peak, which is attributed to the hexagonal (D°,X) transition, as well as the presence of its phonon replica shoulder at 3.35 eV. The other shoulders observed at 3.43 and 3.24 eV are attributed to the (A°,X) and (D°,A°) transition, respectively. Considering the very faint intensity of the shoulders, these results are in good agreement with those reported in a previous CL study of hexagonal GaN.[8] No luminescence is detected in the visible range with either CL or photoluminescence (PL). While these data are consistent with the dominance of the hexagonal phase, some individual crystallites produce spectra which are exceptionally different from the rest of the surface (Fig. 3). In Fig. 3(b), while the 3.47 eV luminescence from location No. 1 is similar to the average, location No. 2 generates a strong spectrum in the 3.27–3.31 eV range. Location No. 2 is also an order of magnitude brighter than the rest of the surface. Figures 3(c) and 3(d) show luminescence images which are spectrally resolved at 3.306 and 3.463 eV, respectively. The complementary features show that the two emission wavelengths are spatially segregated in crystallites 200–500 nm in size. While optical analyses of cubic GaN evaluate its band gap at 3.302 eV,[24] the CL emission from this phase was reported at 3.272 eV at 5 K.[8] Observed on less than 1% of our sample surface, the spatially segregated 3.3 eV emission therefore suggests small amounts of nanometer scale cubic crystallites in a mostly hexagonal material.

B. STL

A typical topographic image of the GaN surface is shown in Fig. 4. This image exhibits crystallites ranging from 200 to 600 nm in size, in agreement with the SEM-CL observations. While the surface reveals an overall roughness of 30 nm rms, the surfaces of individual crystallites are as smooth as 1 nm.

Once a topographic image is acquired, the tip is positioned over a specific crystallite. At constant tip bias, intensity is directly proportional to the tunnel current, suggesting that no saturation effect is taking place. The dependence of luminescence intensity on tip bias at constant current is reported in Fig. 5. The bias is scanned from $V_t = 1.5$ V to $V_t = 4.2$ V with a current of $I_t = 5$ nA, resulting in a total tip retraction of 6 nm from the starting position. Luminescence reveals a voltage threshold of 1.5–1.7 V, followed by the parabolic increase induced by band bending effects.[13,20] Threshold is lower than expected when considering the picture presented in Fig. 1. This observation is further discussed below. No luminescence is observed with a negative tip polarity, as expected from a *n*-doped material.

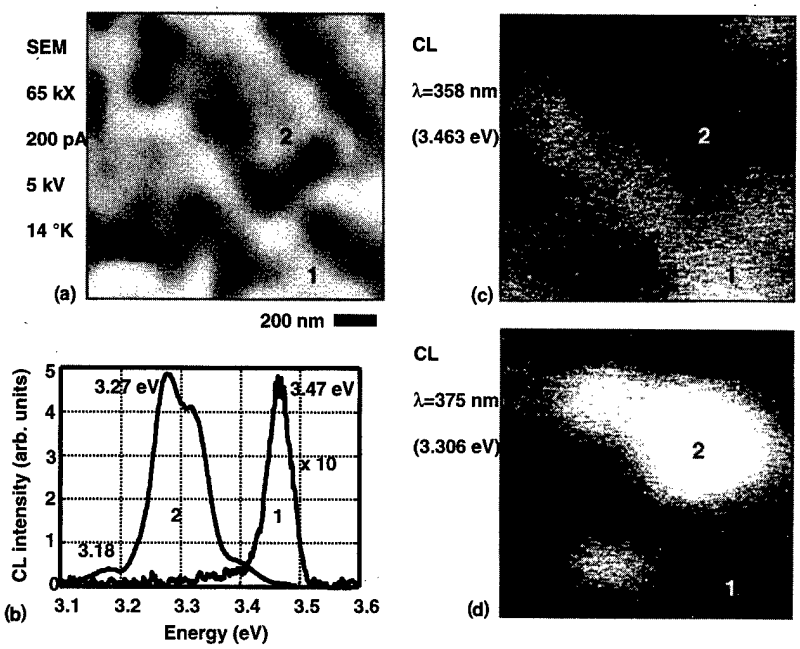


FIG. 3. Spatially resolved CL data of GaN surface (5 kV, 50 pA, $T=14$ K). (a) SEM micrograph (b) spatially resolved spectra. (c) and (d) CL images spectrally resolved at 3.463 and 3.306 eV, respectively.

A set of interference filters covering the visible range is used for a semiquantitative analysis of the emission spectrum. At room temperature, luminescence reveals a faint and broad emission covering the entire visible spectrum and peaking between 700 and 800 nm. No significant signal is detected in the expected 350±35 nm range. With a tip bias and current of $V_t=4.0$ V and $I_t=750$ nA, the average count rates from 1.5 μm×1.5 μm areas is 25 cps (at 350±35 nm) and 315 cps (at 550±35 nm). Under liquid He cooling, however, a drastic increase of band edge emission is observed. Using same sample and tip, these average rates become 1300 cps (at 350±35 nm) and 100 cps (at 550±35 nm), representing a two order of magnitude increase of the "UV" to "visible" emission ratio. While this material is grown by MBE, similar results were observed on MOCVD-grown GaN.[25] The smoother MOCVD material also yields faint visible emission at room temperature, and a sharp increase in the 350±35 nm range at low temperature. This relative increase of near band edge luminescence indicates that temperature dependent phenomena affect the STL process, as is the case in more standard luminescence experiments.

The visible emission observed at room temperature is somewhat unexpected considering that no such luminescence was observed with standard CL. A similar phenomenon was previously reported in Ref. 21 in which, in contrast to PL results, the yellow luminescence dominated the STM-excited luminescence spectrum. However, this study operated the STM tip as a field emitter, thus involving a completely different carrier injection process. In another study involving tunneling injection instead, the previously discussed low voltage threshold at room temperature has also been reported and has been attributed to the yellow luminescence observed in PL.[20] However, no spectral analysis of the STL emission was presented in support of the possibility. In significant

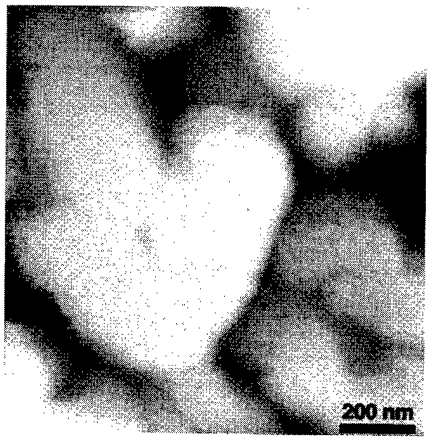


FIG. 4. STM constant-current topographic image of GaN surface ($V_t=3.5$ V, $I_t=1$ nA, room temperature).

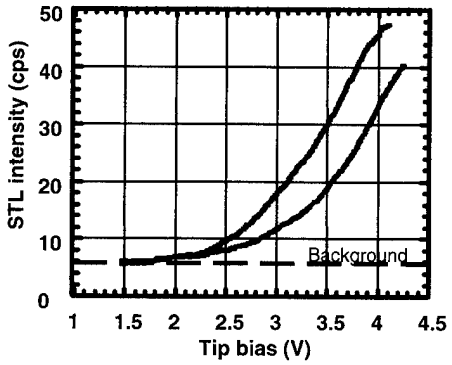


FIG. 5. Dependence of spectrally integrated luminescence intensity with tip bias at constant current (5 nA) and at room temperature. Data are acquired by positioning tip on top of a crystallite. Curves were acquired with same tip on two different crystallites. Each curve represents the average of 64 consecutive slowly ramped measurements.

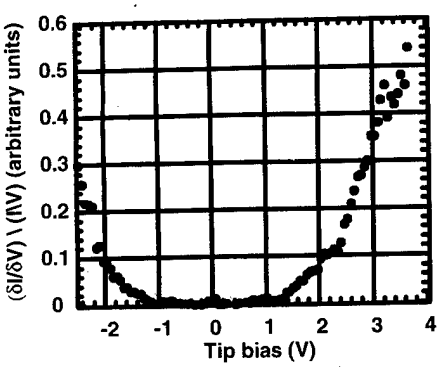


FIG. 6. STM I/V curve of GaN crystallite at room temperature. At low bias, $(\delta I/\delta V)/(I/V)$ is proportional to the surface density of states.

contrast to this last report, neither CL nor PL of our material reveals any defect-induced visible luminescence from the bulk. Nevertheless, our room temperature STL still exhibits a low voltage threshold and a faint visible emission peaking in the near-IR.

A study of the density of state (DOS) through the acquisition of tunnel current versus bias (I/V) spectra offers a preliminary explanation of this disparity. At low voltages, the DOS is proportional to $(\delta I/\delta V)/(I/V)$.[26] A typical DOS spectrum of a GaN crystallite is shown in Fig. 6. Expecting the bulk Fermi level to be close to the conduction band, the origin of the DOS curve should correspondingly be shifted towards the states located at the left of the spectrum. The experimental DOS spectrum is almost centered between the two bands, with the valence band states (positive tip bias) showing a low threshold voltage of $V_t = 1.2$ V. The length of the DOS "gap" is also smaller than the expected bulk band gap of 3.5 eV. These observations imply surface-related phenomena such as surface states, contamination, and Fermi pinning. The faint visible emission and the low bias threshold observed in room temperature STL may therefore have more to do with such surface issues than with the actual properties of the bulk. Tip-related phenomena must also be considered. These issues question the relevance of the luminescence observed at room temperature, and suggest the necessity of low temperature for the imaging of band edge emission.

Interference filters were used to generate spectrally resolved photon emission maps. Figure 7 shows (a) topographic and luminescence images of the surface under LHe cooling. The luminescence images are spectrally filtered at (b) 350±35 nm and (c) 550±35 nm. The peanut-shaped structure at the center of these images may consist of one or two separate crystallites. As illustrated in the profiles found in Fig. 8, the 350±35 nm emission is stronger at the center of the crystallites, and tend to drop off as the tip moves within 100 nm of the structure's edge. On the other hand, the 550±35 nm emission image reveals a modest increase of intensity around the edges of the crystallite, which is especially apparent in the bottom part of the peanut-shaped structure. As discussed previously, other surface-related phenomena may produce significant emission peaking elsewhere in the visible range, and may affect the 550±35 nm emission maps. The correspondence between this last observation and the well documented defect-induced yellow emission therefore remains unclear. However, observed in most other locations, the spatial behavior of the 350±35 nm emission described above suggest that the diffusion of carriers to the crystallite edges inhibits near band edge luminescence through either nonradiative or faint visible radiative processes.

IV. CONCLUSION

We have observed the luminescence of GaN induced by both the scanning tunneling microscope and the scanning

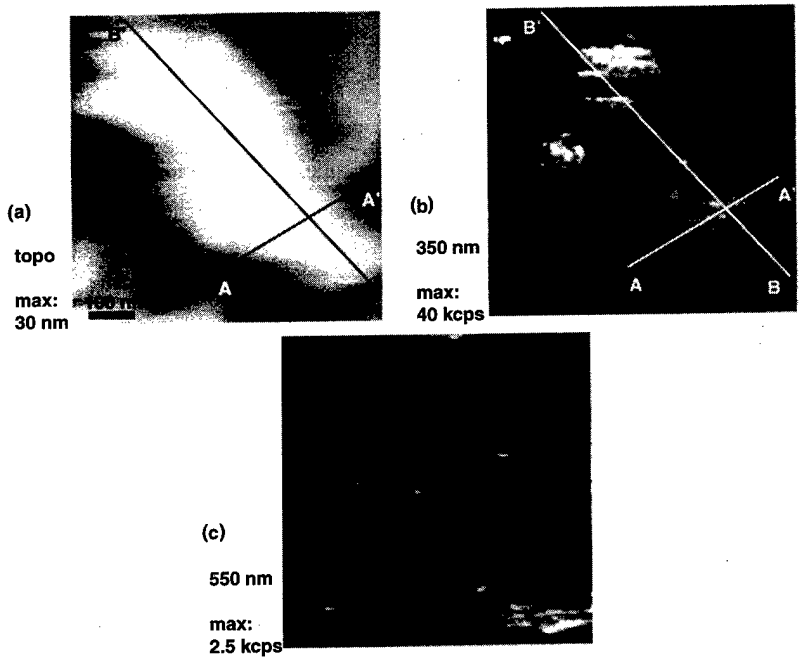


FIG. 7. (a) STM topographic image of GaN region. (b) and (c) STL images of same region. Luminescence is spectrally filtered at 350±35 and 550±35 nm, respectively ($V_t = 4.0$ V, $I_t = 750$ nA, LHe cooled). Central peanut-shaped structure may consist of one or two separate crystallites. Figure 8 shows topographic and band edge luminescence profiles along cut lines.

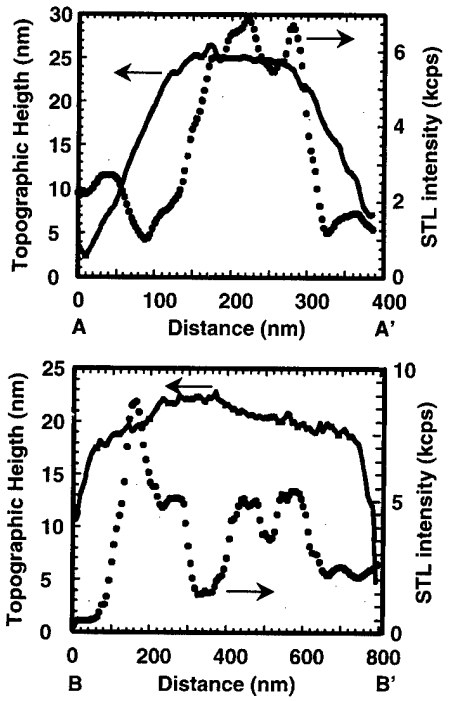


FIG. 8. Topographic (solid line) and STL filtered at 350±35 nm (dotted line). Profiles defined along cut lines shown in Fig. 7 (V_t=4.0 V, I_t=750 nA, LHe cooled).

electron microscope. SEM-CL results reveal the dominance of the ($D°$, X) hexagonal transition at 3.47 eV, with no visible emission within the detection limit. SEM-CL also reveals evidence of small quantities of the cubic phase in the material. The STM was also used for the analysis of individual crystallites. Luminescence threshold voltage, density of states curves as well significant emission covering the visible range suggest surface-related issues and Fermi pinning. Furthermore, the relative intensity of UV emission increases under LHe cooling. While visible luminescence is somewhat stronger at the edges of the crystallites, UV emission is stronger in their centers. This result supports the inhibition of band edge emission at the edges of crystallites.

The STL still requires significant theoretical formalization in order to clarify issues such as lateral and depth resolution, diffusion and recombination carriers, and the nature of the outgoing photons. The impact of surface-related phenomena on the technique should also be addressed. However, these results demonstrate the potential of technique for spatially resolved nanoscale studies of GaN luminescence. Such potential is currently being applied to the analysis of GaN heterostructures, and shows great promises for the development and optimisation of materials for blue emitting devices.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Francesca Pardo and Dr. David Tanenbaum for their invaluable input on the experimental setup. This work was funded by NSF through the Cornell Materials Science Center and the DoD Tri-Services Center for Atomic Scale Spectroscopy.

[1]S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).
[2]S. Nakamura, M. Senoh, S.-I. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Appl. Phys. Lett. **68**, 2105 (1996).
[3]G. E. Bulman, K. Doverspike, S. T. Sheppard, T. W. Weeks, H. S. Kong, H. M. Dieringer, J. A. Edmond, J. D. Brown, J. T. Swindell, and J. F. Schetzina, Electron. Lett. **33**, 1556 (1997).
[4]M. P. Mack, A. Abare, M. Aizcorbe, P. Kozodoy, S. Keller, U. K. Mishra, L. Coldren, and S. DenBaars, Mar. Res. Soc. Internet Journal of Nitride Semiconductor Research **2**, #41 (http://nsr.mij.mrs.org/2/41/).
[5]S. D. Lester, F. A. Ponce, M. G. Craford, and D. A. Steigerwald, Appl. Phys. Lett. **66**, 1249 (1995).
[6]K. Hiramatsu, H. Amano, and I. Akasaki, J. Cryst. Growth **99**, 375 (1990).
[7]D. M. Hofman, D. Kovalec, G. Steude, B. K. Meyer, A. Hoffmann, L. Eckey, R. Heitz, T. Detchprom, H. Amano, and I. Akasaki, Phys. Rev. B **52**, 16 702 (1995).
[8]J. Menniger, U. Jahn, O. Brandt, H. Yang, and K. Ploog, Phys. Rev. B **53**, 1881 (1996).
[9]F. A. Ponce, D. P. Bour, W. Gotz, and P. J. Wright, Appl. Phys. Lett. **68**, 57 (1996).
[10]C. Trager-Cowan, K. P. O'Donnell, S. E. Hopper, and C. T. Foxon, Appl. Phys. Lett. **68**, 355 (1996).
[11]B. G. Yacobi and D. B. Holt, *Cathodoluminescence Microscopy of Inorganic Solids* (Plenum, New York, 1990), p. 83.
[12]D. L. Abraham, A. Veider, Ch. Schonenberger, H. P. Meier, D. J. Arent, and S. F. Alvarado, Appl. Phys. Lett. **56**, 1564 (1990).
[13]P. Renaud and S. F. Alvarado, Phys. Rev. B **44**, 6340 (1991).
[14]L. Samuelson, A. Gustaffsson, J. Lindahl, L. Montelius, M.-E. Pistol, J.-O. Malm, G. Vermeire, and P. Demeester, J. Vac. Sci. Technol. B **12**, 2521 (1994).
[15]M. Pfister, M. B. Johnson, S. F. Alvarado, H. W. M. Salemink, U. Marti, D. Martin, F. Morier-Genoud, and F. K. Reinhart, Appl. Phys. Lett. **65**, 1168 (1994).
[16]M. Pfister, M. B. Johnson, S. F. Alvarado, H. W. M. Salemink, U. Marti, D. Martin, F. Morier-Genoud, and F. K. Reinhart, Appl. Phys. Lett. **67**, 1459 (1995).
[17]L. Samuelson, N. Carlsson, P. Castrillo, A. Gustafsson, D. Hessman, J. Lindahl, L. Montelius, A. Petersson, M.-E. Pistol, and W. Seifert, Jpn. J. Appl. Phys., Part 1 **34**, 4392 (1995).
[18]J. Lindahl, M.-E. Pistol, L. Montelius, and L. Samuelson, Appl. Phys. Lett. **68**, 60 (1996).
[19]M.-E. Pistol, D. Hessman, J. Lindahl, L. Montelius, and L. Samuelson, Mater. Sci. Eng., B **42**, 82 (1996).
[20]B. Garni, J. Ma, N. Perkins, J. Liu, T. F. Kuech, and M. G. Lagally, Appl. Phys. Lett. **68**, 1380 (1996).
[21]A. Liebheit, M. Schwartzkopff, P. Radojkovic, E. Hartmann, E. Hechtl, V. Petrova-Koch, and F. Koch, J. Lumin. **72–74**, 994 (1997).
[22]Private communication from JEOL.
[23]L. L. Smith, S. W. King, R. J. Nemanich, and R. F. Davis, J. Electron. Mater. **25**, 805 (1996).
[24]G. Ramirez-Flores, H. Navarrow-Contreatas, Lastras-Martinez, R. C. Powell, and J. E. Greene, Phys. Rev. B **50**, 8433 (1994).
[25]In collaboration with S. Keller, U. K. Mishra, and S. P. DenBaars, University of California at Santa Barbara.
[26]P. Martensson and R. M. Feenstra, Phys. Rev. B **39**, 7744 (1989).

High efficiency, dual collection mode near-field scanning optical microscope

Stephan J. Stranick,[a] Lee J. Richter, and Richard R. Cavanagh
Surface and Microanalysis Science Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

(Received 26 October 1997; accepted 6 April 1998)

We have developed a near-field scanning optical microscope that provides simultaneous transmission and reflection mode measurements while concurrently recording a topograph of the sample surface. In this microscope design, an ellipsoidal cavity is used to collect and reimage the light reflected from a near-field optic. Reimaging of the near-field optic away from the mechanical components of the microscope eliminates any hindrance that they would otherwise cause. The near-field optical microscope configuration described in this article allows for the efficient, symmetric collection of reflected and transmitted light. This instrumental design has been optimized for chemical problems that require molecular characterization on the nanometer scale. Images of ~100 nm Au particles demonstrate the utility of this instrument. [S0734-211X(98)03804-9]

I. INTRODUCTION

Near-field scanning optical microscopy (NSM) is emerging as a powerful technique for obtaining optical images with spatial resolution that exceeds the diffraction limit of light. This enhanced resolution can be achieved by scanning a subwavelength aperture within a few nanometers of a specimen, such that the spatial resolution is determined by the aperture size and not by the diffractive nature of light.[1–4]

The majority of near-field optical microscope (NSOM) designs make the simultaneous collection of transmitted and reflected light inefficient, restrictive, or impossible. In the standard tip illumination, transmission mode NSOM design, light emanating from the near-field probe passes through the sample and is collected by a far-field optic.[5] Normally, this optic is in the form of a high numerical aperture (NA) microscope objective that is along the same optical axis (coaxial) as the near-field optic. In this configuration, most of the scanning architecture of the microscope would be on the opposite side (reflection side) of the transmission collection optics. This mechanical hindrance typically prevents the use of high NA, short working distance collection optics for reflection mode NSOM. Longer working distance, lower NA optics must be used to collect the reflected light. This results in relatively inefficient collection of the light; and, since the angle of collection is limited, shadowing effects can result from an asymmetric collection of light. These shadowing effects can arise from the collection optic not being coaxial with the tip[6] or from the use of a low NA optic.[7] Shadowing artifacts complicate the image and make reflection measurements difficult to interpret.[6–8] Therefore, it is highly desirable to collect the reflected light symmetrically (coaxial) about the tip. This was the approach taken by several groups that constructed microscopes which were optimized for the collection of reflected light,[7,9–11] and a recent approach that is optimized for both transmission and reflection.[12]

In an alternative approach to the use of optics for reflected light collection, Wei *et al.* have used modified Si detectors positioned in close proximity to the sample surface.[8] This clever approach eliminates many of the difficulties that are encountered with the use of low NA optics, but their scheme limits reflected light analysis to a single detection channel and would make scattering spectroscopy difficult or impossible.

The NSOM design described in this article results in an instrument that can simultaneously record topographic, optical transmission, and optical reflection images. In order to accomplish these simultaneous measurements with high efficiency, a novel optical collection system has been designed and implemented. This collection system is based on the 1 to 1 imaging properties of an ellipsoidal cavity. By placing the near-field microscope inside and at one end (foci) of the cavity, light that is reflected from the microscope's near-field optic (NFO) is collected and reimaged at the other end (foci) of the ellipse. Reimaging the NFO circumvents the mechanical hindrances caused by the microscope scanner and tip holder assemblies. With the elimination of this hindrance, high collection efficiency optics (which generally have working distances <5 mm) can be easily placed at close proximity and *coaxial* to the reimaged NFO. Another advantage of this scheme is that it allows for the use of reflective optics, e.g., Schwarzschild objectives, which are typically larger than refractive optics and consequently even more difficult to position near the NFO. The reflection collection optics do not encumber the collection of transmitted light. Thus transmission measurements can be performed simultaneously, symmetrically, and with high efficiency.

Configuring a NSOM to operate inside an almost completely enclosed cavity requires that the microscope has a compact, rigid design; has an open scanner architecture to minimize the amount of light that it obstructs; and has minimal subsidiary components, e.g., shear-force monitoring optics and electronics. In the sections that follow, we describe

[a]Electronic mail: Stephan.Stranick@nist.gov

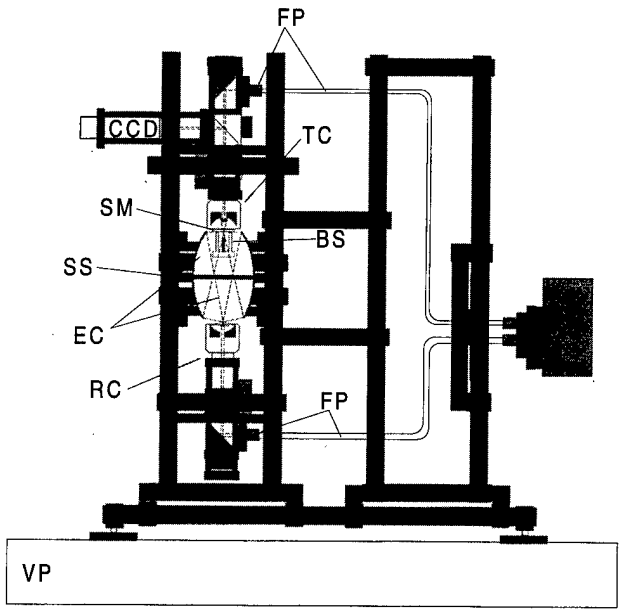


FIG. 1. Schematic of our NSOM design shows the optical ellipsoidal cavity (EC), the beetle style scanner assembly (BS), the sample (SM), the scanner spider (SS), the transmission collection optics (TC), the alignment camera (CCD), the reflection collection optics (RC), the connectorized fiber patch cords (FP), the detectors (DT), the vibration isolation platform (VP), and the four rail optical scaffolding as described in the text.

the design and integration of the components that fulfill these requirements.

II. EXPERIMENTAL APPARATUS

A. Collection optics

A schematic of the microscope is shown in Fig. 1. The collection optics are mounted on a commercially available optical rail system (Spindler and Hoyer Inc.).[13] In this system, four rods are mounted onto plates that can be adapted to support a variety of optical assemblies. An important feature of this support structure is the ease with which one can align the optical elements on a common axis. Additionally, this rail system has proved to be extremely rigid, resistant to thermal and mechanical drifts, and easily modified to meet evolving experimental requirements.

A key component of this design is the ellipsoidal cavity (EC). It provides a high collection efficiency, with a NA $= 1.0$ minus the obstruction of light by the tip holder assembly (13° half angle) and the scanner tripod (12° half angle for each leg). When determining the potential reduction of collection efficiency, only the obstruction of the light by the tripod is relevant. The near-field probe will cause a blind spot at angles approaching the surface normal.[14] As seen in Fig. 1, the ellipsoidal cavity is divided-in-two along its secondary axis. This permits the insertion of the scanner spider (SS) which supports a beetle style scanner (BS). The two halves of the cavity are each mounted on custom made, five-axis mirror mounts. These mounts are used for fine alignment of the cavity to the optical axis of the system. The scanner spider, which is identical in form to a telescope spider, positions the NFO (tip–sample junction) on the optical axis and at the top foci of the elliptical cavity. The light transmitted through the sample (SM) is collected by a far-field microscope assembly (TC), which condenses and collimates the light. A portion of the light (5%) is picked off and projected onto a CCD camera. The camera image is used in the alignment of the near-field fiber probe. Since a well-defined reference point is established by the known position of the focal plane of the far-field microscope and its associated cross hair, the positioning of the fiber probe to the top foci of the ellipsoidal cavity can be conveniently and reproducibly made. The majority of the light is transported through contrasting filters and/or coupled into a 40 μm core optical fiber patch cord (FP). The coupling of the light into an optical fiber serves two purposes: (1) it acts as a spatial filter to reduce the effects of scattered light and (2) it provides a convenient means of transporting the collected light to the detectors. The back scattered light that emanates from the near-field probe positioned at the top foci is reflected off the walls of the mirrored ellipsoidal cavity and reimaged at the bottom foci. It should be noted that the light emanating from the NFO at near-grazing angles will be reimaged and collected at near-normal angles at the second foci of the ellipsoid. This is advantageous since the majority of the light back scattered from a NFO is at near-grazing angles.[14] This light is then collected by a second far-field microscope assembly (RC). This far-field microscope is identical to the assembly for the transmission mode optics—minus the CCD camera and its associated optics. It utilizes high NA optics,[15] that condense and collimate the light which is then transported through contrasting filters and/or coupled directly into a 40 μm core optical fiber patch cord. Both the transmission collection and the reflection collection fibers are connectorized and can be easily routed to a variety of detectors, spectrographs, and analyzers.

B. Beetle style scanner

The basic design of the beetle style scanner has been described elsewhere.[16] Briefly, the scanner is comprised of four piezo tubes and a helical ramped sample holder (SH), see Fig. 2. Three of the piezoelectric tubes, type EBL No. 3, 3.18 mm o.d., 19.05 mm long, with a 0.5 mm wall,[17] form a tripod (TP) that supports and positions the sample holder. The forth piezo is placed at the center of the tripod and holds the tip of the scanned probe microscope. In our application of this scanner scheme, one end of a tripod piezo is glued onto each of the spokes of a three spoked scanner spider (SS) while a ruby ball bearing is placed at the other end and acts as a hard point for seating the sample holder ramps. The sample holder is a metal disk with three helical ramps cut into it that slope 0.75 mm over a 120° rotation. Coarse approach of the sample holder to the tip is accomplished by applying a sawtooth waveform to the tripod piezo tubes. This causes an "inertial jerking" of the piezos tangential to the sample holder resulting in the rotation of the sample holder and subsequently a change in the tip-sample separation (z direction). Similarly, a sawtooth waveform can be applied

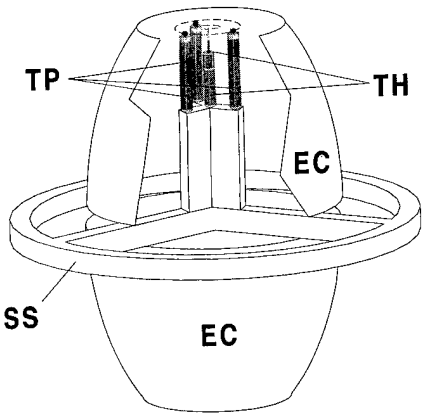


FIG. 2. Schematic of the beetle scanner embedded in an ellipsoidal cavity (EC). The scanner assembly is comprised of three tripod piezoelectric tubes (TP) that surround the tip holder assembly (TH) and shear-force piezoelectric tubes. The scanner assembly is mounted on the three arms of the scanner spider (SS). See the text for details.

that translates the sample holder laterally (xy direction). This lateral, macroscopic translation enables one to move to a particular region on the sample surface before scanning. By applying a triangle waveform to the tripod, the sample holder is reproducibly scanned in the xy plane. The tripod is also used for the ultrafine z direction translations that are made in response to feedback loop error signals. These signals are used to maintain a fixed gap between the probe and sample. This scanner provides an xy scan range of 27 μm and a z scan range of 1.7 μm.

In our application only the sample holder is moved. The tip *must* remain stationary to maintain the alignment of the NFO with the spatial filters (optical fiber patch cords). The center piezo assembly is comprised of two tubes stacked one on top of the other. The bottom tube is glued to the center of the three spoked scanner spider. It measures 9.5 mm in length but is otherwise identical to the tripod piezoelectric tubes to provide approximate compensation for thermal drift. This piezoelectric tube is used for making ultrafine optical alignments, and makes alignment to the spatial filters straight forward. The second piezoelectric tube, type EBL No. 2, 3.18 mm o.d., 3.18 mm long, with a 0.25 mm wall,[17] is used for shear-force detection which is described in the section below. It should be noted that it is important to electrically and mechanically decouple the alignment piezo tube from the shear-force piezo tube. Therefore, two separate tubes are used rather than one segmented tube. The tip holder is comprised of a precision machined pin receptacle, which has been glued inside a ceramic washer. The ceramic washer is placed on top of the shear-force piezo tube. The pin receptacle is the type used in precision integrated circuit sockets and has a beryllium copper multifinger spring contact. After opening the normally closed end of the receptacle, a 23 gauge stainless steel hypodermic tube can be inserted inside with slight force. This method of holding a tip housed inside a metal hypodermic provides excellent mechanical coupling of the tip to the shear-force piezoelectric tube. A high degree of mechanical coupling is essential for reliable operation of this shear-force scheme.

The beetle style scanner, originally designed as a scanning tunneling microscope, has proved to be highly reliable and robust. This scanner is now commonly used in all forms of scanned probe microscopes including the application found here, NSOM.[18–20] Apart from the beetle scanner's proven stability and adaptability, the scanner's wide open, compact design makes it well suited for integration into the reflection and transmission mode optical train found in this NSOM.

C. Shear-force detection

Shear-force detection is used to maintain the tip a fixed distance from the surface. While numerous optical detection schemes for shear-force sensing have been developed,[7,21,22] we use a solid-state, piezoelectric detection method that is similar to the scheme developed by Marti and co-workers.[23] This method eliminates the need for alignment of ancillary, shear-force optics, and keeps the obstruction of reflected light to a minimum. By using one half of a piezo electric crystal to drive the mechanical resonance of a near-field optical probe and the other half of the piezoelectric tube to sense the amplitude of this resonance, a change in resonance amplitude can be measured as the probe interacts with a surface. The detection of the resonance frequency change is made by a lock-in amplifier. The measurable change in resonance is sufficient to allow for nanometer and subnanometer surface variations to be detected and imaged. This solid-state, piezoelectric detection method greatly simplifies shear-force sensing. However, for this scheme to work reliably, three design criteria must be achieved: (1) good mechanical coupling must be made between the fiber probe and the shear-force piezo; (2) the shear-force piezo must be kept thermally stable; and (3) the mechanical loading of the shear-force piezo must be kept stable and to a minimum. With regards to the first criterion, we first glue the fiber probe into a 1 cm long, 23 gauge stainless steel hypodermic tube. The tip is then inserted into the tip holder assembly described above. The second criterion is significant, as thermal fluctuations can alter piezoelectric gains. This causes drift of the detected shear-force signal and can result in unstable control of the tip height. Measures have been taken to reduce thermal effects and are described in the section below. Finally, the third criterion is realized by making all electrical connections to the piezo tube with fine wires that are properly secured to provide adequate strain relief. Other factors that lead to improved performance and reliability of this shear-force scheme are the use of the higher resonance modes of the tip[24] and tuning the tip resonance to a featureless region of the shear-force piezo's response spectrum.[23] By using the higher modes of the tip resonance an order-of-magnitude improvement in shear-force sensitivity has been achieved. With optical shear-force techniques, higher modes are more difficult to use due to their nodal nature. However, this is not the case with this scheme. In fact, the higher frequency modes generally move the detected shear-force resonance out of the

lower kilohertz region where most of the piezo resonance frequencies reside. When the tip and piezo resonance frequencies coincide, an instability in the shear-force response results. The instability comes from an amplification, due to resonance, of any changes to the piezoelectric coefficients. Another way of avoiding overlap of the shear-force resonance with piezo resonance frequencies is to tune the tip resonance away from this region. This is done by adjusting the length of the fiber that extends out of the stainless steel hypodermic tube.[23,24] We have found that after initial optimization following insertion of a new tip, no additional adjustments to the shear-force scheme are needed for day-to-day operation.

D. Vibration and thermal isolation

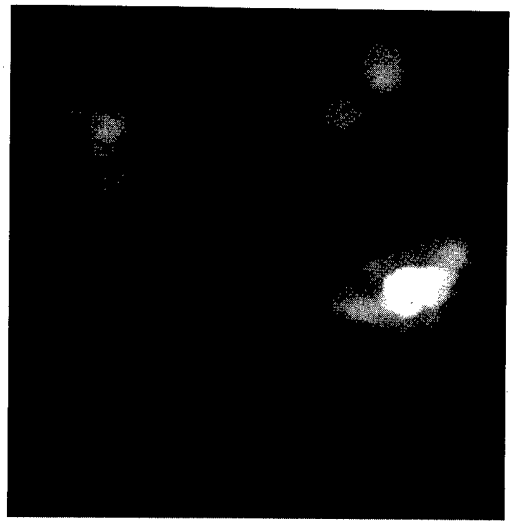
The microscope sits on a commercially available, vibration isolation workstation. This workstation is composed of a pneumatically floated, vibration damping optical breadboard that has a smaller breadboard mounted on it. The smaller breadboard is supported on rubber feet and provides an additional level of isolation. No other precautions are taken at the level of the microscope scanner to reduce or isolate vibrations. The vibration isolation workstation is housed in an acoustic isolation chamber. The acoustic isolation chamber not only eliminates the effects of environmental noise, but it also acts as a thermal barrier to help stabilize the NSOM from any thermal fluctuations in the laboratory. All instrumentation and light sources are housed separately from the microscope.

E. Power monitor

We have developed a passive coupler for monitoring the laser light coupled into the fiber pigtail of the NFO. This is accomplished by placing an amplified Si photodiode near a portion of the fiber pigtail and measuring the scattered light out of the fiber. This procedure allows for quick and convenient identification of laser intensity or fiber launch variations that could otherwise complicate an NSOM image. To measure the low levels of scattered light out of the fiber cladding, a gain of 10^{10} V/A is required. This level of amplification is provided by a simple operational amplifier circuit. The effects of stray light are minimized by placing a laser-line pass filter in front of the Si detector. We have used this monitoring method to identify variations in laser intensity and to detect launch drift. In the NSOM images shown in Figs. 3(b) and 3(c), the recorded variation of laser intensity coupled into the fiber was less than 1% over the time course of the image.

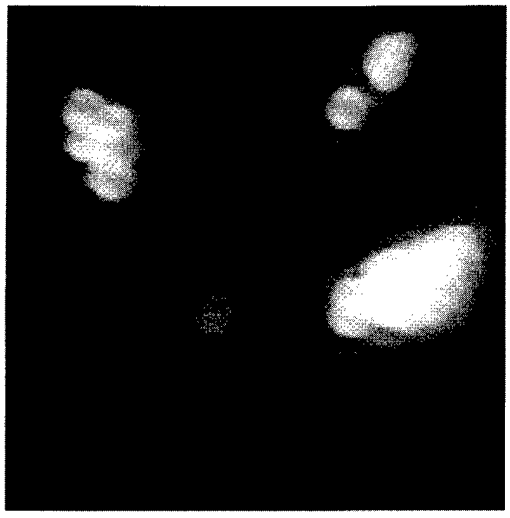
III. RESULTS

Figures 3(a)–3(c) show simultaneous recordings of shear-force topography, NSOM transmission mode, and NSOM reflection mode images. These images are of a 4.5 μm×4.5 μm region of a dilute coverage of 100 nm Au particles on a silanized glass substrate.[25] The near-field probe was a pulled, Al coated optical fiber[26] with a physical aper-



(a)

(b)



(c)

FIG. 3. Simultaneously recorded topographic (a), transmission intensity (b), and reflection intensity (c) images of a 4.5 μm×4.5 μm region. The images show the topographic and optical features of 100 nm Au particles on a silanized glass substrate. The grayscale represents a 170 nm altitude change in (a), a 58.5% intensity change in (b), and an 80.5% intensity change in (c).

ture size of no greater than 100 nm. The physical size of the tip was determined by imaging a high-aspect-ratio feature on the surface which in turn imaged the tip structure. The light source was an air cooled, Ar^+ ion laser operating at 488 nm and is fitted with an intensity stabilizer. The grayscale representation in Fig. 3(a) shows an altitude change of 170 nm from black to white. The features present in the topography image are assigned as being 100 nm Au spheres. As seen in this image, the Au spheres can be found in isolated, paired, and aggregated forms. The topographic features assigned as being the Au spheres appear larger than 100 nm in this image. This is due to the bluntness of the tapered near-field probe (nominal outer diameter of 270 nm), and the Au sphere having a comparable aspect ratio to the shear-force asperity. Vertical (z) resolution better than 1 nm has been achieved using this shear-force scheme with 3 nm asperities routinely imaged. An encouraging observation is the ability of this shear-force scheme to image isolated Au particles with no obvious perturbation of the particle's position. Particle sizes as small as 15 nm have been imaged with this system.

In both of the near-field optical images, intensity changes in the transmitted and reflected light are represented by a linear grayscale. The range of intensity variation in the transmission image is 58.5% while features in the reflection image represent an intensity change of 80.5%. Of note in these images are the intensity changes that are in near registry with the topographic position of the Au particles, ± 40 nm from the topographic center of the particles. Internal structure is apparent in both the transmission and reflection images. Also apparent in the images are intensity changes that surround the Au particles. These features have no topographic counterparts, i.e., they do not correlate to any topographic features nor are they a result of topographic artifacts.[27,28] These wavelike patterns have been tentatively assigned as interference patterns caused by coherent scattering between the particles and the near-field probe itself. The periodicity of the waves, antinode to antinode, is $\sim\lambda/2$ or 244 nm. Of importance here is the fact that these wave patterns have a relatively symmetric shape and show no signs of asymmetry due to collection artifacts. The only break in symmetry is the node that passes through the peak of the sphere position and also bisects the waves. These features are associated with and affected by the polarization state of the light emanating from the probe. The wave and nodal features are currently a topic of both experimental and theoretical works and will be the main topic of a future publication.

IV. CONCLUSIONS

In this article we have described a near-field optical microscope capable of simultaneous operation in transmission and reflection modes with high collection efficiency. To accomplish this we have employed an ellipsoidal cavity and a scanner of open architecture. The scanner is compact, rigid, and thermally compensated, all of which make it ideal for this application. The scanner assembly has been used in both shear-force and scanning tunneling microscopy modes for tip height control. In both cases a high degree of stability and reproducibility were observed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the technical assistance of Adville A. Bell and Ken K. Takagi. We would also like to thank Professor Michael J. Natan and Dr. Chris Keating of the Pennsylvania State University for providing the Au particle samples used in this study.

[1]E. H. Synge, Philos. Mag. **6**, 356 (1928).
[2]E. A. Ash and G. Nicholls, Nature (London) **237**, 510 (1972).
[3]D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984).
[4]A. Lewis, M. Isaacson, A. Harootunian, and A. Murray, Ultramicroscopy **13**, 227 (1984).
[5]E. Betzig and J. K. Trautman, Science **257**, 189 (1992).
[6]J. A. Cline and M. Isaacson, Appl. Opt. **34**, 4869 (1995).
[7]R. D. Grober, T. D. Harris, J. K. Trautmen, and E. Betzig, Rev. Sci. Instrum. **65**, 626 (1994).
[8]P. K. Wei, J. H. Hsu, W. S. Fann, and K. T. Tsai, Appl. Opt. **35**, 6727 (1996).
[9]C. L. Jahncke, M. A. Paesler, and H. D. Hallen, Appl. Phys. Lett. **67**, 2483 (1995).
[10]C. Durkan and I. V. Shvets, Ultramicroscopy **61**, 227 (1995).
[11]U. Ben-Ami *et al.*, Appl. Phys. Lett. **68**, 2337 (1996).
[12]C. L. Jahncke and H. D. Hallen, Rev. Sci. Instrum. **68**, 1759 (1996).
[13]Certain commercial equipment, instruments, or materials are identified in this article to specify adequately the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.
[14]C. Obermuller, K. Karrai, G. Kolb, and G. Abstreiter, Ultramicroscopy **61**, 171 (1995).
[15]The numerical aperture of the reflection collection system is determined by the numerical aperture of the microscope objective that is used to condense and collimate the reimaged point source (near-field optic).
[16]K. Besocke, Surf. Sci. **181**, 145 (1987).
[17]Stavely Sensors, East Hartford, CT 06108.
[18]B. Gasser, A. Menck, H. Brune, and K. Kern, Rev. Sci. Instrum. **67**, 1925 (1996).
[19]L. A. Silva, Rev. Sci. Instrum. **68**, 1300 (1997).
[20]S. J. Stranick and P. S. Weiss, Rev. Sci. Instrum. **65**, 918 (1994).
[21]R. Toledo-Crow, P. Yang, T. Chen, and M. Vaez-Iravani, Appl. Phys. Lett. **60**, 2957 (1992).
[22]E. Betzig, P. L. Finn, and J. S. Weiner, Appl. Phys. Lett. **60**, 2484 (1992).
[23]J. Barenz, O. Hollricher, and O. Marti, Rev. Sci. Instrum. **67**, 1912 (1996).
[24]B. Vohnsen, S. Bozhevolnyi, and R. Olesen, Ultramicroscopy **61**, 207 (1995).
[25]The samples were provided by Professor M. J. Natan of the Department of Chemistry, Pennsylvania State University.
[26]E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251**, 1468 (1991).
[27]B. Hecht, H. Bielefeldt, Y. Inouye, D. W. Pohl, and L. Novotny, J. Appl. Phys. **81**, 2492 (1997).
[28]S. Bozhevolnyi, J. Opt. Soc. Am. B **14**, 2254 (1997).

Probing conducting particles buried in a $Ni_x(SiO_2)_{1-x}$ composite by conducting atomic force microscopy

E. Z. Luo,[a] I. H. Wilson, J. B. Xu, and J. X. Ma[b]
Electronic Engineering Department and Materials Science & Technology Research Centre, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong

X. Yan
Physics Department, The Hong Kong University of Science and Technology, Kowloon, Hong Kong

(Received 29 September 1997; accepted 27 April 1998)

In this article, we present an experimental study on probing conducting particles buried in a $Ni_x(SiO_2)_{1-x}$ composite with x around the percolation threshold x_c by conducting atomic force microscopy (C-AFM). The buried conducting particles were "observed" via the electric current image of C-AFM at constant bias. The current from buried conducting particles originates from field assisted tunneling through the insulating layer. Examples of measuring the thickness of the insulating layer will be given. The analysis shows that it is possible to probe the buried metal particles as deep as several nanometers underneath the surface. By correlating the surface topographic and the current image, the profile of the metal-insulator interface can be measured. General issues on spatial resolution will also be discussed in this article. © *1998 American Vacuum Society.* [S0734-211X(98)03704-4]

I. INTRODUCTION

Scanning tunneling microscopy (STM) and atomic force microscopy (AFM) have been widely used in profiling the local electronic structure and morphology with high spatial resolution of various surfaces. However, due to their extreme sensitivity to the physical properties of the topmost surface, e.g., the local electronic structure (band structure, density of states, surface states, etc.) for STM and the topography for AFM, it is not easy to probe the physics interests in subsurface. For instance, it is quite difficult to probe a conducting particle buried by an insulating layer by either STM or AFM. The insulating layer prevents the use of STM and the normal AFM is chemically insensitive. In practice, "seeing" the interesting physics at subsurface, particularly at conductor-insulator interfaces with high resolution is very much desired, especially for the fields of materials science and semiconductors physics. Although there are some theoretical suggestions[1,2] on this question, so far there are few experimental attempts. Conducting AFM, based on the principle of conventional AFM, with conducting tip, has been widely used in probing the local electric properties of various materials[3–5] and the local dielectric strength of ultrathin insulating layers like the SiO_2/Si interface.[6,7] In this article we present a detailed experimental study of probing the metallic phase buried by a SiO_2 layer by conducting AFM.

Granular materials consist of nanometer size metal particles embedded in a nearly immisicible insulating medium. Electron transport in this kind of material occurs mainly via connected conducting networks (CNs), when the concentration of metal phase exceeds the percolation threshold x_c. When the AFM tip rasters the surface of a metal-insulator composite as shown in Fig. 1, there are several possibilities. For simplicity, we take the case of $x>x_c$. At the points P1 and P2 in Fig. 1, the tip contacts metal particles, which are parts of the CNs that extend to the surface, so there is current flow between the tip and the sample if a bias is applied. The electron transport should be Ohmic in this case. At the points like P3 and P4, the metal particles are deeply buried by a thick insulating layer, therefore, there is no current flow. For an intermediate case like the point P5, the metal particles are buried by a thin insulating layer, electrons can flow from the tip through the insulating layer by direct tunneling or Fowler–Nordheim emission. In the field emission range, the $I-V$ characteristic in these areas is described by the following expression:[6]

$$\ln\left(\frac{I}{V^2}\right)=\ln B-\frac{C}{V}, \tag{1}$$

where $B=(\alpha e^2/8\pi h)(m_0/m^*)[1/t^2(E)](\beta^2/\phi s^2)$ and $C=(8\pi/3)(\sqrt{2m^*e}/h)\nu(E)(s/\beta)\phi^{3/2}$, i.e., it should be linear in the $\ln(I/V^2)-1/V$ plot. In Eq. (1), I/α is the current density, s is the insulator thickness, e is the electron charge, h is Planck's constant, m_0 is the mass of a free electron, m^* is the effective mass of an electron in insulating layer, ϕ is the barrier of insulating layer, $\beta(E)$, $t(E)$ and $\nu(E)$ are correction factors for the nominal electrical field $E=V/s$, in which $t(E)$ and $\nu(E)$ account for image charge lowering the barrier height and $\beta(E)$ is the field enhancement factor arising from the nonplanar geometry of the tip.[6] Note α is the effective emission area of the electrode. In C-AFM measurement, α is determined by the tip conditions such as tip radius and loading force.[7] From Eq. (1), the buried conductor could be "observed" via the current image at constant bias voltage. The thickness of the insulating layer could be derived from the slope of the $\ln(I/V^2)-1/V$ plot provided that the barrier ϕ, and other parameters are known. Furthermore, if the param-

[a]Electronic mail: ezluo@ee.cuhk.edu.hk
[b]Present and permanent address: Physics Department, Nanjing University, Nanjing 210093, China.

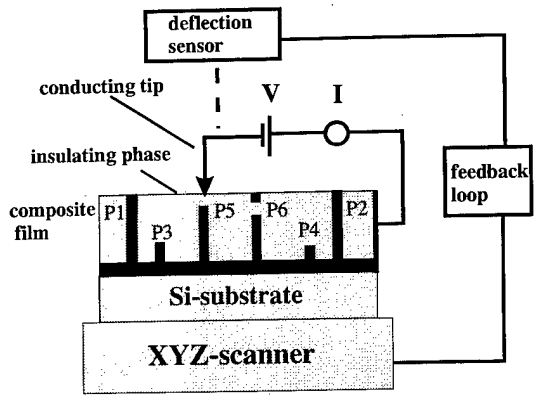


FIG. 1. Schematic diagram of the experimental setup and principle of measurement. The vertical black bars represent conducting particles: exposed to the surface (P1 and P2), buried by thin SiO_2 layer (P5) or by thick SiO_2 layer (P3 and P4), point P6 represents a metal particle exposed to surface but with a dielectric junction inside the sample. The $I-V$ characteristic of P6 should be similar to point P5, however, they can be distinguished by experiment, see the text for details. The conducting network is represented by the horizontal black bar.

eters like α, β, t, ϕ, ν in Eq. (1) are constant, then at constant bias V, to the first order approximation, the current is a measure of the thickness s underneath the tip and the current distribution is a measure of thickness distribution of the insulating layer. By correlating the topographic image, it is possible to probe the roughness of the insulator-conductor interface. In this article we employ C-AFM and based on the above concept to probe buried particles in a strongly disordered metal-insulator composite $Ni_x(SiO_2)_{1-x}$. This system was chosen because it exhibits giant Hall effect (GHE) as the x approaches x_c and the mechanism of GHE remains unclear.[8] It has been shown that the information obtained from C-AFM studies is helpful to understanding the peculiar properties in such a percolating system.[9]

II. EXPERIMENT AND RESULTS

In this article, we will focus on experimentally identifying buried conducting particles and measuring the thickness of the insulating layer above the buried conductor in the samples as described above. Experimentally, it is easy to distinguish the buried conductors from exposed conducting paths (type P1 and P2 in Fig. 1) simply by conductance measurement, since buried conductors have very low conductance. However, there are also some exposed very low conductance CNs because of the random nature of the samples we used. To exclude those low conductance but exposed conducting paths, one can measure the $I-V$ on fixed point of interest. Since Ohmic behavior dominates the conducting paths but Fowler–Nordheim (FN) emission of Eq. (1) dominates the buried conductors, we can distinguish the nature of conducting paths by measuring $I-V$. Even if $I-V$ characteristic shows FN emission behavior, one can still not be sure that those points are buried conductors since there is another possibility as shown by P6 in the Fig. 1. At point P6, a metal particle is exposed to the surface, but there is tunneling junction inside the sample. Therefore, both the conductance and $I-V$ characteristic will behave in the same way as a buried conductor P5. This could be examined by the effect of C-AFM tip induced anodization of metal particles exposed to the surface. Since the experiments were carried out in the ambient atmosphere, tip induced anodization of the metallic particles will take place. This effect is widely used in nanofabrication or modification on conducting surfaces by either STM and C-AFM.[10,11] Although nickel is highly resistant to oxidation, however, in our experience, C-AFM tip induced anodization still took place, which was evidenced by the fact that the current decreases gradually when surface was continually scanned over the same area at fixed bias. We found that the oxidation effect was stronger at higher bias (>2 V). The nickel oxide layer grew with repeated scanning. In this case, one had to confirm that the tip conditions were unchanged. In this way, one can distinguish the type P5 and P6 since the P5 is protected from anodization by the thin insulating layer. Therefore the current will not decrease with repeated scanning. Although the growth of an unknown layer on SiO_2 under C-AFM operation was reported,[6] it was not a severe problem in this study. In the data shown below all the above mentioned effects have been taken into account ensuring that we are studying the buried conductors.

The conducting AFM used in our experiments was constructed by modifying a commercial Nanoscope III AFM (Digital Instruments, Inc.) working in contact force mode with a homemade preamplifier and a bias unit with maximum voltage ±13 V. Surface topographic and current image were acquired via different data channels simultaneously. The noise of the preamplifier and the bias unit is below ±1.5 pA and 0.1 mV, respectively. The offset current of preamplifier is around 10 pA determined on pure SiO_2. The C-AFM worked at constant force and constant bias mode and the maximum current range is ±10 nA. Commercial Si_3N_4 cantilevers coated with a Cr (10 nm) plus a Gold layer of about 100 nm were used. There were no significant changes of the tip conductivity even after more than 50 images. The sample used in this study was a 1 μm thick co-sputtered $Ni_x(SiO_2)_{1-x}$ thin film on Si substrate with x around x_c, where $x_c \approx 0.53$, is the percolation threshold confirmed by either conventional resistivity measurements[8] or quantitative C-AFM measurements.[9] The details of sample preparation can be found elsewhere.[8] The C-AFM experiments were carried out in the ambient atmosphere. In the current image data channel, both real time and offline "planefit" (an option of data processing provided by Nanoscope III software) was turned off in order to avoid any artifacts induced by software processing and therefore, all the current data shown below are absolute values.

Figure 2 shows a typical current image for $x=0.532$, which is just around the percolation threshold, at a constant bias of 14.5 mV. The visible spikes in the current image correspond to conducting paths (CPs) formed inside the sample of the type P1 and P2 in Fig. 1. At constant bias, the higher the current, the more conductive the paths. As shown by the image, the typical current of the conducting paths is 0.3–1.2 nA, which gives a typical conductance of 2.0–8.3

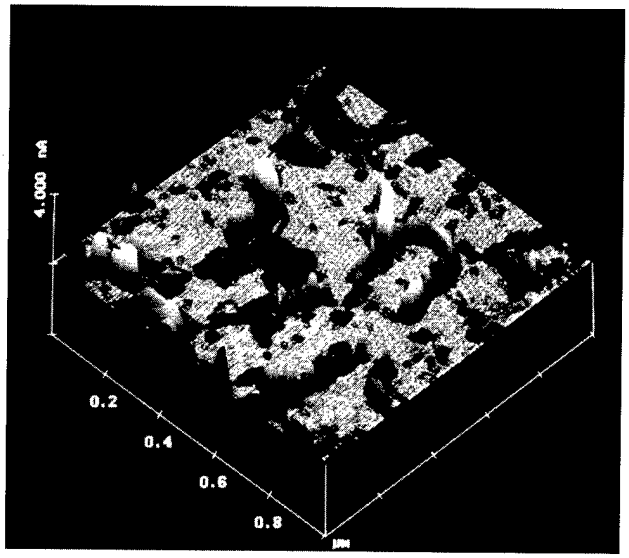
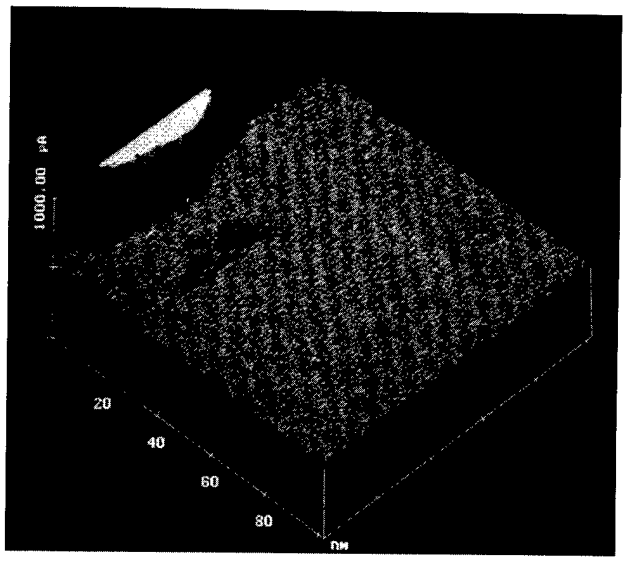
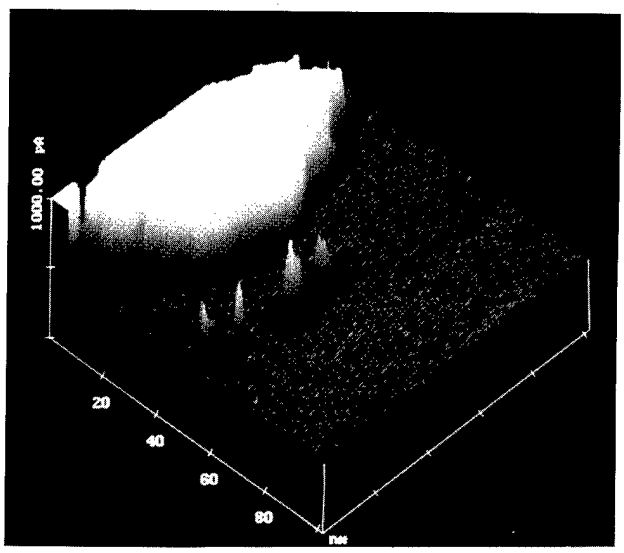


FIG. 2. Electrical current image at a bias of 14.5 mV on a $Ni_x(SiO_2)_{1-x}$ sample with $x=0.532$. The current spike corresponding to conducting paths (type P1 and P2 in Fig. 1). The scan area is 1 μm$\times$1 μm and the vertical scale is 4 nA.

$\times 10^{-8}\ \Omega^{-1}$ with some distribution. These conducting paths exhibit fairly good Ohmic behavior in terms of either single path or averaged over all the paths.[9] Moreover, it has been proved that the conductance distribution is independent on the bias applied.[9] Between these conducting paths, there are insulating areas characterized by a uniform offset current background of 10 pA. Zooming in such insulating areas and at much higher bias the insulating area becomes conductive as shown in Figs. 3(a) and 3(b), which were obtained at biases of 5.5 and 6.6 V, respectively. The high current point in Figs. 3(a) and 3(b) corresponds to the same CP, which is saturated at such high bias voltage since the current range of our setup is limited to 10 nA. The points near the conducting path, which have a typical current of about 22 and 310 pA, respectively, are interpreted as corresponding to buried metal particles based on the following considerations. First, at those points, effective conductance is about $4.0\times10^{-12}\ \Omega^{-1}$ and $47\times10^{-12}\ \Omega^{-1}$ at $V=5.5$ V and 6.6 V, respectively, which is nearly four orders of magnitude lower than that of the CPs as shown in Fig. 2. On the other hand, if those points are from "low conductance" conducting paths, they should obey Ohm's law. However, the above estimation indicates that the conductance depends strongly on the bias. In order to identify the buried metal particles more quantitatively, another suspected buried metal particle was chosen as shown in Fig. 4. The image was obtained at a bias of 7.0 V. Like the case of Fig. 3, the high current point in Fig. 4 corresponds to a CP. In order to measure the local $I-V$ plot on the suspected locations, the tip was located on the center of the Fig. 4. With our setup it is possible to measure the local $I-V$ by ramping the voltage, however, the capacity spike, which was induced by the AC coupling between tip and sample, makes the data unreliable, which was also observed by O'Shea *et al.*[6] To avoid this, we measured the local $I-V$ by manually changing the voltage and reading the



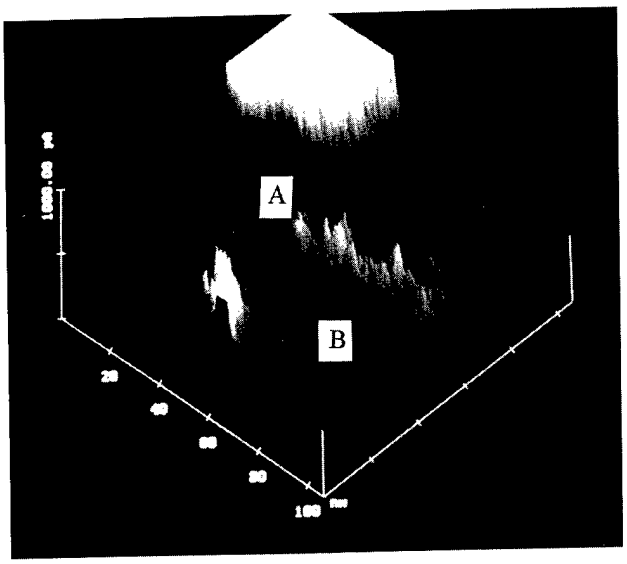
(a)



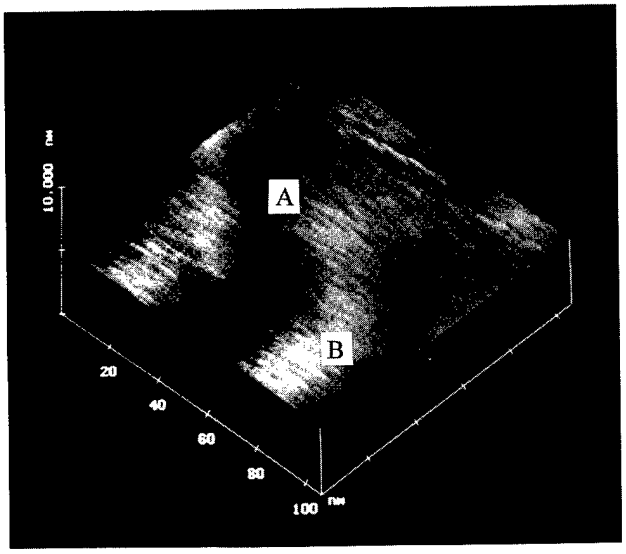
(b)

FIG. 3. Electrical current image in a nonconductive area in Fig. 2. The high current point corresponds to a conducting path, which is saturated at such a high bias. The small spikes are from buried metal particles. The imaging bias is 5.5 V (a) and 6.6 V (b). The scan area is 102 nm$\times$102 nm and the vertical scale 1 nA.

current from the current data channel. During this measurement, the feedback of AFM was turned on to maintain the tip and sample in contact. Before and after this measurement, current images at 7.0 V were acquired in the same area and no changes were observed, indicating that the measurement did not change or modify the sample. Therefore, the growth of an unknown layer or tip induced anodization, or the breakdown of SiO_2 layer, or tip damage did not occur. After subtracting the offset current the $I-V$ plot is shown by the inset in Fig. 5. Below 5 volts, the current could be hardly measured and above that, the current increases rapidly with voltage. The $\ln(I/V^2)-1/V$ plot as illustrated by Fig. 5 indicates a good agreement with theory. The thickness of the insulating layer s can be obtained by the slope of Fig. 5 if the param-



(a)



(b)

FIG. 4. Electrical current (a) and topography image (b) in one of the nonconductive areas in Fig. 2. The imaging bias is 7.0 V. The current peak in the center corresponds to buried metal particles and high current point represents conducting path. The lateral separation of A and B is 15 nm and the current is 330 and 40 pA, respectively. The height difference measured from topography image is about 0.2 nm. The scan area is 105 nm×105 nm. The vertical scale in (a) is 1 nA and in (b) is 10 nm.

eters of ϕ, $\beta\, t(E)$, $\nu(E)$ are known. As an estimation, taking $\beta=1$, $t(E)=1.013$, $m^*=0.42m_0$ in SiO_2, and $\nu(E)=0.925$,[6] the average barrier height of $Au/SiO_2/Ni$ is unknown, we take the value of $Au/SiO_2/Si$ for reference, which is 4.2–5.4 eV,[6] the least square fit gives $s=2.8$ and 1.9 nm, respectively. On the other hand, the effective emission area α of the tip is related to constant B in Eq. (1). Taking s from the above fit, one obtains $\alpha=51$ and 85 nm^2 for $\phi=5.4$ and 4.2 eV, respectively. As discussed by Ruskell *et al.*,[7] the effective area of emission α equals the effective contact area of tip and sample, i.e., $\alpha=\alpha_c=\pi r_c^2$. Here r_c is the contact radius, which is given by $r_c^3=3/4(k_1+k_2)FR_{\text{tip}}$, where F, R_{tip} are the force applied and the tip radius, respectively, k_i is related to the elastic properties of tip and sample materials via $(1-\nu_i^2)/E_i$, where ν_i and E_i $(i=1,2)$ are the Poisson's ratio and Young's moduli of gold tip and SiO_2 layer. In our experiment, the loading force was about 100–200 nN and the original Si_3N_4 tip has a radius of about 25 nm, if we assume that the tip radius increases to 50 nm after gold coating, then the contact area is about 43–68 nm^2 from the above estimation, which agrees well with the experimental results.

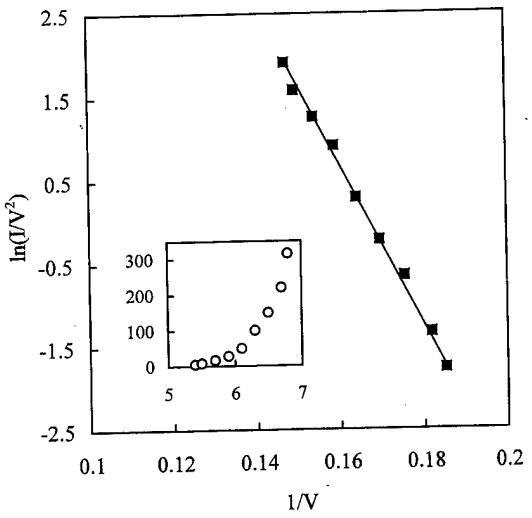


FIG. 5. $\text{Ln}(I/V^2)$ vs $1/V$, showing Nordheim–Fowler emission behavior. Inset: correspond $I-V$ plot. In those two plots, the units of voltage and current are volts and pA, respectively.

It has been shown that from the $I-V$ measurement at a fixed point it is possible to quantitatively determine the thickness of the insulating layer. On the other hand, a current image at constant bias provides information of the distribution of insulating layer. If the variation of insulating layer thickness is small, Eq. (1) can be approximated and rewritten as: $I/\alpha=B'(V^2/S_{\text{av}}^2)\exp[-C'(s/V)]$, where S_{av} is the average thickness of the insulating layer. In this approximation, the current image at constant bias is similar to STM, $I(r)\propto\exp[-s(r)/\xi]$ or the exponential approximation.[1] Therefore, to the first order of approximation, the current image is a map of thickness $s(r)$ through $\ln[I(r)]$. In Fig. 4(a) the two points A and B, which are laterally 15 nm apart, have a current 330 and 40 pA, respectively. Point A is the point that the above $I-V$ analysis applied, so that $s(A)=2.8$ nm or 1.9 nm depending on the barrier height. If all parameters in Eq. (1) are the same for points A and B (position independent), then $s(B)\approx 5.8$ or 4.0 nm from the above exponential approximation. In general, the thickness of the insulating layer is expressed by $s(r)=H(r)-h(r)$, where H and h are the profile of surface and inner interface as shown by Fig. 6. In C-AFM setup, the surface profile H could be obtained simultaneously as the current image. Combining the two images, the interface profile is given by $h(r)=H(r)-s(r)$. Measured from the topographic image as shown in Fig. 4(b), the height difference $H(A)-H(B)\approx 0.2$ nm, which gives $h(A)-h(B)\approx 3.2$ or 2.3 nm. Therefore, one is able to determine

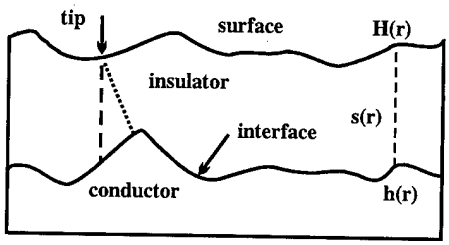


FIG. 6. Relation of surface $H(r)$, interface $h(r)$ and thickness of insulating layer $s(r)$. Because of the complexity of the electrical field, the electron tunneling path is not necessarily the path just beneath the tip (dashed line), but the most probable tunneling path (dotted line). If the two paths are not the same, then a lateral shift of the object and image will occur.

the separation of two metal particles in both lateral and vertical directions.

III. DISCUSSION

It has been demonstrated that C-AFM is a suitable method to probe buried metal particles either qualitatively (image) or quantitatively to determine the thickness of the insulating layer. Experimentally, reliable $I-V$ measurement is a crucial factor for quantitative analysis. Many factors, such as change of the effective contact area, change of tip conductance, growth of unknown layer which in fact increases the insulating layer thickness, drift of sample and break down of the oxide layer, make the measured data unreliable. For the data presented in Fig. 5, they are the average value of two measurements, i.e., voltage increasing and decreasing, the data are reproducible with very small error. Before and after the $I-V$ measurements, the current images are also almost unchanged, indicating that all these factors are negligible. Determination of thickness s with high accuracy demands detailed knowledge of all the parameters in Eq. (1), especially the barrier height. Due to the randomness of the sample in this study, the electrical field in the SiO_2 layer is more complicated than planar interface system like SiO_2/Si.[6,7] This has two main impacts on the above analysis; (1) parameters β, $t(E)$ and $\nu(E)$ in Eq. (1) could be different from those of a planar interface or even, they might be position dependent, (2) the thickness s is not necessarily the perpendicular distance between tip and metal particle since electron tunneling depends on the field distribution in SiO_2. Tunneling could occur via the most probable tunneling paths as discussed in Refs. 1 and 2 and shown by the dotted line in Fig. 6. It is straightforward that these considerations will affect the precise location of the metal particles and the spatial resolution in both vertical and lateral directions. In this respect, more systematic studies are needed. Therefore, the results of s are only for reference.

We have proposed that the electrical current image maps thickness. This can be extended to more general cases. This interpretation is only meaningful provided that (1) the parameters in Eq. (1) are position independent, (2) the contact area α_c keeps constant during scanning. The lateral resolution depends mainly on the effective contact area and also on the vertical separation of the metal particles. This dependence can be demonstrated as follows. Let's take the Fig. 3 as an example. We presume that buried particles, which correspond to small current spikes in Fig. 3, have a similar s, since the current at all points is nearly the same and has a similar behavior with bias. It is certain that between those particles there are other deeply buried particles because of randomness of the sample. There is no current at this imaging bias since the field is too small for those deeply buried ones for field emission to occur. Increasing the bias makes the deeply buried particles visible in the current image, however, the current at the shallow buried particles increases and result in a diffuse current image, which lowers the lateral resolution. Even more, break down may occur in the insulating layers on the shallow buried particles. Therefore, when the lateral resolution is considered, it should be noted that it depends on s. If two particles have nearly the same s as shown in Fig. 3, a lateral resolution smaller than 8 nm can be reached.

IV. SUMMARY

It has been shown for the first time that metal particles buried by an insulating layer can be probed via the C-AFM's current image, which originates from field assisted emission. The thickness of the insulating layer (how deep a metal particle is buried), can be quantitatively measured by analyzing the local $I-V$ characteristic according to the FN emission theory. By comparing the current at different locations, the lateral and the vertical separation among an ensemble can be obtained. In this study, a lateral resolution smaller than 8 nm has been obtained.

ACKNOWLEDGMENTS

This work is in part supported by the Research Grant Council of Hong Kong via grants CUHK501/92E and CUHK4137/97E. One of us (E.Z.L) is supported by the Post-Doctoral Fellowship of CUHK.

Presented at the AVS 44th National Symposium, 20–24 1997, San Jose, CA.

[1]M. Sumetskii and H. Baranger, Appl. Phys. Lett. **66**, 1352 (1995).
[2]M. Sumetskii, J. Phys.: Condens. Matter **2**, 2487 (1990).
[3]E. Z. Luo, J. X. Ma, J. B. Xu, I. H. Wilson, A. Pahmaov, and X. Yan, J. Phys. D **29**, 3169 (1996).
[4]C. Shafai, D. J. Thomson, M. Simard-Normandin, G. Mattiussi, and P. J. Scanlon, Appl. Phys. Lett. **64**, 342 (1994); P. De Wolf, J. Snauwaert, J. Clarysse, W. Vandervorst, and L. Hellmans, Appl. Phys. Lett. **66**, 1530 (1995).
[5]F. Houze, R. Meyer, O. Schneegans, and L. Boyer, Appl. Phys. Lett. **69**, 1975 (1996).
[6]S. J. O'Shea, R. M. Atta, M. P. Murrell, and M. E. Welland, J. Vac. Sci. Technol. B **13**, 1945 (1995).
[7]T. G. Ruskell, R. K. Workman, D. Chen, and D. Sarid, Appl. Phys. Lett. **68**, 93 (1995).
[8]A. B. Pakhomov, X. Yan, and B. Zhao, Appl. Phys. Lett. **67**, 3497 (1995).
[9]E. Z. Luo, I. H. Wilson, X. Yan, and J. B. Xu, Phys. Rev. B (in press).
[10]J. A. Dagata, J. Schneir, H. H. Harary, C. J. Evans, M. T. Postek, and J. Bennett, Appl. Phys. Lett. **56**, 2001 (1990).
[11]P. M. Campbell, E. S. Snow, and P. J. McMarr, Appl. Phys. Lett. **66**, 1388 (1995).

Microroughness of polymer thin films studied by total-reflection x-ray fluorescence and atomic force microscopy

W. L. Wu and W. E. Wallace[a)]
Polymers Division, National Institute of Standards and Technology, Building 224, Room B320, Gaithersburg, Maryland 20899

(Received 20 November 1997; accepted 22 May 1998)

The surface roughness of polymer thin films supported on nickel-coated silicon substrates was analyzed by angle-dependent total-reflection x-ray fluorescence (TXRF) in conjunction with atomic force microscopy (AFM). For highly rubbed polystyrene (PS) surfaces (500 cm rubbing length at a load of 2 g/cm^2 over a velour cloth) displaying sharp grooves and ridges, the TXRF showed no significant change while the AFM results revealed an anisotropic 6.08 nm root-mean-square roughness with an average peak-to-peak distance of 170 nm. These results were compared to the isotropic, gradually varying sinusoidal surface roughness of phase-separated polystyrene/poly(vinyl methyl ether) blend thin films. The AFM results were very similar to the results from the rubbed PS in terms of root-mean-square roughness and average peak-to-peak distance; however, the TXRF results revealed enhanced nickel fluorescence at incident angles smaller than the polymer critical angle for reflection. This discrepancy highlights some of the qualitative differences in surface topography between rubbed and phase-separated polymer thin films, and demonstrates the utility of angle-dependent TXRF to study thin film roughness and planarity. [S0734-211X(98)10304-9]

I. INTRODUCTION

Thin-film smoothness and planarity is of central importance to many aspects of modern electronics. For example, the rubbing of polymer thin-film surfaces to create microscopic grooves is commonly used to create an alignment template for liquid crystals in flat panel display applications.[1,2] Extensive work has been conducted to determine the molecular-level reorganization that results from rubbing, particularly of polyimides.[3–7] The surface topology has been studied by atomic force microscopy (AFM) and, not unexpectedly, many microscopic grooves and ridges parallel to the rubbing direction were observed.[8,9] However, since the tip of a typical AFM stylus has a radius of about 20 to 40 nm, any surface features with lateral dimensions on the order of this size will be significantly broadened. With respect to possible liquid crystal alignment mechanisms, these lateral dimensions and the sharpness of the grooves and ridges on the rubbed surface are critical to the proper performance of the display. Since the AFM-observed lateral spacing is on the order of 100–200 nm, the finite radius of the stylus tip significantly contributes to broadening of these features. For example, with a 40 nm radius AFM tip, a surface with sharp parallel ridges 80 nm apart will yield an image that appears to be a smooth sinusoidal surface even if these ridges are idealized as spikes the lateral dimension of which is infinitesimally small.

The effect of surface roughness on angle-dependent total-reflection x-ray fluorescence (TXRF) has been considered both experimentally[10,11] and theoretically[12] for the fluorescent species directly resident on the rough surface. However, the case of a thin film with a given surface roughness on top of a smooth buried layer made up of the fluorescent element has not been fully explored. This situation is shown in Fig. 1. The angle-dependent total-reflection x-ray fluorescence in this case should be a sensitive measure of smoothly varying thin-film roughness and a novel way to characterize rapidly the planarity of thin films. This could be especially useful for studying dielectric planarization over metal topography in microelectronics applications.

For a smooth thin-film surface, analyzed by means of a well-collimated incident x-ray beam, the fluorescence intensity is zero if the x-ray penetration depth is less than the film thickness. At a grazing incident angle below the critical angle for total reflection, the penetration depth is typically limited to a few nanometers depending on the x-ray index of refraction of the thin-film material. (Recall that the real part of the x-ray index of refraction is less than unity for most materials.) In this case only a negligible fluorescence signal can be measured for films thicker than the evanescent wave penetration depth, typically 10 nm. At the critical angle for reflection and above, the penetration depth goes to infinity (absent effects of x-ray absorption) leading to a steep rise in observed fluorescence. Such a sharp increase in the buried-layer fluorescence near the critical angle gets damped out if the thin-film surface is rough. This is because on a rough surface the local incident angle can vary significantly from the macroscopic incident angle. The sharp transition between the not observing fluorescence and observing fluorescence at the critical angle can disappear completely for very rough surfaces. This increase in fluorescence at grazing angles is expected to depend strongly on the lateral dimensions of the roughness of the thin-film surface. For example, on surfaces with sharp spike-like ridges the change in TXRF intensity from the substrate is expected to be minimal since only the

a)Electronic mail: william.wallace@nist.gov

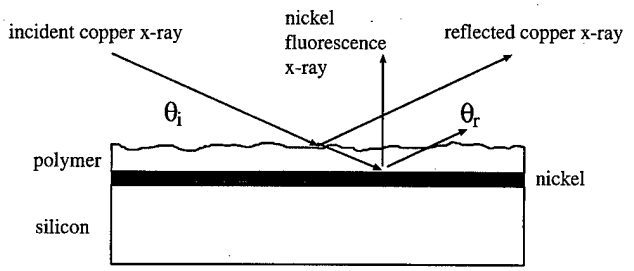


FIG. 1. Schematic diagram of the sample configuration along with the beam paths for the specular reflection of the incident beam and the fluorescence of the nickel buried layer. The incident beam angle is given as θ_i, while the specularly reflected beam angle is given as θ_r. The fluorescence is measured normal to the sample surface.

area under the bottoms of the ridges will be penetrated by the incoming x-ray beam at grazing angles less than the macroscopic critical value. For a sinusoidal surface, on the other hand, a significant fraction of the surface area will have a local incident angle that is greater than the critical angle, and therefore will permit the penetration of the incident x-ray beam. This, in turn, will give rise to fluorescence at incident angles below the macroscopic critical angle. It is the objective of this work to demonstrate that angle-dependent TXRF can provide roughness information that is quite different from that provided by AFM.

II. EXPERIMENT

Buried-layer substrates for the polymer thin films were prepared by first sputtering films of 1000 nm thick metallic nickel, which served as the fluorescent tag, onto polished silicon single-crystal wafers. The polystyrene (PS) films were made by spin coating at a spin speed of 1000 rpm using a polystyrene with a relative molecular mass of 200 000 g/mol prepared in a toluene solution having a mass fraction of 0.04 polymer. The resultant films were on the order of 500 nm thick. The rubbing of these polystyrene films was conducted at a load of 2 g/cm^2 at a speed of 1 cm/s over an acetate/rayon velour cloth. A rubbing distance up to 500 cm was applied to each sample with examination of the evolution of the surface roughness at various intermediate distances. Polymer thin films with sinusoidally varying surface roughness were prepared using a phase-separated blend of polystyrene and poly (vinyl methyl ether) (PS/PVME). The PS/PVME blend thin films were prepared in the same fashion as the single-component polystyrene films. The polystyrene for the blend work had a relative molecular mass of 220 000 g/mol and a polydispersity of 1.02. The PVME was synthesized via anionic polymerization and had a relative molecular mass of 345 000 g/mol and a polydispersity of 1.45. A molar ratio of 3/7 was used to make the PS/PVME blend. This blend is rather unstable in thin-film form,[13] consequently phase separation proceeds readily upon heating. The blends were annealed at 150 °C in vacuum for various times to create thin-film surfaces with varying degrees of sinusoidal roughness.

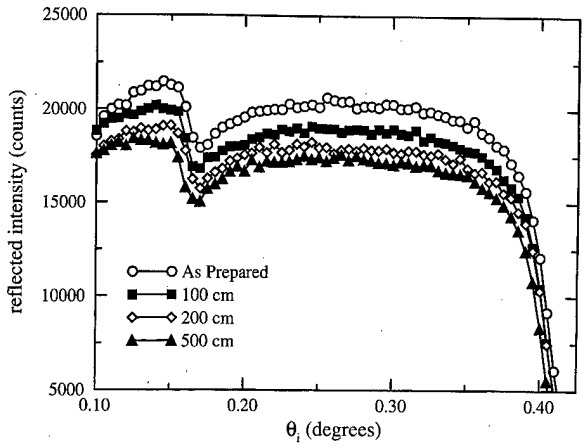


FIG. 2. Specular reflectivity from the polystyrene thin films in the as-cast condition and at three rubbing lengths. The curves are offset vertically for clarity. The minimum at 0.154° is the critical angle for polystyrene and that at 0.41° is the critical angle for the nickel buried layer.

The specular reflectivity and the angle-dependent TXRF measurements were conducted in a theta-theta configuration with a copper x-ray tube. Slits were used to define the incident and reflected x-ray beams. The incident beam had a divergence of 0.067°. No incident-beam monochromator was available on the instrument used. For this reason, the copper x-ray tube was operated at a relatively low accelerating voltage of 9 kV. The copper $K\beta$ transition, with an energy of 8.90 keV, was used to excited the nickel fluorescence. The nickel absorption edge, at 8.3 keV, lies between the copper $K\alpha$ energy (8.04 keV) and the copper $K\beta$ energy. Thus, by judicious selection of the x-ray tube voltage and the underlying fluorescence layer element the experiments could be performed without a monochromator. Addition of a monochromator to the incident beam would markedly improve the resolution of the TXRF measurements. The fluorescence was measured with a second energy-dispersive detector placed

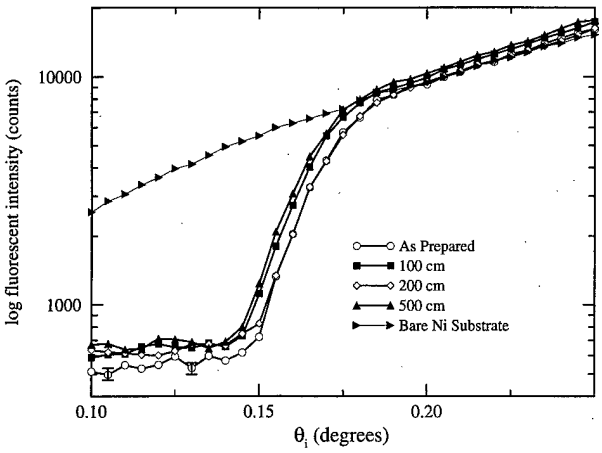


FIG. 3. Angle-dependent total-reflection x-ray fluorescence from the nickel buried layer for the same rubbed polystyrene samples as shown in Fig. 2. For comparison, fluorescence from the bare nickel layer is also shown. Note that there is no change in shape among the rubbed polystyrene curves. The uncertainty is based on Poisson counting statistics and is only significant in the low-count region where it is shown for a few typical points.

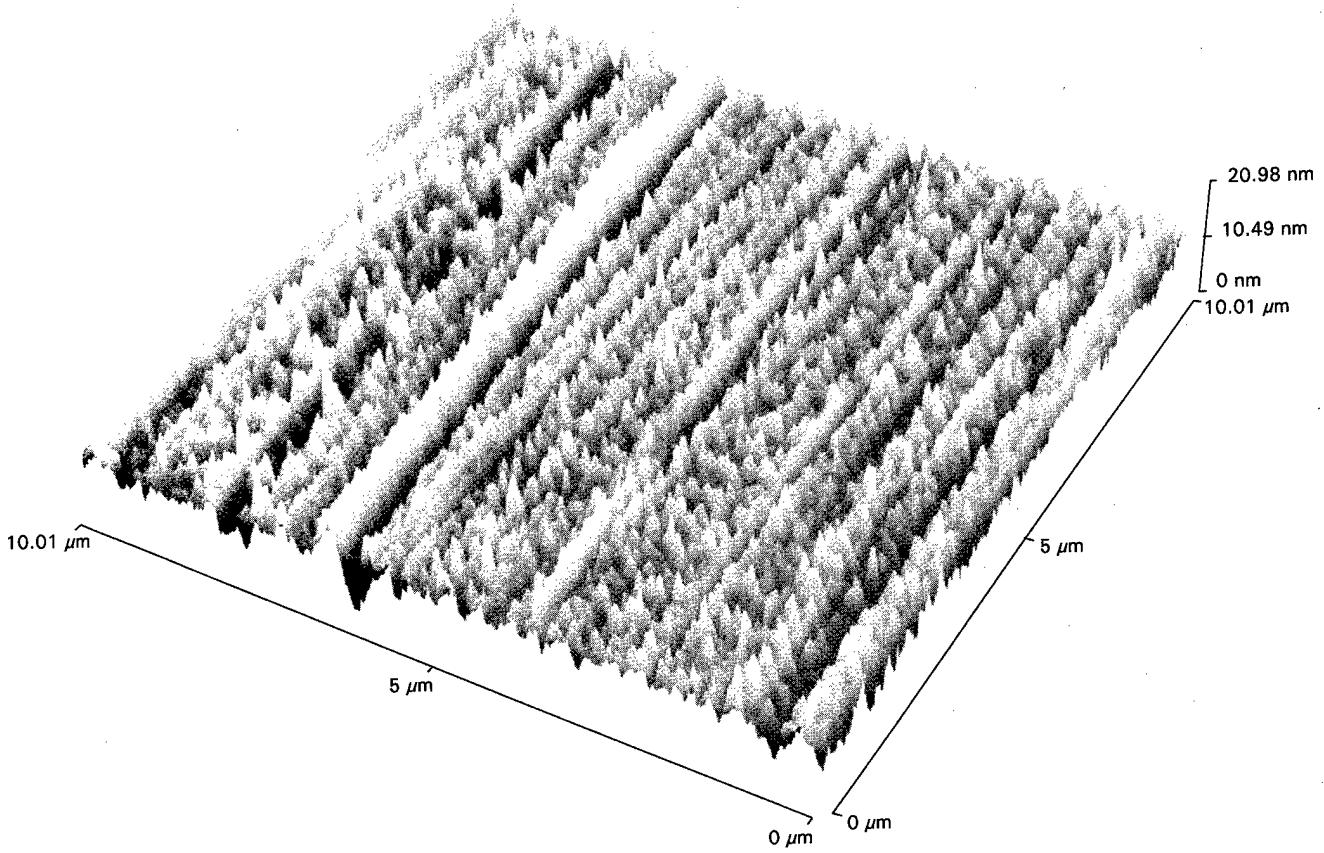


FIG. 4. Noncontact-mode atomic force microscope image of the polystyrene film after 500 cm of rubbing. The anisotropic nature of the roughness is due to the fact that the rubbing was performed in only one direction. The scanned region is 10 μm×10 μm with a vertical scale 21 nm.

directly above the illuminated area at a distance of 5 mm. The energy window was set at 7.47 keV, that of the nickel $K\alpha$ line and had a window width of 500 eV. Scatter of the incident copper $K\alpha$ into the fluorescence detector was found to be minimal.

The specularly reflected beam was monitored via an energy-dispersive detector with its energy window set at 8.04 keV, that of the copper $K\alpha$ line. The width of this window was 200 eV. The use of an energy-dispersive detector to measure the specular reflectivity means that a monochromator was not needed for these measurements either.

The atomic force microscope was operated in a noncontact mode with a scanning speed of 10 μm/s over an area of about (10×10) μm. Rescan of the same area revealed no difference in image indicating that this scanning scheme did not introduce any measurable artifacts onto the polymer thin-film surfaces.

III. RESULTS AND DISCUSSION

The specular reflectivity of the incident x-ray beam near the critical edge of the PS film is shown in Fig. 2 with the curves offset vertically for clarity. Data for samples subjected to three different rubbing distances, 100, 200, and 500 cm, are presented along with data from the unrubbed surface. The reflectivity was performed normal to the rubbing direction and did not show any discernible change upon rubbing. The uncertainty for this experiment, and the angle-dependent total-reflection x-ray fluorescence, is derived from Poisson counting statistics and is given by the square root of the total number of counts. In most cases this is much smaller than the symbols plotted in the figures. In a few of the low-count regions of the TXRF curves the uncertainty is somewhat larger and is indicated on those figures. The cusp near the grazing incident angle of 0.154° corresponds to the critical angle for reflection of the PS films; the sharp fall off at 0.41° is the critical angle for nickel. No difference was found between the reflectivity data collected with the incident beam parallel or perpendicular to the rubbing direction, at least in the angular range studied. The angle-dependent total-reflection x-ray fluorescence results were also collected with the beam perpendicular to the rubbing direction in order to maximize the deviation between the local incident angles and the macroscopic incident angle. These fluorescence results are presented in Fig. 3 together with the fluorescence data from the bare nickel surface. Again rubbing did not result in any systematic changes in the angle-dependent TXRF results. That is, all of the curves have essentially the same shape with only a very slight increase in fluorescence intensity below the critical angle for polystyrene. (The lateral shift in the curves is due to a slight sample misalignment

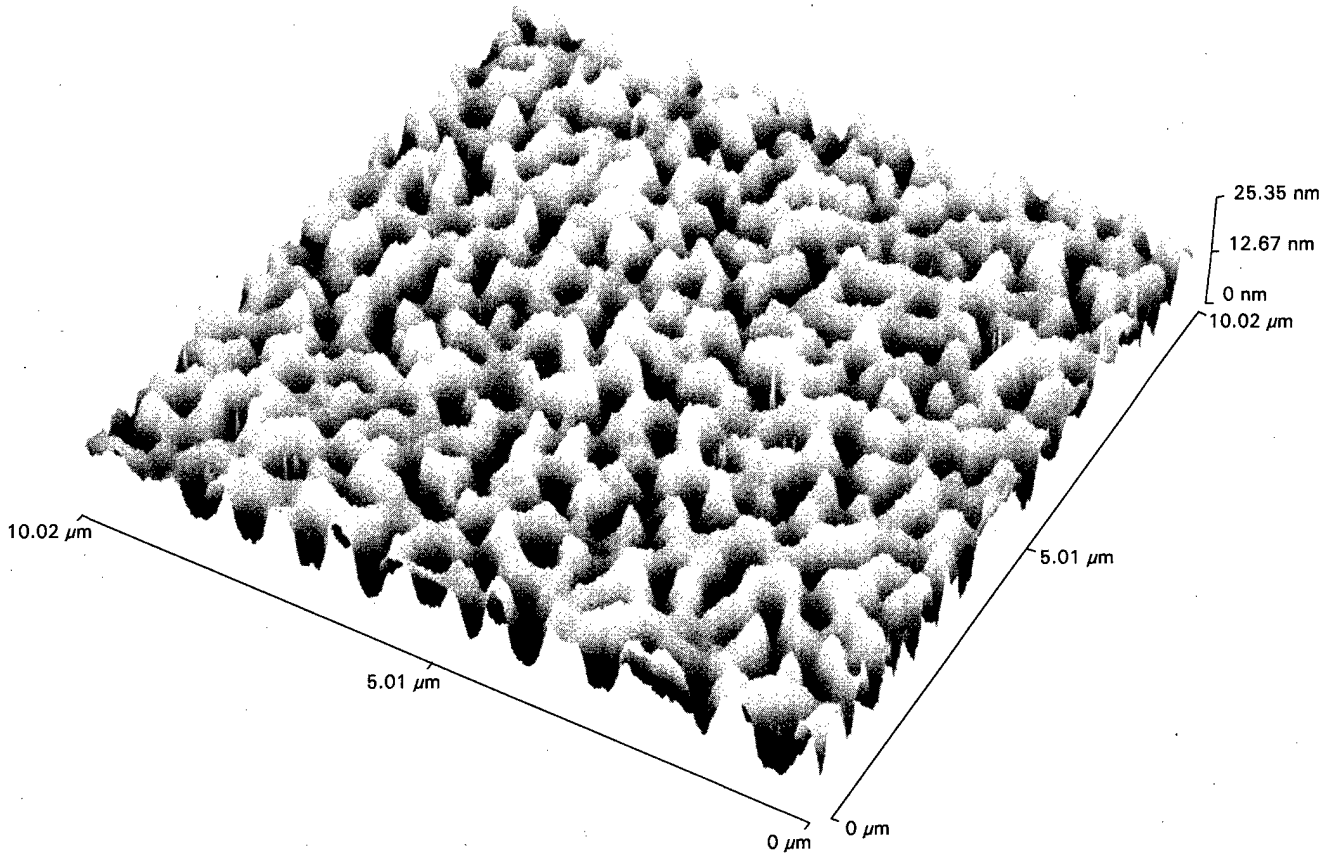


FIG. 5. Noncontact-mode atomic force microscope image of the phase-separated polystyrene/poly(vinyl methyl ether) blend film after 0.5 h annealing at 150 °C. The isotropic roughness is typical for phase-separated thin-film blends. The scanned region is 10 μm×10 μm with a vertical scale of 25.35 nm.

incurred by removing the sample from the reflectometer to perform the rubbing. This will not affect the shape of the curves.) The dramatic increase of the nickel fluorescence intensity at 0.154° coincides with the critical angle for PS. Conversely, fluorescence intensity from the bare nickel surface exhibits a gradual increase with increasing grazing incident angle. (Note that this bare nickel surface itself had a root-mean-square roughness of about 10 nm which is why its fluorescence intensity increases gradually and not abruptly.) Since the critical angle of nickel for copper $K\alpha$ x rays is 0.41°, then over the incident angles studied by TXRF, 0.1° to 0.25°, the fluorescence intensity is expected to increase only very slightly due to increasing penetration of the evanescent wave into the nickel film below the nickel critical angle. A somewhat greater increase than expected is observed due to the change in the size of the incident beam footprint which decreases as the incident angle increases. In particular, the fluorescence detector window is only 2.5 mm square but at low incident angle, for example, 0.1°, the footprint of the incident beam 29 mm and decreases to 11 mm at 0.25°. Thus, the change in incident beam footprint leads to a gradual rise in the fluorescence signal. Neither the specular reflection nor angle-dependent TXRF data was corrected for this effect.

The AFM results from the rubbed PS film with 500 cm rubbing distance is given in Fig. 4. The roughness is anisotropic since the rubbing was performed in one direction only. Line profile analysis of this surface in a direction perpendicular to rubbing yielded a root-mean-square (rms) roughness of 6.08 nm, a marked increase from 1.37 nm of the PS surface prior to rubbing. The average peak-to-peak distance was 170 nm and the average peak angle was 1.78°. For the unrubbed surface, the peak spacing was 80 nm with an average peak angle of only 0.39°. This clearly demonstrates that rubbing leads to a noticeable increase in film roughness even though the specular reflectivity and the TXRF failed to discern the difference between the unrubbed and the rubbed surface. This discrepancy between the TXRF and AFM results seems to indicate that AFM is more sensitive than TXRF in detecting this type of the roughness on a rubbed PS surface. That is, if the PS roughness consists of sharp grooves and hills, while the majority of the surface remains unaltered, the TXRF does not detect any significant changes. Conversely, for rough surfaces with gradual slope, TXRF should detect differences. To test this conjecture a surface with gradual type of roughness was examined by both AFM and TXRF.

An atomic force microscope image of the PS/PVME surface after 0.5 h annealing at 150 °C is shown in Fig. 5. As expected, annealing resulted in a pronounced increase in the surface roughness. The roughness here is isotropic as expected from random nucleation in a phase-separated thin

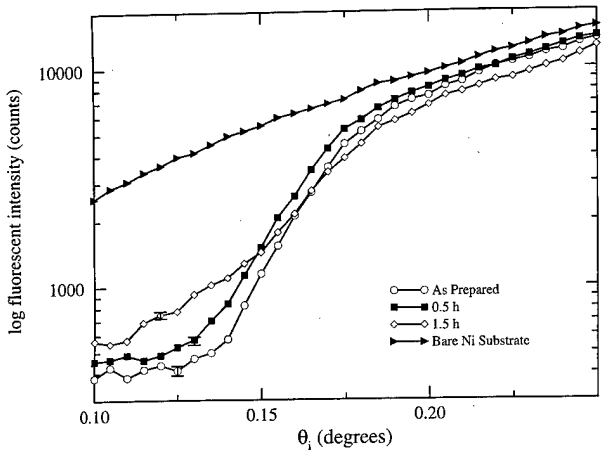


FIG. 6. Angle-dependent total-reflection x-ray fluorescence from the nickel buried layer for the phase-separated polystyrene/poly(vinyl methyl ether) blend films. For comparison, fluorescence from the bare nickel layer is also shown. Note the change in shape of the curves as a function of annealing time. The uncertainty is based on Poisson counting statistics and is only significant in the low-count region where it is shown for a few typical points.

film.[14,15] The results from line analysis indicate that the rms roughness increased from 3.39 to 6.43 nm. The average peak-to-peak distance was 290 nm with an average peak angle of 1.06°. These surface roughness parameters are not markedly different from the 500 cm rubbed PS surface except in the peak spacing: the rubbed surface has more tightly spaced roughness than the annealed PS/PVME. Further annealing to 1.5 h total time resulted in an increase of roughness to 22.9 nm.

The TXRF results of the PS/PVME surfaces are given in Fig. 6 for the as-cast sample and after annealing times of 0.5 and 1.5 h together with the data from the bare nickel surface as the fluorescence intensity reference. Annealing has a definite effect on the TXRF intensity in an angular region below the critical angle of the polymer film. The fluorescence intensity increases with annealing time and is consistent with the AFM results. At higher angles of incidence, the fluorescence intensity follows that of the bare nickel because the absorption of x rays by the thin polymer coating is negligible.

It should be noted that the electron density of PVME and PS are about the same. Consequently, the multiple phase nature of these phase-separated thin films has no effect on the reflection of the x-ray beam. The phase-separated film acts as a single phase material with a rough surface topology. It should also be noted that the TXRF result of Fig. 3 from the rubbed PS series seems to be different from that of the PS/PVME film even in the unannealed condition shown in Fig. 6. A sharp rise of the fluorescence intensity occurs near the polymer critical angle for the rubbed PS films whereas a more rounded transition occurs in the unannealed PS/PVME sample. According to the AFM results, the roughness of the unrubbed PS was 1.37 nm while that for the unannealed PVME/PS film was 3.39 nm. The high initial roughness of the PS/PVME surface is likely a result of some minor phase separation occurring during the sample preparation stage. The above observations demonstrate the sensitivity of TXRF for detecting surface roughness of the type exemplified by the PS/PVME blend. Since the surface topology of phase-separated blends is controlled by the surface tension,[16] it is reasonable to expect that the roughness is the type with smooth and gradual slopes. As to the nature of the surface roughness of rubbed PS, the fact that TXRF data from severely rubbed samples still exhibited a sharp transition near the polymer critical angle, almost identical to that of surface prior to rubbing, suggests a different type of roughness. One with sharp spikes and narrow and steep grooves is consistent with the TXRF and AFM results.

These results also bear on the interpretation of grazing-incidence x-ray diffraction as applied to probe the surface chain orientation of rubbed polyimide.[3] By comparing the diffraction patterns with the incident beam parallel and perpendicular to the rubbing direction while keeping the incident beam at a constant grazing angle below the polyimide critical angle, information regarding chain orientation was obtained. One concern of the above measuring technique is the penetration depth of the x rays at these two incident beam orientations. With the incident beam perpendicular to the rubbing direction, the local incident angle in many locations may exceed the experimental incident angle. If so, one expects that the x-ray beam will penetrate to a greater depth than expected, hence the diffraction pattern originates from a deeper surface layer than that with an incident beam parallel to the rubbing direction. If this were true the diffraction results would be ambiguous since the diffraction volume would be different in each case. Now this concern can be removed since the angle-dependent TXRF results indicated, at least for polystyrene films, that rubbing did not result in a surface topography that will significantly enhance the x-ray penetration through the films.

IV. CONCLUSIONS

Angle-dependent total-reflection x-ray fluorescence, in conjunction with atomic force microscopy, was used to study the surface roughness of polymer thin films deposited onto nickel-coated silicon wafers. Velour-rubbed polystyrene produced a type of roughness characterized by sharp parallel ridges on an otherwise unadulterated surface which was detectable by AFM but was not readily apparent in the TXRF spectra. However, for phase-separated polystyrene/poly (vinyl methyl ether) thin films, the roughness of which consisted of continuously varying sinusoidal hills and valleys and for which the calculated AFM roughness was very similar to that for the rubbed polystyrene, the TXRF results were very different. These results demonstrate how angle-dependent TXRF may be useful as a probe of thin film roughness and, for example, may be useful in measuring the ability of dielectric films to planarize underlying metal topography.

[1]S. W. Depp and W. E. Howard, Sci. Am. **268**(3), 90 (1993).
[2]J. S. Im and A. Chiang, MRS Bull. **21**(3), 27 (1996).

[3]M. F. Toney, T. P. Russell, J. A. Logan, H. Kikuchi, J. M. Sands, and S. K. Kumar, Nature (London) **374**, 709 (1995).
[4]M. G. Samant, J. Stohr, H. R. Brown, T. P. Russell, J. M. Sands, and S. K. Kumar, Macromolecules **29**, 8334 (1996).
[5]N. A. J. M. van Aerle and A. J. W. Tol, Macromolecules **27**, 6520 (1994).
[6]Y. Ouchi, I. Mori, M. Sei, E. Ito, T. Araki, H. Ishii, K. Seki, and K. Kondo, Physica B **208/209**, 407 (1995).
[7]H. Kikuchi, J. A. Logan, and D. Y. Yoon, J. Appl. Phys. **79**, 6811 (1996).
[8]Y. M. Zhu, L. Wang, Z. H. Lu, Y. Wei, X. X. Chen, and J. H. Tanh, Appl. Phys. Lett. **65**, 49 (1994).
[9]J. Y. Huang, J. S. Li, Y. S. Juang, and S. H. Chen, Jpn. J. Appl. Phys., Part 1 **34**, 3163 (1995).
[10]R. Klockenkamper, *Total-reflection X-ray Fluorescence Analysis* (Wiley, New York, 1997).
[11]K. Tsuji, T. Yamada, T. Utaka, and K. Hirokawa, J. Appl. Phys. **78**, 969 (1995).
[12]D. K. G. de Boer, Phys. Rev. B **53**, 6048 (1996).
[13]Q. Pan and R. J. Composto, in *Phase Separation Studies of Confined Thin Film Polymer Blends*, edited by J. M. Drake, S. M. Troian, J. Klafter, and R. Kopelman (Materials Research Society, Pittsburgh, 1995), Vol. 366, p. 27.
[14]F. Bruder and R. Brenn, Phys. Rev. Lett. **69**, 624 (1992).
[15]L. Sung, A. Karim, J. F. Douglas, and C. C. Han, Phys. Rev. Lett. **76**, 4368 (1996).
[16]S. Puri and K. Binder, Phys. Rev. E **49**, 5359 (1994).

Plasticity study of deformed materials by *in situ* atomic force microscopy*

C. Coupeau,[a] J. C. Girard, and J. Grilhé
Laboratoire de Métallurgie–Physique, Université de Poitiers, BP 179—86960 FUTUROSCOPE Cedex, France

(Received 25 September 1997; accepted 22 May 1998)

The emerging of a single dislocation at a surface creates a step with a height equal to the Burgers vector component normal to the surface. By coupling an atomic force microscope with a tensile tester, the fine slip line structure can be analyzed. This equipment is particularly suitable to follow the course of plastic flow from the emergence of the first few dislocations from bulk crystal to the stages of work hardening. The motivations, instrumentation and results on LiF single crystal are described. © *1998 American Vacuum Society.* [S0734-211X(98)09504-3]

I. INTRODUCTION

Observation of the surfaces of materials after or during deformation has always been an important and interesting way to study plasticity. The discovery of slip lines on deformed single crystals in the 1930s,[1] is thus at the origin of plastic theories based on slip systems and the dislocation concept.[2–4] Transmission electron microscopy (TEM) is used extensively to determine both dislocation Burgers vectors and slip planes. Nevertheless, surface observation is always used since it remains the simplest method to determine the active slip planes in a given material.

Among the recent techniques to analyze the fine slip line structure one can note the kinetic observation method by optical microscope proposed by Nëuhauser and co-workers[5–7] and more recently the development of tensile testing inside a scanning electron microscope.[8–10] However, these techniques do not have sufficient resolution to follow the emergence of single dislocations characterized by formation at the surface of low-height steps.

This kind of experiment may be successfully carried out taking advantage of the atomic-scale resolution offered by the scanning probe microscopy.[11–14] Moreover, it was clear that the development of an experimental system allowing *in situ* surface observations during deformation would be a real advantage for plastic deformation studies. Few attempts have thus been undertaken with bending measurements performed both in scanning tunneling microscopy and atomic force microscopy (AFM).[15,16]

In this article an apparatus is described which consists of a tensile tester coupled to an atomic force microscope. Details of stress/strain measurements are given obtained on LiF single crystals.

II. EXPERIMENTAL APPARATUS

The apparatus developed is shown in Fig. 1. The system operates in air at room temperature and is composed of a stand alone AFM (SAAFM) connected to a tensile tester.

The choice of the compression test depends on the Schmid factor of the primary slip system which increases to the detriment of nonactivated systems,[17] so that the compression test is a stable configuration. Moreover, specific specimens with well-defined characteristics are not needed and this is of distinct interest. The specimen is deformed by two piezoelectric translators (PZT) in a symmetrical configuration (No. 5 in Fig. 1), which introduces minimal mechanical vibrations into the system and limits the sample shift under the scanning tip. Each translator is able to extend up to 60 μm in length by minimum increments of 6 nm and can support 30 kN in force.

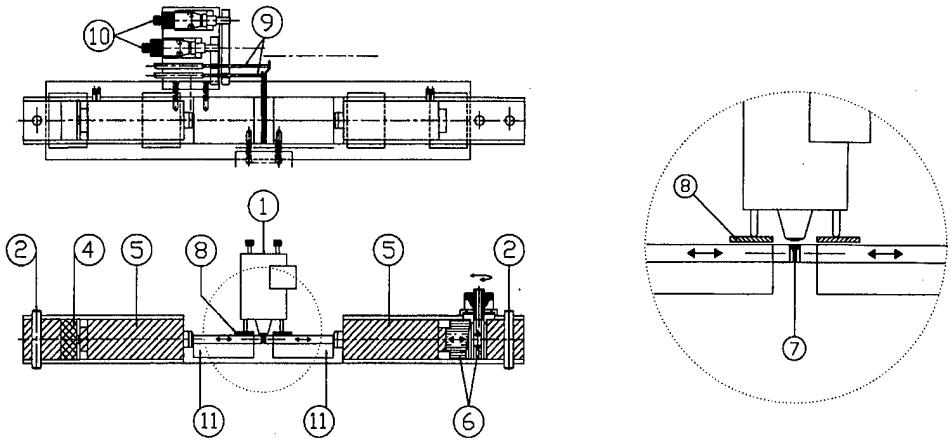
Two linear slide bars (No. 11) have been installed to insure perfect alignment of the applied stress. This is determined by a load cell (No. 4) included behind one of the two PZT at the extremity of the system. It is assumed there is no friction between the parts. All the parts of the device are held rigidly in an Invar tube (near zero thermal expansion coefficient). This apparatus is supported by a pneumatic isolation system to reduce as much as possible transmission of external vibrations. A LABVIEW program is used with an additional computer to control the compression machine entirely.

The expansion of a PZT is not proportional to the electric field strength. At a given voltage the expansion of the translator depends on whether it has been operated previously at higher or lower voltage. Figure 2 shows the theoretical elongation (or applied ramp voltage) and the real expansion of the PZT. In a voltage/expansion graph this nonlinear behavior is characterized by a hysteresis curve.

Furthermore, if the operating voltage on a PZT is changed, a slow drift can be observed after the actual expansion in the same direction. This drift is closely associated with the hysteresis behavior and is caused by a follow-up polarization of the ceramic. It is very small and decreases logarithmically in time so that a certain percentage of the path is moved every time cycle. The drift is about 1% of the distance traveled per time decade and is only noticeable in practice in the first few seconds after the positioning process.

A piezotranslator, just like any other solid, is an elastic body with a given stiffness. If a force acts on it, then it will be compressed by a certain amount ΔL_R. In this case, the force acting on the translator is dependent on the path; the positioning element is pressed against a spring C_S, which is the sample one. The more rigid the constraint is, i.e., the greater the spring constant, the smaller the path covered by the PZT operating at maximum field strength (Fig. 3). A part

*No proof corrections received from author prior to publication.
[a]Electronic mail: Christophe.Coupeau@Imp.univ-poitiers.fr



(1) Stand Alone AFM
(2) End walls
(4) Load cell
(5) Piezo-translators
(6) Rough approach system
(7) Sample
(8) Translation stage
(9) Optical fibers
(10) Micrometer screws
(11) Linear slide bars

FIG. 1. Experimental apparatus cross sections.

of the expansion generated by the piezoeffect is lost in the compression due to the elasticity of the element.

All these factors show that it is not possible to determine accurately the strain from the voltage applied on the translators.

Consequently, an additional system has been included (Nos. 9 and 10 in Fig. 1). This system consists of two groups of optical fibers bundled together. This allows the determination of the movement of the sample edges at the contact surfaces with each piezotranslator. The working principle is based on the ratio intensity between the light emitted by one group of optical fibers and that reflected from the probe surface and received by the other fiber group [Fig. 4(a)]. The voltage-distance curve is presented in Fig. 4(b). Two linear parts are clearly observed, before and after the optical peak. In spite of a small displacement range, the first linear portion offers high-accuracy resolution. For this reason, two mirror parallel bars lying perpendicularly to the sample holder have been installed.

A compression test has been carried out on a LiF sample. In Fig. 5 is reported the position evolution of each contact sample surface. Two domains are clearly distinguishable. The first one, zone A from $t=0$ to 480 s, corresponds to the elastic deformation of the sample as can be seen on the stress-time curve presented in Fig. 6. The piezotranslators have difficulty compressing the sample because of its stiffness. The asymmetry observed during loading can be explained by the fact that the translators are not identical and have, consequently, different internal characteristics. After the critical shear stress is reached (zone B from $t=480$ to $t=1050$ s), the sample stiffness becomes negligible compared to the translators. All the expansion generated by the PZT is transmitted entirely to the specimen. The same behavior is observed during the down-loading movement.

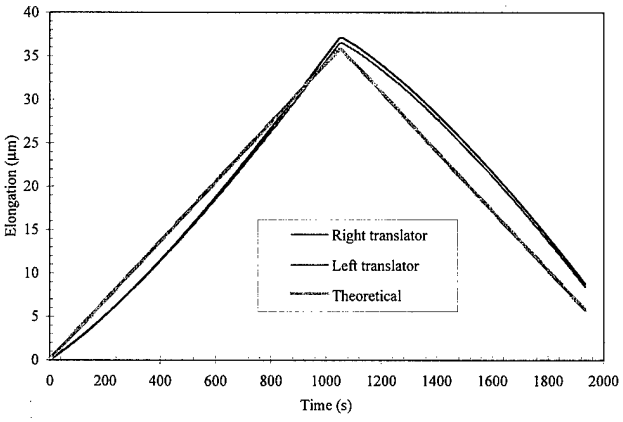


FIG. 2. Expansion-time curve of piezotranslator showing the hysteresis phenomenon.

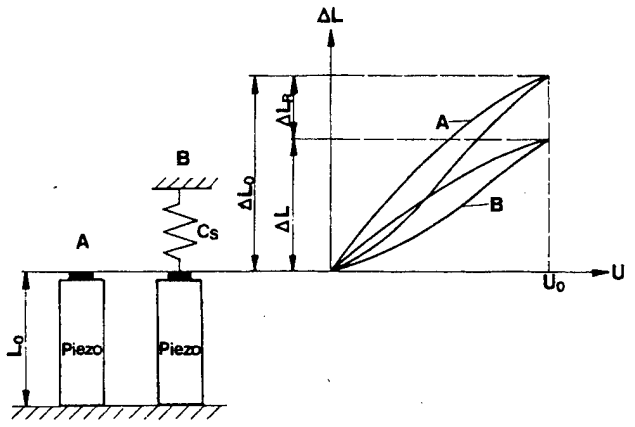


FIG. 3. Reduction of PZT expansion in the case of load of the piezo with an external spring.

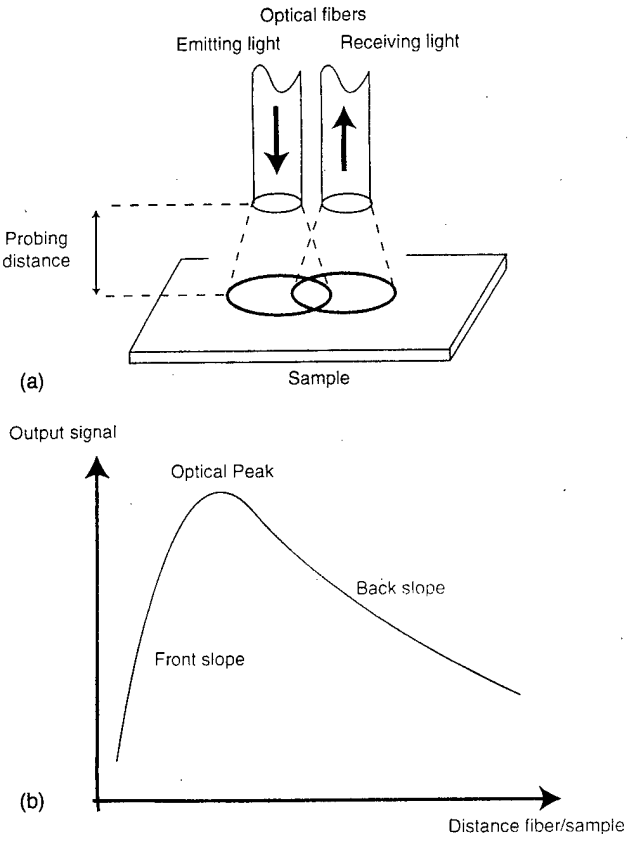


FIG. 4. Displacement measurement system: (a) working principle and (b) voltage vs distance curve.

It is concluded that the nonlinearity of the deformation rate depends on the specimen elastic or plastic deformation domain. A quasilinear deformation rate is observed beyond the critical shear stress and the apparatus permits the determination at any time of both sample strain and applied stress with a high degree of resolution. The channel sampling is at 50 kHz so that it provides a resolution of ± 5 nm for the contact-surface displacement and ± 10 or ± 350 g for the force depending on the load cell used in the machine, respectively. The specimen dimensions are nominally $2.5 \times 2.5 \times 5\ \text{mm}^3$, which provides a minimum deformation rate of approximately $10^{-6}\ \text{s}^{-1}$.

The SAAFM works in contact mode so that atoms on both the tip and sample repulse one another. Contrary to a classical AFM, the SAAFM relies on an optical interference method to detect the deflection of the cantilever. The interference system is mounted directly at the end of the piezotube so that adaptation of the deformation machine to the lower part of the microscope is possible. The usual motorized stage used to bring the sample into contact with the scanning tip has been replaced in the SAAFM by a manual adjustment system made up of three finely threaded screws whose rounded ends rest on the work surface (table or specimen). The maximum observation scan size of this microscope is $18 \times 18\ \mu\text{m}^2$ with a vertical displacement of 4.3 μm. The resolution of the SAAFM is comparable to that of a standard AFM, i.e., atomic resolution for specific materials freshly cleaved and a lateral resolution between two experimental image points of approximately 2–3 nm due to con-

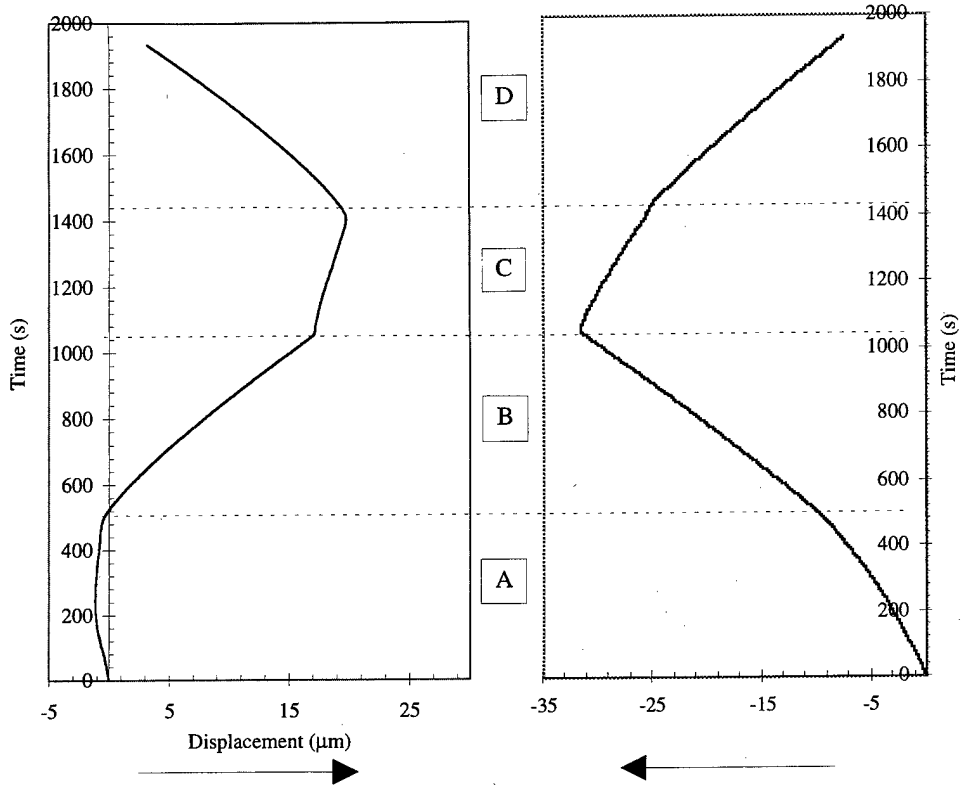


FIG. 5. Evolution of left and right sample contact surface during deformation process. A and B represent the elastic and plastic deformation domains of the sample, respectively, and C and D represent the unload movement.

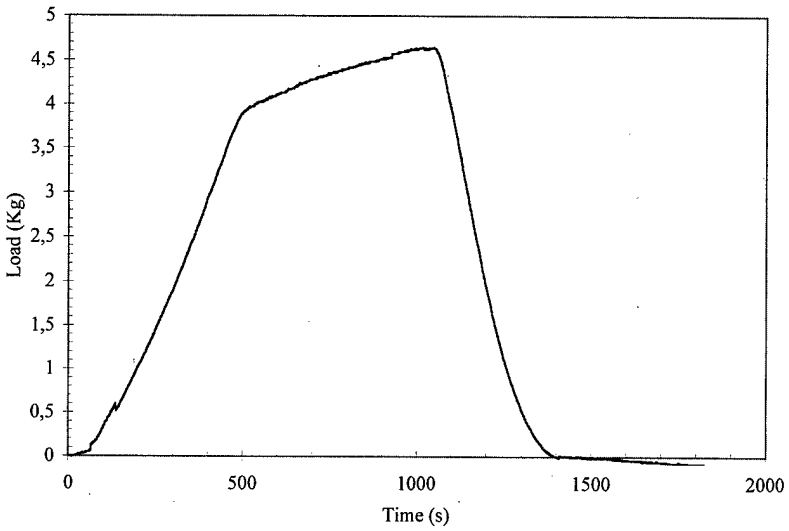


FIG. 6. Experimental load-time curve.

tamination phenomena and a fraction of angstroms vertically. However, manual engagement results in enhanced risks of tip breaking, difficult control of the tip–surface applied force and mechanical instabilities problems. The SAAFM is held on an $X-Y$ micrometer stage allowing translation of the microscope above the sample surface (Nos. 7 and 8 in Fig. 1). Microfabricated silicon nitride cantileversare used with a spring constant of about 0.1 N/m, an opening angle of 35° and a 10–15 nm curvature radius (Digital Instruments).

During a compression test the scanning tip is always at the surface contact and the sample is scanned continuously, in spite of additional noise created by the movement of the sample under the microscope. Such an image, on which one scan line corresponds to one strain value assuming the deformation and scan rates are synchronous, is of great interest. However, to study plasticity and analyze the fine structure of slip lines the compression machine is stopped just prior to scanning the surface. An entire image corresponds to one specific strain value so that cinephotography can be undertaken. Slip line statistical analysis can be used to determine the relevant experimental parameters.

III. EXPERIMENTAL RESULTS ON LIF

LiF single crystals have been studied because of their low chemical reactivity and the possibility of obtaining smooth surfaces by cleavage techniques. Compression specimens were prepared by cutting, polishing and finally cleaving the material just prior to scanning the surface. The samples were cut with the compression axis at 0° from the cube axis [001] and observation faces parallel to the (100) planes [Fig. 7(a)]. The primary slip system in LiF is the $\{110\}\langle 110\rangle$ with a Burgers vector of $a/2[110]$, where a represents the lattice parameter. By calculating the Schmid factor one, therefore, expects slip lines to lie in the [010] direction at 90° from the compression axis and a step height of an emerging single dislocation equal to its surface perpendicular Burgers vector component,

$$h_e = b_\perp = b\cos\frac{\pi}{4} = \frac{a}{2} = 0.201 \text{ nm}.$$

Because of the activation of two symmetrical slip planes, a structure of increasing and decreasing steps is expected [Fig. 7(b)].

AFM images of the slip line pattern have been performed at room temperature and pressure. The images shown in the following are taken from top to bottom unfiltered in trace direction, but no significant effects have been noticed in relation to the scanning direction. The compression axis is at $7° \pm 1°$ from the horizontal base of each image.

In Fig. 8 is reported the experimental strain versus time curve. As discussed previously, one can see that the deformation rate is almost constant when the critical shear stress is reached and is approximately 2.47×10^{-5} s^{-1} (dashed line).

The AFM image shown in Fig. 9 represents the first stage

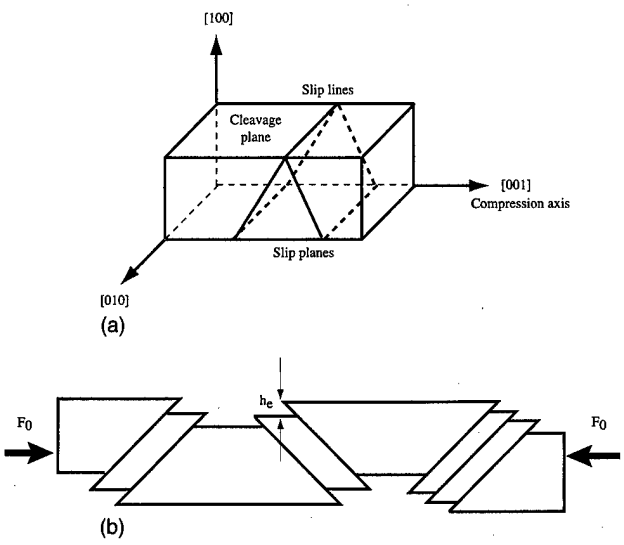


FIG. 7. Sample configuration showing (a) crystallographic orientations and expected slip lines and (b) expected step structure and unit step height equal to the Burgers vector component perpendicular to the surface.

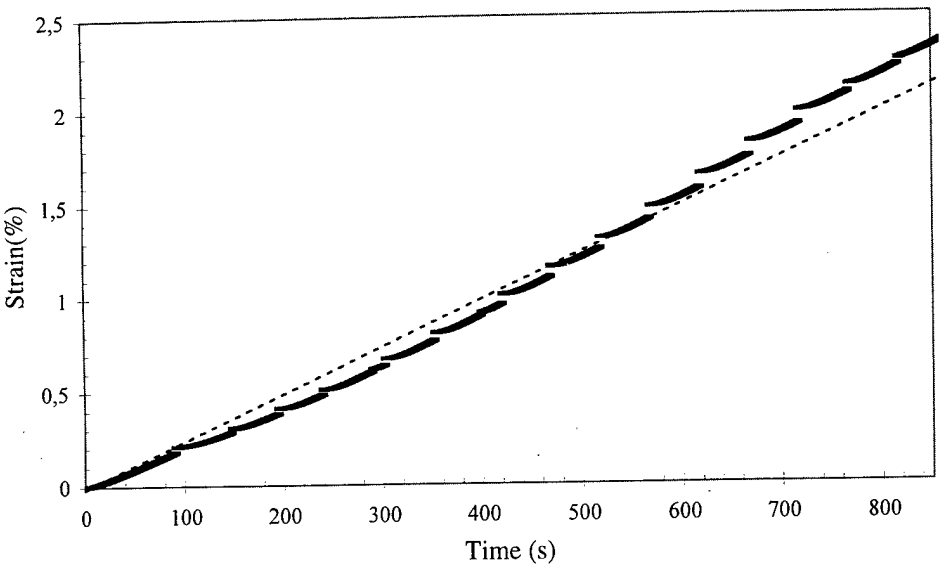


FIG. 8. Experimental deformation rate. The discontinuities (slight increases of strain) observed on the curve are the result of the slight drift of the piezoactuators after each stop of the deformation device.

of a plastic deformation process characterized by the emergence of the first few dislocations. The measured strain is 0.17% and the scan size of the image is about $15\times15\ \mu m^2$ with the grey scale bar extending a few nanometers. Rivers of cleavage steps are clearly distinguishable running from the upper left to the lower right corners. This pattern is the result of crack boundary intersections with Frank network dislocations. As expected, the surface exhibits slip lines running in preferred crystallographic direction [010]. A higher resolution scan has been carried out. A profile plot obtained by averaging several consecutive sections perpendicular to the slip lines is represented in the top left corner of Fig. 9 and shows that this slip line consists of two individual steps corresponding to the emergence of, respectively, six and two dislocations. This demonstrates the potential use of SAAFM to analyze fine slip line structures on a nanometer scale.

Figure 10 shows both the experimental stress–strain curve and three-dimensional (3D) views of LiF surface evolution during deformation. The critical shear stress is about 10 MPa. The irregularities observed on the curve are accounted for by both the drift of the piezotranslator as described previously and the specimen stress relaxation due to the dislocation network reorganization. AFM image scan size is approximately $16\times16\ \mu m^2$ and the height extension is about a few tenths of a nanometer. The grey arrows are used to position the slip line structures of each image from each to the

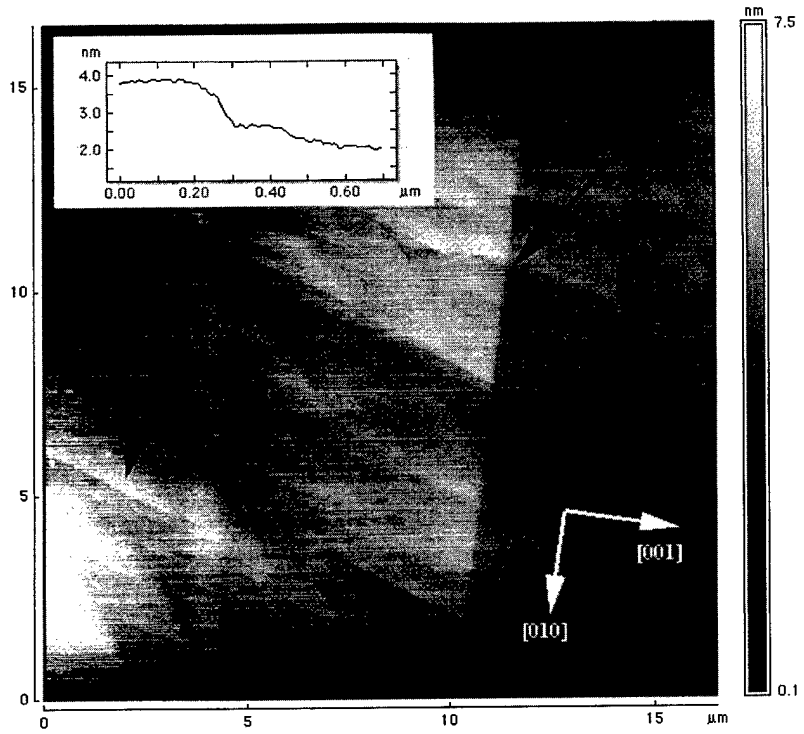


FIG. 9. Appearance of first slip lines at 83.2° from the horizontal line, which corresponds to the crystallographic direction [010]. Measured strain is 0.17%. Cross section of the fine slip line structure is observed in the top left of the image.

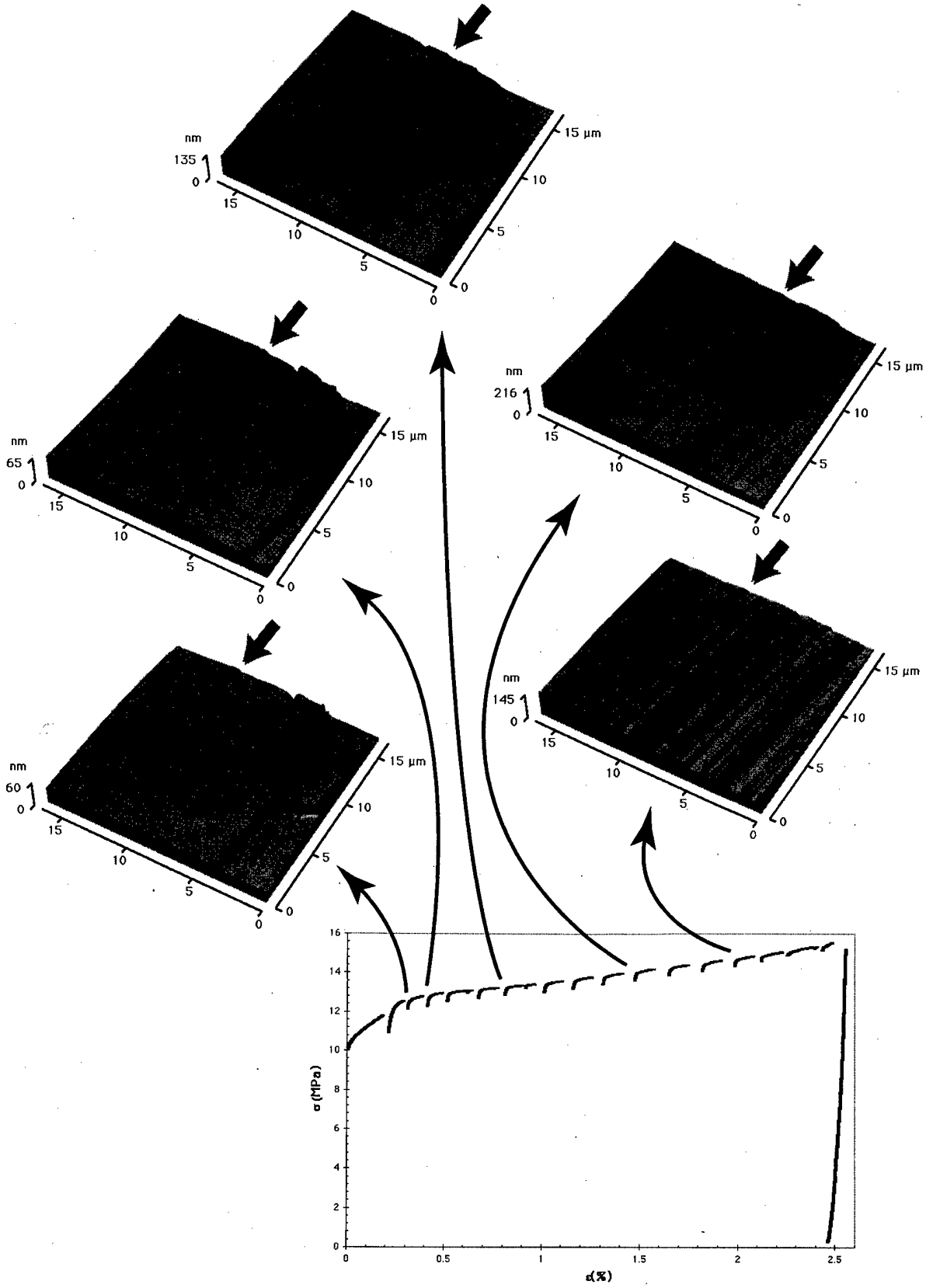


FIG. 10. Evolution of the LiF surface during deformation. Scan size of the each image is approximately $16 \times 16\ \mu m^2$ with a grey bar fully extended of few tenths of a nanometer. Each 3D view image is linked to the compression curve.

other. Furthermore, each image is linked to the compression curve (black arrows) so that the emergence of dislocations and the evolution of the slip line structure is observed *in situ* during deformation.

IV. CONCLUSION

An experimental apparatus is presented which consists of a special atomic force microscope connected to a microcompression machine. This apparatus is particularly suited to analyze slip line structures at a nanometer scale taking advantage of the atomic resolution offered by this class of scanning probe microscopy and allows observation *in situ* of the evolution of material surfaces during deformation from the emergence of the first few dislocations from bulk crystal through the stage of work hardening to secondary slip system activation, so that cinemovies can be realized.

The stress and strain of the sample are determined as a function of time with high-accuracy resolution. AFM images are linked to the compression curve, and hence, a statistical analysis of slip lines may be considered to compare the behavior of various materials and to characterize plastic surface deformation. These experiments permit determination of relevant experimental parameters during deformation, such as the average number of dislocations per slip line or average terrace width.

This method for characterizing plastic deformation *in situ* by using AFM provides information complementary to that obtained by transmission electronic microscopy. Indeed, TEM analysis is principally realized on single dislocations or isolated pileup and is very difficult to perform on work-hardened areas. AFM is particularly interesting for analyzing the fine structure of slip lines, i.e., the distance between the slip planes of dislocations relates precisely to work-hardened areas. Another factor is that TEM allows observations of dislocations remaining in the crystal whereas slip line analysis by AFM allows determination of dislocations which have emerged during the deformation process.

ACKNOWLEDGMENTS

This work was partially financial supported by the DRET agency and the Région Poitou-Charentes. The authors would like to thank L. Cartz for fruitful discussions and useful suggestions.

[1]E. Schmid and W. Boas, *Kristallplastizitat* (Springer, Berlin, 1936).
[2]J. Friedel, *Dislocations* (Pergamon, Oxford, 1964).
[3]J. P. Hirth and J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968).
[4]D. Hull, *Introduction to Dislocations* (Pergamon, Oxford, 1975).
[5]H. Neühauser, R. Rodloff, H. Flor, and Ch. Schwink, Proceedings of the 4th International Conference on the Strength of Metals and Alloys, Nancy (1976).
[6]Ch. Schwink and H. Neühauser, Phys. Status Solidi **6**, 679 (1964).
[7]H. Neühauser and R. Rodloff, Acta Metall. **22**, 375 (1974); Sov. Phys. Solid State **49**, 445 (1978).
[8]M. Ignat, M. Ducarroir, M. Lelogeais, and J. Garden, Thin Solid Films **220**, 271 (1992).
[9]M. Ignat, P. Scafidi, E. Duloisy, and J. Dijon, Mater. Res. Soc. Symp. Proc. **338**, (1994).
[10]M. Ignat and P. Scafidi, The Japan Institute of Metals (1994).
[11]G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).
[12]G. Binnig, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 9 (1986).
[13]F. J. Giessibl and G. Binnig, Ultramicroscopy **42-44**, 7 (1992).
[14]M. Ohta *et al.*, Jpn. J. Appl. Phys., Part 1 **32**, 2980 (1993).
[15]Th. Fries, K. Oster, and K. Wandelt, Acta Metall. Mater. **42**, 3129 (1994).
[16]M. Göken, H. Vehoff, and P. Neumann, Scr. Metall. Mater. **7**, 1187 (1995).
[17]B. Jaoul, *Etude de la Plasticité* (Dunod, Paris, 1965).

Secondary ion mass spectrometry and atomic force spectroscopy studies of surface roughening, erosion rate change and depth resolution in Si during 1 keV 60° O_2^+ bombardment with oxygen flooding

Z. X. Jiang[a] and P. F. A. Alkemade
Department of Applied Physics, Delft Institute of MicroElectronics and Submicron Technology (DIMES), Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

(Received 19 December 1997; accepted 3 April 1998)

Oxygen flooding during secondary ion mass spectrometry (SIMS) depth profiling is often used to achieve very short surface transients. However, for low-energy and obliquely incident O_2^+ primary beams, oxygen flooding can be detrimental. In this work we have measured as a function of depth the erosion rates, surface topographies and depth resolution for B and Ge in Si bombarded by 1 keV 60° O_2^+ with and without oxygen flooding. Using B and Ge deltas we showed that the erosion rate under oxygen flooding was not constant. The effect was most pronounced at intermediate flooding pressures; at saturation pressures, a drop of 25%–30% was found within ~25 nm below the surface. Atomic force microscopy measurements revealed that the erosion rate change was related to the onset of surface roughening. Oxygen flooding influenced the depth resolution in terms of the decay length for B and Ge in different ways. With oxygen flooding, the Ge decay length was larger than without flooding due to oxide-enhanced segregation; while the B decay length was smaller, due to swelling of the sample. In terms of the delta peak width, best depth resolution was always obtained without flooding. This study showed that oxygen flooding impedes accurate ultrashallow SIMS depth profiling. © *1998 American Vacuum Society.* [S0734-211X(98)00404-1]

I. INTRODUCTION

Crucial for rapid progress in microelectronics is the development of complementary metal–oxide–semiconductor (CMOS) transistors with ultrashallow junctions.[1] Within ten years junction depths as shallow as ~20 nm are expected to be in production. Shallow junctions require a good control and, thus, a good understanding of shallow doping of Si. This puts forward an increasingly stringent demand for accurate ultrashallow dopant analysis. Secondary ion mass spectrometry (SIMS) has proven to be the most powerful technique for dopant depth profile analysis, but quantification of dopant concentrations in the outermost several nanometers is still not at all trivial.[2] A major problem arises because a SIMS analysis exhibits a surface transient before achieving equilibrium in the sputtering process.[3,4] The surface transient can be reduced to a few nanometers if a low-energy primary beam is being used, but a further reduction is hindered by the presence of the native oxide on Si wafer materials.[5,6] Very low-energy (i.e., sub-keV) primary beams for SIMS became available only recently.[6–9]

Oxygen flooding during SIMS analysis is a well established method for enhancing the positive secondary ion yields.[10,11] It is also used in depth profiling to reduce the surface transient; ion yields reach equilibrium levels almost immediately.[5] Moreover, a number of studies have shown that the depth resolution for B in Si bombarded by 3–8 keV O_2^+ is better with oxygen flooding.[12,13] However, a recent study by Wittmaack and Corcoran[14] suggested that with oxygen flooding, the erosion rate in Si for an oblique 2 keV O_2^+ beam changes significantly after sputter removal of ~20–40 nm. Jiang and Alkemade[15] measured the change in erosion rate in Si for a 1 keV 60° O_2^+ beam using a Si sample with ten equidistant SiGe delta layers. They demonstrated that the erosion rate change with oxygen flooding was related to the onset of surface roughening. This result seems to contradict the work by Elst *et al.*,[16,17] who did not observe roughness by AFM for high-energy O_2^+ bombardment with saturated oxygen flooding. At intermediate flooding pressures, very rough surfaces developed.[16,17] It can therefore not be excluded that the roughness and the related change in erosion rate reported in Ref. 15 were caused by the presence of Ge. Regarding its wide-spread use, studies of the side effects of oxygen flooding on shallow depth profiling are of major importance. Especially for high depth resolution analysis at low primary beam energies, these studies are sparse, and hence, thorough investigations are urgently needed.

In this work we study in detail the phenomena related to sputter profiling of Si by 1 keV 60° O_2^+ with and without oxygen flooding. The erosion rate change, surface roughening, and the depth resolution for B and Ge are being measured. In order to clarify the alleged role of local changes in composition, we use different samples: two samples with multiple delta layers (Ge or B) and one that is homogeneously doped with B. The multiple delta layers are used to establish an intrinsic depth scale for the ion beam sputter profiling and to measure the depth resolution. The topography of the crater bottoms is measured with atomic force microscopy (AFM).

[a]Electronic mail: jiang@cerberus.dimes.tudelft.nl

II. EXPERIMENT

A 150 nm Si layer with four B delta layers was grown onto a 100 mm p-doped Si(100) wafer with molecular beam epitaxy (MBE) at 500 °C. We have chosen this temperature as a compromise between initial roughness of the surface and B diffusion. The areal density was $\sim 1\times10^{14}$ B atoms/cm^2 for each delta. The nominal depth of the first B delta was 6.0 nm, while the nominal distance between the four deltas was 15 nm. This sample is referred to as "Si(Bδ)." A second sample — "Si(Geδ)" — was a Si wafer with a 150 nm Si top layer having ten $Si_{70}Ge_{30}$ delta layers of 0.8 nm thickness (nominally), grown by atmospheric-pressure chemical vapor deposition (APCVD) at 700 °C. The nominal depths of the deltas were multiples of 13 nm. This sample was also used in a previous study.[15] In addition, we have analyzed an unprocessed Si wafer homogeneously doped with B at $\sim 8\times10^{18}$/cm^3; viz. the "Si(B)" sample. At the moment of the SIMS measurements, the Si(Bδ), Si(Geδ), and Si(B) samples had been kept in air for four days, five months and more than four years, respectively.

SIMS measurements have been performed on a Vacuum Generators IX70S magnetic-sector microprobe.[8] We bombarded the samples with a 1 keV O_2^+ beam of 50 nA at an incidence angle of 60° to the surface normal. The day-to-day reproducibility of the current measurement was $\sim 5\%-10\%$. The drift of the beam current during a one-day's measurement was usually less than 5%. Incidence angles in our instrument have been calibrated previously using the so-called trajectory-mapping method.[18] In general, final crater depths ranged between 10 and 180 nm; in one measurement we profiled down to 1.0 μm. For the oxygen flooding measurements, the crater size was 0.48×0.74 mm^2; for the UHV measurements, the crater area was 31% larger. Previous work[8] has shown that under these conditions (1 keV at 60° in our instrument) crater bottoms remain flat, at least for profiling of Si without flooding. The base pressure of the analysis chamber was lower than 5×10^{-9} Pa. For the flooding experiments, an oxygen flux was directed onto the sample from a 1 mm diameter pipe, 6 mm from the sample surface. We have measured SIMS profiles for all three samples under UHV conditions (thus, without flooding), at an intermediate oxygen pressure of 2.0×10^{-5} Pa, and at a saturation pressure of 1.3×10^{-4} Pa. (Saturation pressures are pressures for which secondary ion intensities of the matrix elements reach equilibrium almost at the start of the SIMS profiling.) Additional measurements at incremental oxygen pressures — from UHV to 1.4×10^{-4} Pa — have been done for the Si(Geδ) sample. The pressures were measured with a gauge close to a pump, 40 cm from the wafer surface. We estimate the pressure near the sample to be almost one order of magnitude higher. Secondary ion intensities of $^{11}B^+$ or $^{70}Ge^+$, together with those of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ were recorded from the central 6% of the rastered area. No attempt was made to ensure equal secondary ion transmission in all measurements. For each sample and flooding conditions, we produced craters of various depths. A Dektak profilometer was used to measure the crater depths. The final topography of the crater bottoms was measured with a Nanoscope III AFM with silicon tips manufactured by Digital Instruments; the nominal radius of the tip was 5 nm. The scanned areas were either 2×2 μm^2, 1×1 μm^2 or 500×500 nm^2. The AFM was operated in the tapping mode.

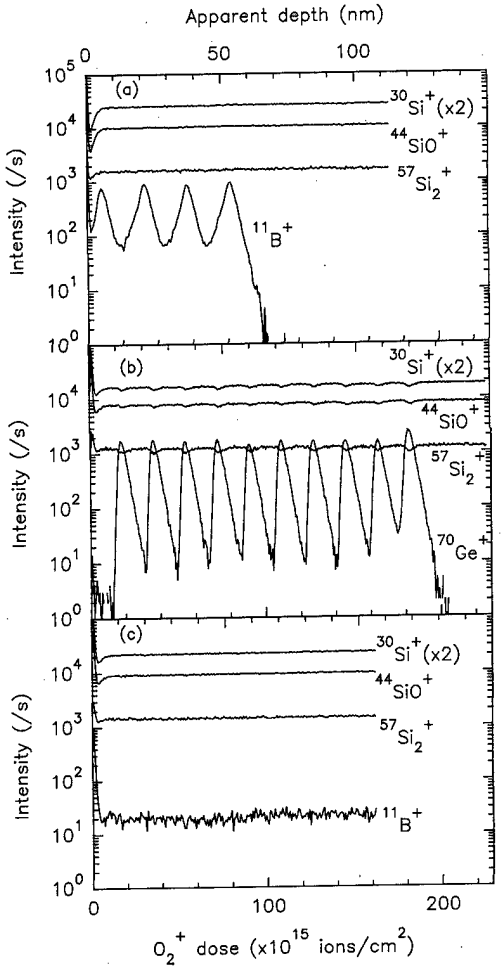


FIG. 1. SIMS profiles of secondary ions from the Si samples (a) with B deltas, Si(Bδ); (b) with Ge deltas, Si(Geδ); and (c) with homogeneous B, Si(B), analyzed with a 1 keV 60° O_2^+ beam at UHV.

III. RESULTS

A. SIMS profiles

Figure 1 shows the depth profiles of the secondary ions for the samples Si(Bδ), Si(Geδ) and Si(B), measured under UHV. The *apparent* depth was established via the measured final crater depths, assuming a constant erosion rate. After a short surface transient ($\sim 6\times10^{15}$ ions/cm^2), the intensities of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ stabilized, although the Ge deltas in Fig. 1(b) caused some variations. Note that the dips just after the surface peaks of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ are different for the different samples. Obviously this was due to the differences in native oxide thickness: the one stored longer in air had a thicker oxide.[19] The apparent depths (viz., peak maxima) of the B deltas are summarized in Table I. Their mutual distances are almost constant: 15.9 nm, 6%

TABLE I. Apparent depths of the Ge and B deltas in the Si wafers analyzed with 1 keV 60° and 50 nA O_2^+ beams at UHV, at an intermediate flooding pressure (2×10^{-5} Pa), and at a saturated flooding pressure (1.3×10^{-4} Pa). The last row "i+1−i" indicates the mean apparent distance between Ge deltas 5 to 10.

Number of delta layer	Apparent depth of delta (nm)		
	UHV	Intermediate pressure	Saturation pressure
Sample Si(Bδ)			
1	5.5	3.2	3.7
2	21.4	12.0	17.3
3	37.4	21.9	32.5
4	53.3	38.0	48.3
Sample Si(Geδ)			
1	11.7	7.8	9.1
2	23.6	15.5	21.3
3	35.5	23.9	33.7
4	47.5	35.3	46.3
5	59.4	49.3	58.8
$i+1-i$	12.0	14.0	12.9

larger than nominal. The Si cap layer, however, appears to be 0.5 nm thinner than nominal. For the Si(Geδ) sample, the thickness of the cap layer is 11.7 nm; the distances between the Ge peaks are 11.9 ± 0.1 nm. Close inspection of Fig. 1(a) reveals that the B leading slopes have two parts; first, a rise with a typical (exponential) rise length of 3.1 nm and then, just before the maximum intensity, a slightly faster rise. We attribute the slower rise to diffusion of B during the growth, the result of the relatively high growth temperature. Comparable samples that were grown at lower temperatures did not show the two-slope behavior,[8] but had rougher surfaces (0.6 nm standard deviation, as measured by AFM). They were, therefore, not useful for this study. The Ge profiles are better: the rise length of the first Ge peak is very short (0.3 nm) and there is no indication of segregation. Differences between the various peaks are marginal.

At intermediate oxygen flooding — 10^{-6}–10^{-5} Pa — the stability of the secondary ion intensities was lost. Figure 2 shows the profiles at a flooding pressure of 2.0×10^{-5} Pa. The intensities of all ions changed markedly after a dose of $\sim50\times10^{15}$ ions/cm^2; most intensities rose, especially those of $^{44}SiO^+$; only those of $^{57}Si_2^+$ dropped. Moreover, the delta peaks got broader, indicating a rapid loss of depth resolution. The outermost B and Ge deltas appear to be shallower now, see Table I. The apparent distances between the Ge deltas are $\sim$14 nm after the transition. We note that this value of 14 nm is partly fortuitous; it depended on the final crater depth. Additional measurements showed that the depth of the significant change in the secondary ion intensities was very sensitive to the current density and the oxygen pressure. Therefore, all measurements of Fig. 2 were completed in a single day and terminated after the same bombardment time. Although we got nearly identical profiles of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ for the samples Si(Geδ) and Si(B) [Figs. 2(b) and 2(c)], the profiles for sample Si(Bδ) were slightly different: the change in intensities occurred earlier. Measurement of

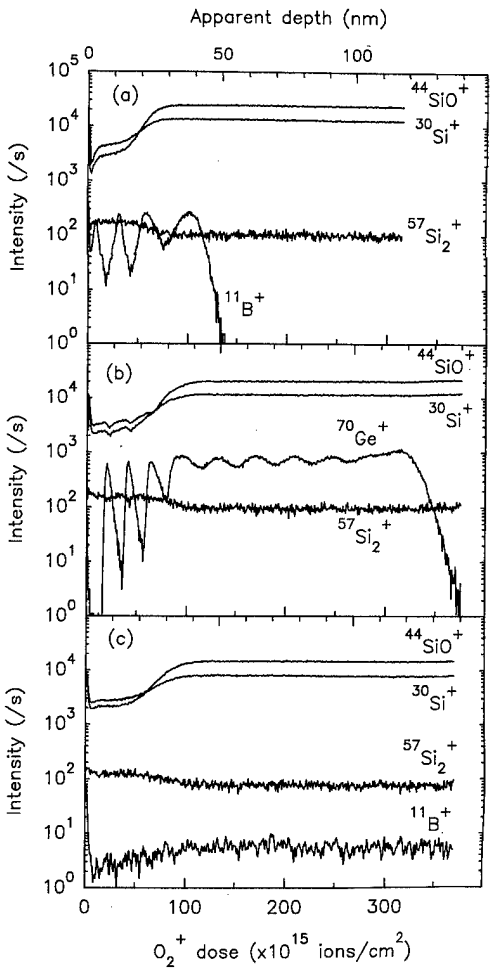


FIG. 2. SIMS profiles of secondary ions from the Si samples (a) Si(Bδ); (b) Si(Geδ); and (c) Si(B) analyzed with a 1 keV 60° O_2^+ beam at an intermediate oxygen flooding pressure (2.0×10^{-5} Pa).

the beam current after the SIMS profiling of the sample Si(Bδ) revealed that the beam current had incidentally dropped from 50 to 45 nA. The total erosion depth was 16% less.

With saturation oxygen flooding — typically above 4×10^{-5} Pa — the intensity of $^{30}Si^+$ reached equilibrium very rapidly after the start of the measurement. Figure 3 shows the profiles for our samples at an oxygen flooding pressure of 1.3×10^{-4} Pa. Note that the intensities of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ in Figs. 3(a)–3(c) and $^{11}B^+$ in Fig. 3(c) were constant, almost from the start. However, there was a small drop of 5–15% for $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ after $\sim120\times10^{15}$ ions/cm^2. The apparent depths of the B deltas and the outermost Ge deltas are also shown in Table I. These deltas appear to be shifted towards the surface as compared to the UHV measurement, but the shifts are less than for the intermediate pressures.

B. AFM images

The change of the measured SIMS intensities suggests a significant alteration of the bombarded surface. To investigate the state of the surface after SIMS analysis with and

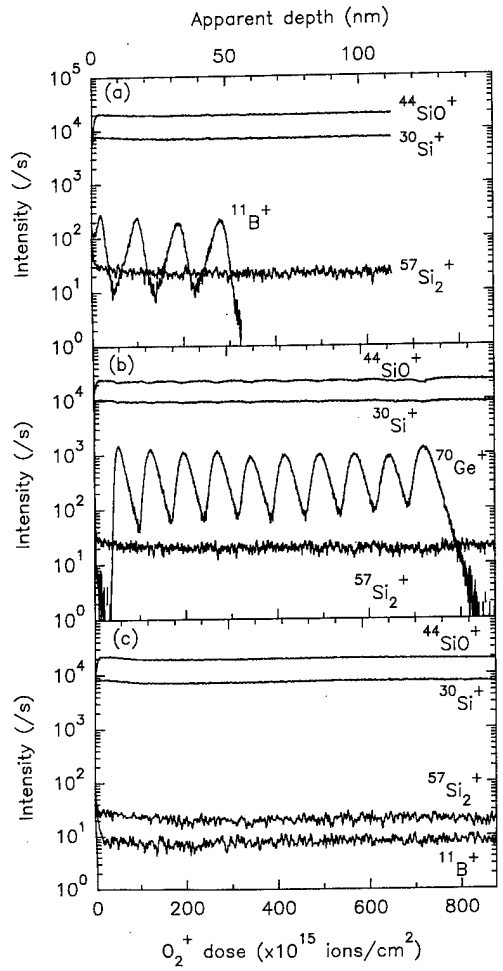


FIG. 3. SIMS profiles of secondary ions from the Si samples (a) Si(Bδ); (b) Si(Geδ); and (c) Si(B) analyzed with a 1 keV 60° O_2^+ beam at a saturation oxygen flooding pressure (1.3×10^{-4} Pa).

without oxygen flooding, the crater bottoms were profiled with AFM. Figures 4, 5, and 6 show the surface topography of the crater bottoms of our samples bombarded under UHV and flooding conditions. The craters are those of the measurements of Figs. 1–3. Clearly, one sees that regular ripples have developed under all conditions. However, with oxygen flooding the surface acquired an additional, irregular type of roughness. The differences among the three types of samples are marginal.

To understand the influence of oxygen pressure on surface roughening in more detail, crater bottoms were produced in the Si(Geδ) sample at various pressures. Figure 7 shows the 1×1 μm^2 AFM images of the ~160 nm deep crater bottoms. With oxygen flooding at 1×10^{-6} Pa, the topography was composed of irregular large bumps and hollows. With more flooding — 6×10^{-6} and 9×10^{-6} Pa — also thick and irregular ripples were observed. At an oxygen pressure of 2×10^{-5} Pa — i.e., the transition between the intermediate and saturation pressures — the ripples got thinner and more regular. At and above 3.6×10^{-5} Pa, the topography was composed of very regular fine ripples and irregular, but relatively weak, large bumps and hollows. Interestingly, the topography is very similar to that obtained by Mayer *et al.* on SiO_2 bombarded by 1 keV 55° Xe^+.[20] Figure 8 shows the root-mean-square (rms) of the height variations and the typical peak-to-valley height difference of the ripples as a function of the oxygen pressure. At UHV and the intermediate pressures, the ratio of the peak-to-valley value to rms was ~2.5. At the saturation pressures, the ratio was 1.5. The amplitude of the ripples at saturation was comparable to that at UHV, but there was a distinct additional irregular roughness. The surface was smoothest at UHV; at the intermediate pressures, it was roughest. Above 4×10^{-5} Pa, the roughness was independent of pressure, but still larger than at UHV. We note that the craters of Figs. 7(d) and 7(h) were also profiled with AFM in a previous study.[15] The reported rms and peak-to-valley values were then ~30% lower, probably due to the poor quality of the tip in that study.

Often the surface roughness depends on the erosion depth.[21,22] Figures 9–11 show the bottoms of craters at various depths. The Si(B) sample was bombarded at UHV, at an intermediate pressure (1.6×10^{-5} Pa), and at a saturation pressure (1.3×10^{-4} Pa), respectively. The depths of the deeper craters were measured with the profilometer, while for the craters shallower than 30 nm they were deduced from

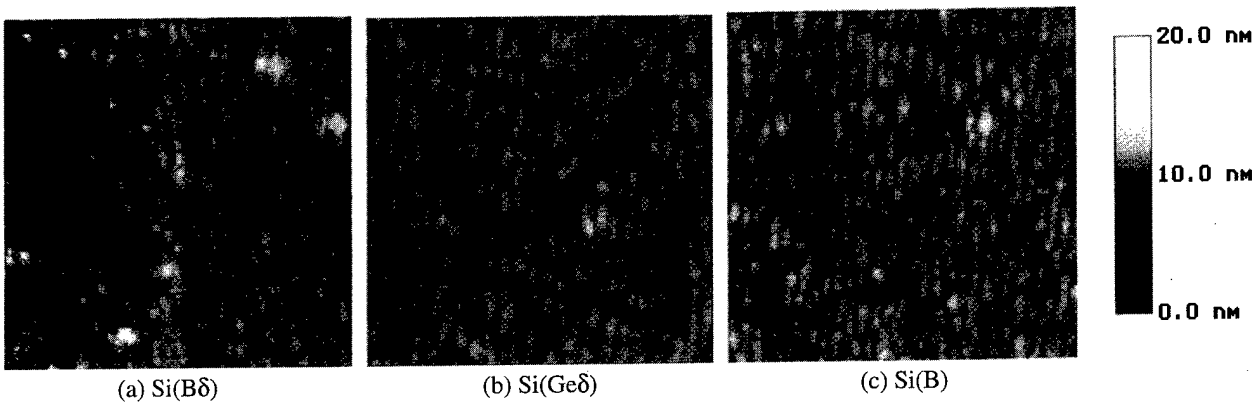


FIG. 4. 1×1 μm^2 AFM images of the crater bottoms after SIMS profiling at UHV (see Fig. 1). Beam entered from the right.

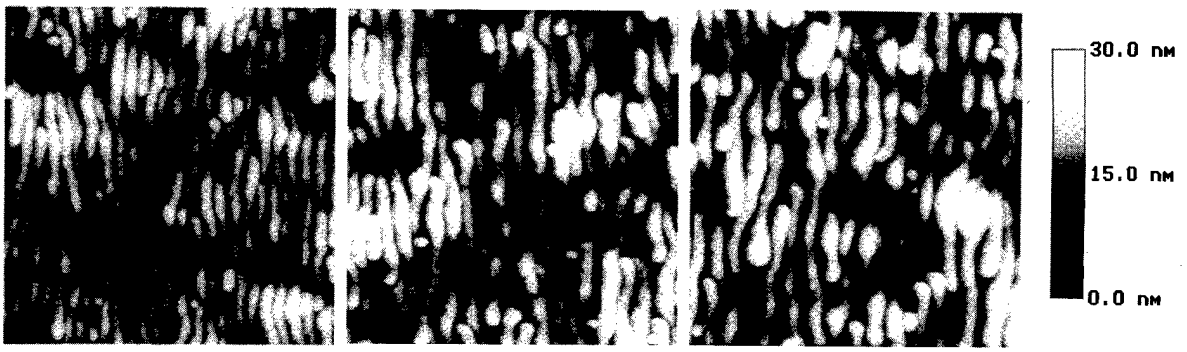


FIG. 5. 1×1 μm^2 AFM images of the crater bottoms after SIMS profiling at intermediate flooding (see Fig. 2). Beam entered from the right.

the erosion time and the erosion rate. (The accurate determination of the depths will be discussed in the next section.) The topography of the 27 nm deep UHV crater bottom did not differ markedly from that of the original surface. At a depth of 52 nm, ripples became visible; beyond 100 nm, the ripples were very clear [Fig. 9(e), see also Fig. 4]. At the intermediate pressure of 1.6×10^{-5} Pa, the development of the surface roughness was complex (Fig. 10). The ripples observed at the depth of 29 nm were very similar to the UHV ones beyond 100 nm (see, e.g., Fig. 4). At 56 nm, the topography was composed of many thin bars and small bumps. At 70 nm, the bars were thicker and the bumps larger. The topography beyond 100 nm was mainly composed of thick, irregular ripples. Figure 11 shows the AFM images for a saturation pressure of 1.3×10^{-4} Pa. The regular and fine ripples appeared already at a depth of 9 nm. As the erosion proceeded, the magnitude of the ripples increased and the distance between adjacent ripples rose slightly from 19 nm [Fig. 11(b)] to 22 nm [Fig. 11(e)]. An irregular pattern became visible in the 24 nm deep crater [Fig. 11(b)]. The irregular bumps and hollows grew wider with increasing erosion depth. In Fig. 12 we summarize the roughness — viz., rms values — for the images of Figs. 9–11. One sees that at UHV conditions, the roughness rose only slightly with erosion depth. The initial rms was 0.15 nm; at 1 μm it was still only 0.9 nm. At the intermediate pressure, however, the rms values increased first rapidly to ~3.5 nm at 50 nm and then slower. For the saturation pressure, there was a rapid rise near 20 nm. Thereafter, it rose slowly to ~1 nm at 170 nm, about two times higher than under UHV conditions.

IV. DISCUSSION

The increasing demand for analysis of ultrashallow junctions in Si stimulates research on surface transients during SIMS profiling. Although the introduction of oxygen flooding seems beneficial — the surface transient in terms of the ion yields is very short — the erosion rate change within the initial tens of nanometers is indicative of a much longer transition.[14] Understanding of the complex phenomena related to oxygen flooding is therefore indispensable for proper analysis of ultrashallow profiles. In this section we discuss the measured change in erosion rate, the surface roughening and the depth resolution as functions of the oxygen flooding pressure and the erosion depth. First, however, we will show that these phenomena are intrinsic to Si and not related to the added doping elements.

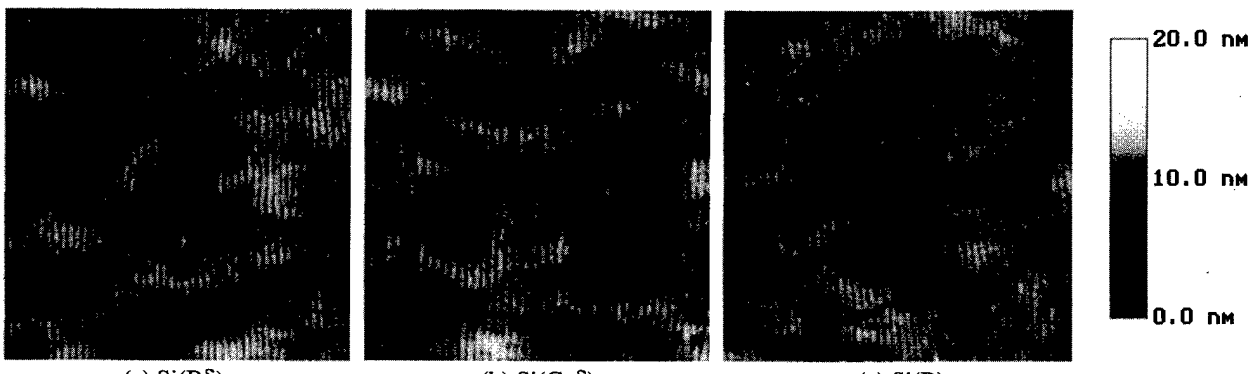


FIG. 6. 1×1 μm^2 AFM images of the crater bottoms after SIMS profiling at saturation flooding (see Fig. 3). Beam entered from the right.

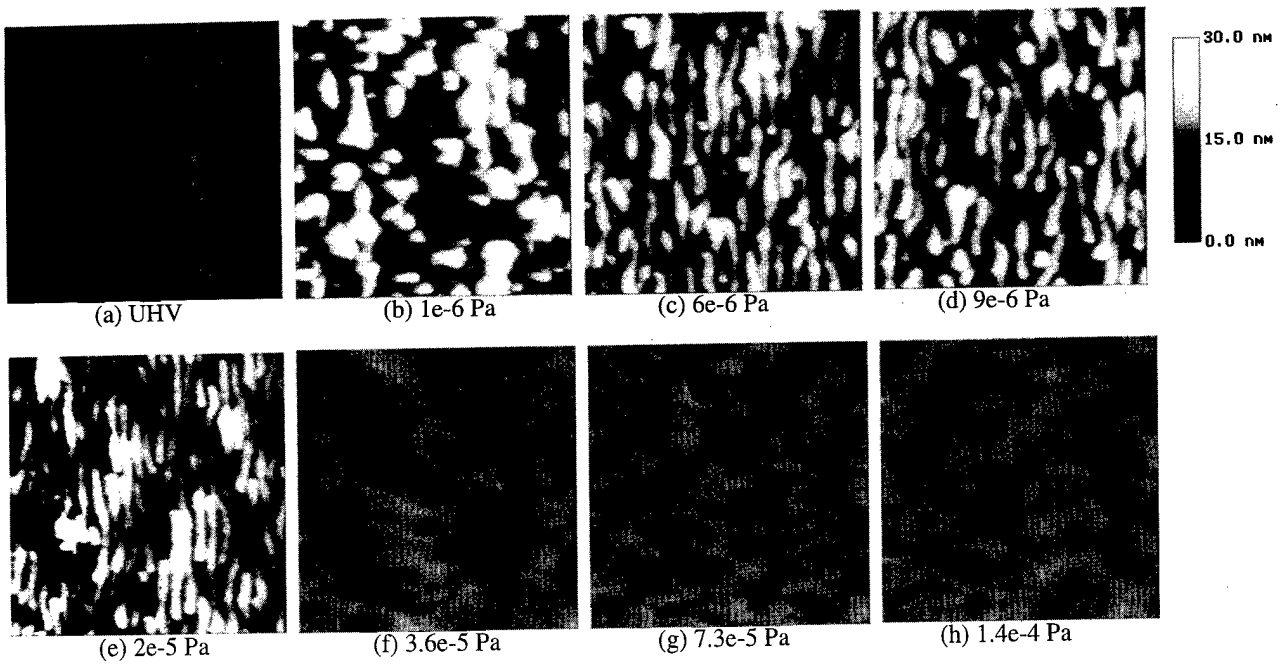


FIG. 7. 1×1 μm^2 AFM images of ~160 nm deep crater bottoms of the sample Si(Geδ), profiled at various oxygen flooding pressures. Beam entered from the right.

A. Surface roughening of Si samples with various dopings

In a previous study[15] we have shown that the erosion rate change in the Si sample with the ten Ge delta layers [Si(Geδ)] was related to the onset of surface roughening. It could, however, not be excluded that the roughening was induced by the presence of Ge. Local strain in a layer with a high Ge concentration might have initiated the roughness. In the present work we analyzed, besides the Si(Geδ) sample, a plain Si wafer evenly doped with 8×10^{18} cm^{-3} B [Si(B)] and a Si wafer with four B deltas [Si(Bδ)]. The measured SIMS profiles of $^{30}Si^+$, $^{44}SiO^+$, and $^{57}Si_2^+$ were similar for all three samples at all selected flooding conditions (Figs. 1–3). Only for the sample Si(Bδ) at the intermediate flooding pressure (2×10^5 Pa), the change of the secondary ion intensities occurred earlier, which was very likely the result of a lower primary beam current. We note that the 10% drop in the current induced a 16% drop in the average erosion rate. This nonlinear behavior is probably associated with the dynamic process of oxygen incorporation in the surface during ion beam bombardment with oxygen flooding.[23] The AFM measurements of the crater bottom topographies (Figs. 4–6) confirmed that the effect of oxygen flooding was the same for the three types of samples studied. And again, the Si(Bδ) sample at the intermediate pressure (2×10^{-5} Pa) deviated slightly: the ripples were thinner and the topography looked like that of a crater produced with the usual beam current of 50 nA but at a slightly higher oxygen pressure. Attributing the minor differences to small drifts in the experimental conditions, we conclude that there is no effect of the dopants on surface roughening. This result is not surprising because the highest Ge concentration after ion beam mixing was between 5% and 10%; for B it was less than 0.1%. Surface roughening during SIMS under oxygen flooding is, thus, an intrinsic property of silicon.

B. Change in erosion rate

Because of the varying erosion rate under oxygen flooding conditions, depth calibration in the outermost tens of nanometers is not trivial. In principle, the erosion rate varia-

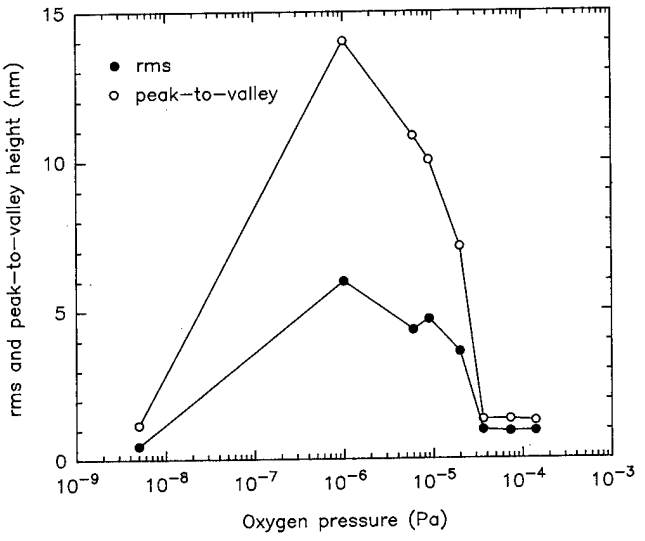


FIG. 8. Root-mean-square (rms) of the surface roughness and the typical peak-to-valley height of the ripples vs oxygen flooding pressure (data of Fig. 7).

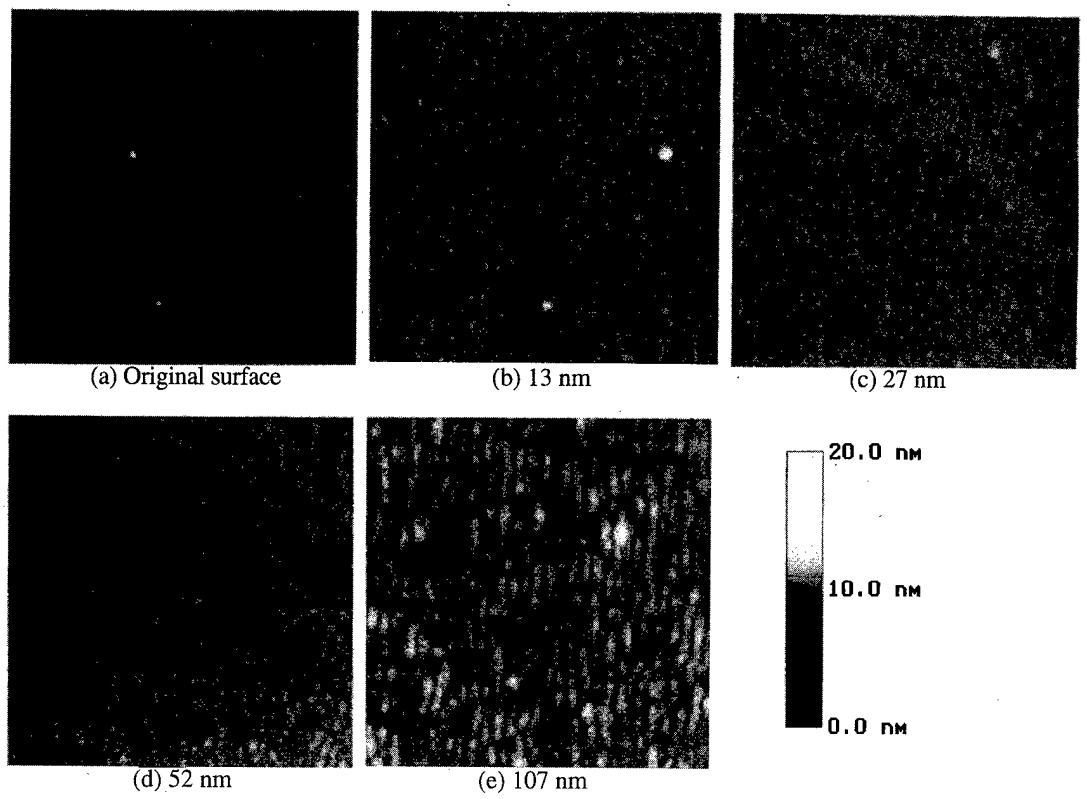


FIG. 9. 1×1 μm^2 AFM images of the crater bottoms of various depths for the sample Si(B) eroded at UHV.

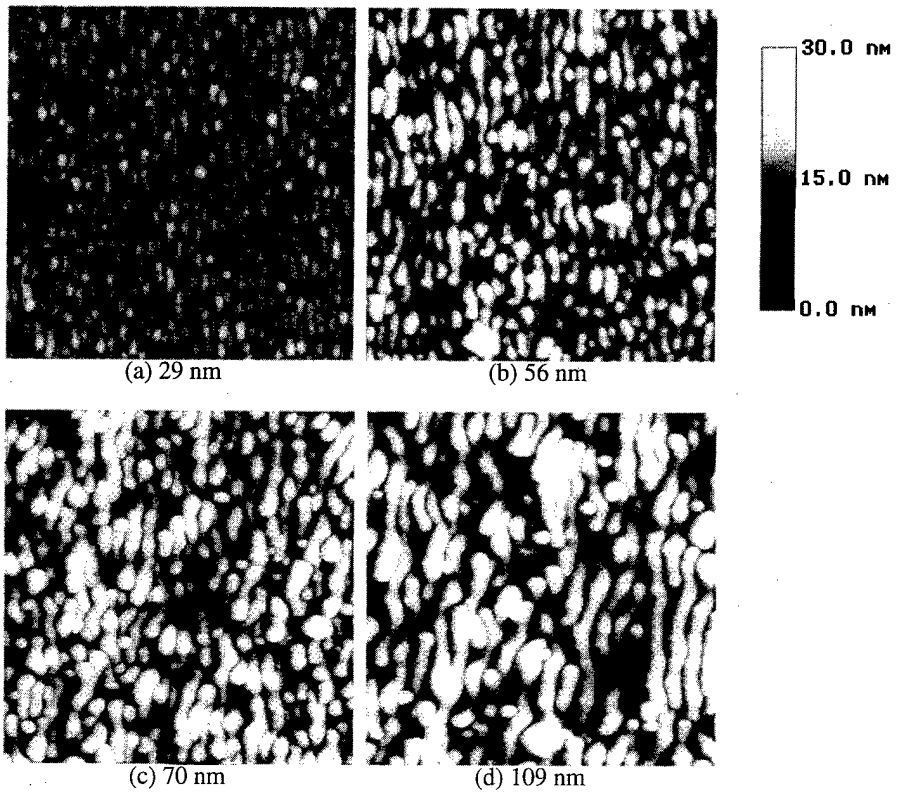


FIG. 10. 1×1 μm^2 AFM images of the crater bottoms of various depths for the sample Si(B) eroded at an intermediate oxygen flooding pressure (1.6×10^{-5} Pa).

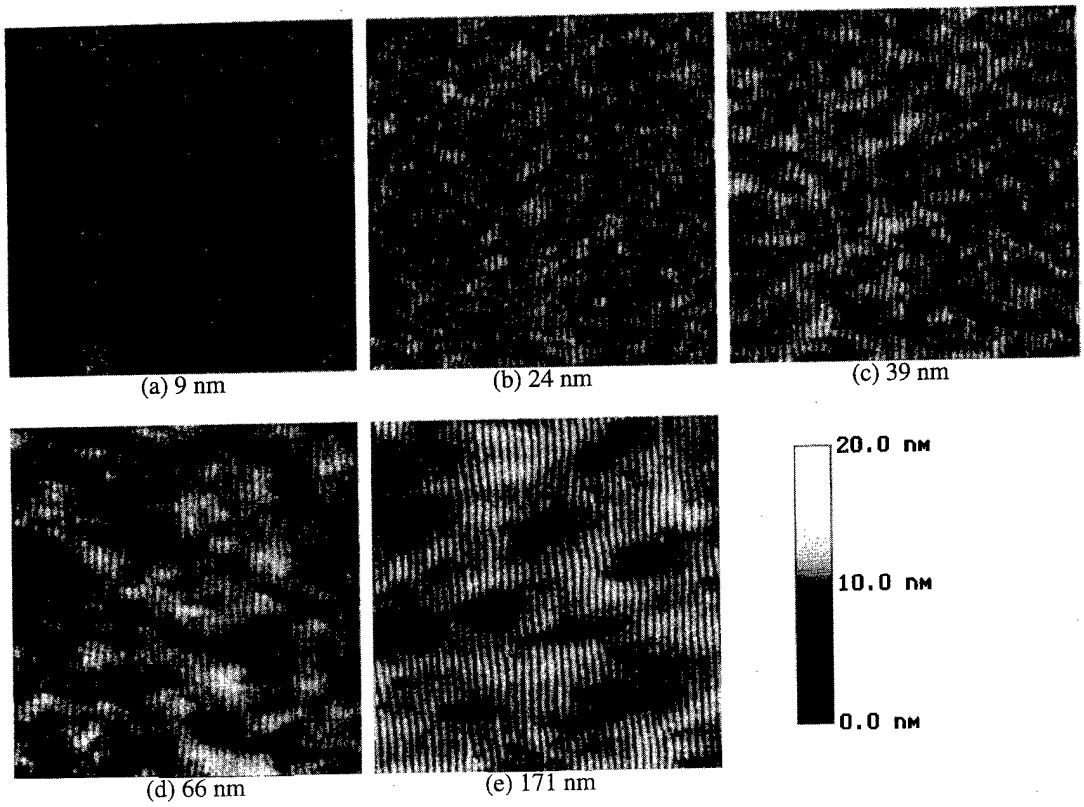


FIG. 11. 1×1 μm^2 AFM images of the crater bottoms of various depths for the sample Si(B) eroded at a saturation oxygen flooding pressure (1.3×10^{-4} Pa).

tion can be measured by terminating the SIMS analysis after various O_2^+ doses and measuring the crater depths with a profilometer. But the relative inaccuracy of shallow crater depth measurements is very large.[24] Therefore, we used in this study the delta layers for an intrinsic depth calibration. Neglecting the small differential shifts,[2,19] we regard the apparent depths of the deltas measured under UHV as the real depths. Indeed, the observed constancy of the sputter time between the peaks is reassuring. Assuming for the moment that with oxygen flooding, the shift of the measured Ge peak positions was induced only by the change in erosion rate, we calculated the erosion rate (nanometers per unit of dose) vs the real erosion depth for various oxygen pressures. The result is shown in Fig. 13. It is evident that the erosion rate was not constant during depth profiling with oxygen flooding. For an oxygen pressure of 1×10^{-6} Pa, a drop of $\sim10\%$ occurred at the depth of around 80 nm. At 9×10^{-6} Pa, the erosion rate dropped by nearly a factor of 2 between 30 and 70 nm. For the saturation pressures, the erosion rate was about four times lower than at UHV and there was a drop of 22% within 20 nm below the surface. The B data — i.e., the open circles in Fig. 13 — show the same trend: a 30% drop within the outermost 30 nm. We attribute the difference in the final erosion rate between the two samples to the inaccuracy in the measurement of the primary beam current. Obviously, if the depth were calibrated in the usual way, i.e., via a simple mean erosion rate (see the dotted lines in Fig. 13), it would have resulted in too shallow and too narrow profiles of the shallow layers.

In the above determination of the depths of the deltas, we have discarded the differential shift at UHV.[2] In a separate study,[19] we have measured a shift towards the surface of ~0.2 nm for a Ge delta and ~0.5 nm for a B delta. Neglect of the UHV differential shift in the present work resulted in an underestimation of the erosion rate between the surface and the first delta under flooding conditions of $\sim2\%$ for Ge

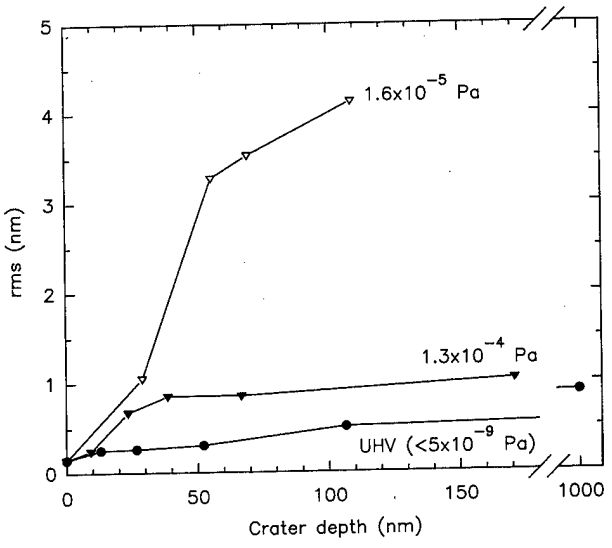


FIG. 12. Root-mean-square (rms) of the surface roughness vs crater depth for various flooding pressures. Data from Figs. 9–11, but data point at 1000 nm from the sample Si(Ge).

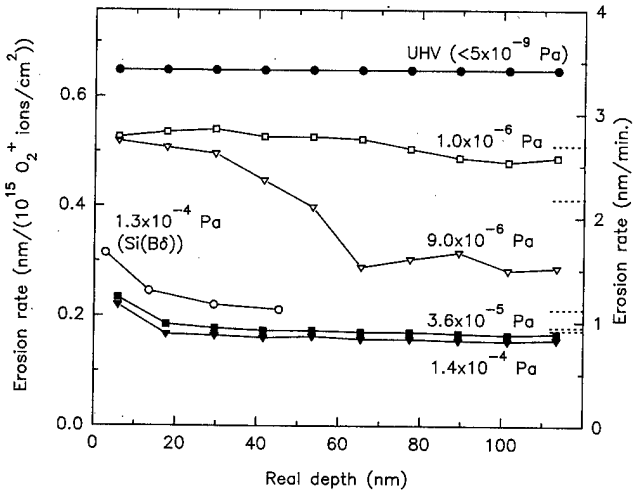


FIG. 13. Erosion rate vs real erosion depth for various oxygen pressures. The right axis is the erosion rate in nm/min. for flooding conditions. The dotted lines indicate the corresponding average erosion rates (=final crater depth/total erosion time).

and ~9% for B. Furthermore, also the chemically-induced relocation under flooding conditions has not been considered. With saturation flooding, the surface was fully oxidized almost from the start. Since Ge segregates towards the oxide/substrate interface,[25] a larger chemically induced relocation is expected as compared to UHV. Thus, if we had taken this chemical effect for Ge under flooding properly into account, the drop in the erosion rate in the outermost ~20 nm would have been even larger than 22%. Boron atoms in Si, on the other hand, behave similar to the Si atoms.[25] Thus, no chemical effects are expected for B. Indeed, the B measurement showed a larger drop in erosion rate in the outermost 30 nm than Ge.

C. Surface roughening

Beam-induced surface roughness has been intensively studied (e.g., Refs. 22, 26–29). It is known that surface roughening depends on material properties and primary beam species, impact energy, and incidence angle. Generally, surface roughness increases with erosion depth.[21] Moreover, it can be different for different ambients.[16,17] Also the details of the topography (e.g., ripples, cones, or pits) can be quite different.

Stevie *et al.*[27] and Wittmaack[28] showed that for Si wafers bombarded by 5.5–10.5 keV O_2^+ under UHV conditions, ripples formed at erosion depths between 0.3 and 4.2 μm when the incidence angle was within the range of 32° to 58°. Outside that range no roughness developed.[27] In contrast, ripples became already visible at an erosion depth of only 50 nm for our 1 keV 60° O_2^+ beam under UHV [Fig. 9(d)]. Apparently, surface roughening occurs much earlier for lower impact energies. A similar behavior was found by Cirlin *et al.* on AlGaAs/GaAs superlattices bombarded by 1.0–7.0 keV O_2^+ beams at a fixed incidence angle of 40°.[22] Although we observed surface roughening by the 1 keV 60° O_2^+ beam at a rather early stage, its subsequent development was slow (Fig. 12). The topography of a 1 μm deep crater bottom was similar to that of a 100 nm deep one; the rms height variations were only slightly (0.4 nm) larger. Furthermore, we observed a small but gradual increase (~5%–10%) in the secondary ion intensities of SiO^+ and Si^+ in the outermost 200 nm. Between 200 nm and 1 μm, the yield remained constant within 2%. In general, the intensity changes that accompany the onset of surface roughening are much larger.[27,28]

Under flooding conditions, surface roughening depends on the oxygen pressure. Elst *et al.* observed that a Si surface bombarded by an 8 keV 37° O_2^+ beam roughened quickly and heavily at intermediate flooding pressures;[16,17] at a depth of 1 μm, the rms value was as large as ~150 nm. However, the authors did not observe roughness at a saturation flooding pressure. Studies by Zalm and Vriezema[12] and Erickson and Brigham[13] suggested that saturation flooding can suppress surface roughening as compared to UHV. In the present study at low impact energy, the surface roughened not only at the intermediate, but also — though less pronounced — at the saturation pressures (Fig. 7). The roughness at saturation was about two times larger than under UHV. Futhermore, the roughness under flooding developed initially very rapidly (Fig. 12); later on, i.e., beyond 40–50 nm, the topography of the surface changed only little.

Interestingly, the topography at the saturation pressures was composed of very regular fine but long ripples superposed on an irregular bumpy pattern [Figs. 6, 7(f)–7(h) and 11]. The wavelength of the ripples — viz., the distance between adjacent ripple maxima — was on average 23 nm and changed only slightly with the erosion depth (Fig. 11). For our UHV conditions, the wavelength was ~45 nm. At higher beam energies much larger wavelengths (hundreds of nanometers) were observed for Si and other semiconductor materials.[16,30,31] Unfortunately, the formation of ripples during ion bombardment is not fully understood despite the numerous studies on this subject.

In general, the onset of significant surface roughening causes a change in erosion rate.[28] For our intermediate flooding pressures, this was clearly the case. Very likely, the onset of surface roughening at ~20 nm for saturated flooding (Fig. 12) is also directly related to the change in the erosion rate at about the same depth (Fig. 13). Tian and Vandervorst suggested a different mechanism of erosion rate change under flooding conditions.[32] The outermost surface layer was and remains SiO_2 from the beginning of sputtering under saturation flooding. Nevertheless, a thicker oxide layer needs to be built up in the sample, which leads to a gradual reduction in sputter rate. However, quantification of their model is as yet not possible.

D. Depth resolution

In general, the depth resolution of SIMS is influenced by atomic mixing, surface roughening, segregation, etc.[33–36] Oxygen flooding affects the depth resolution because it influences surface roughening and induces additional chemical effects. Zalm and Vriezema[12] showed that for P, Ga, and Sb

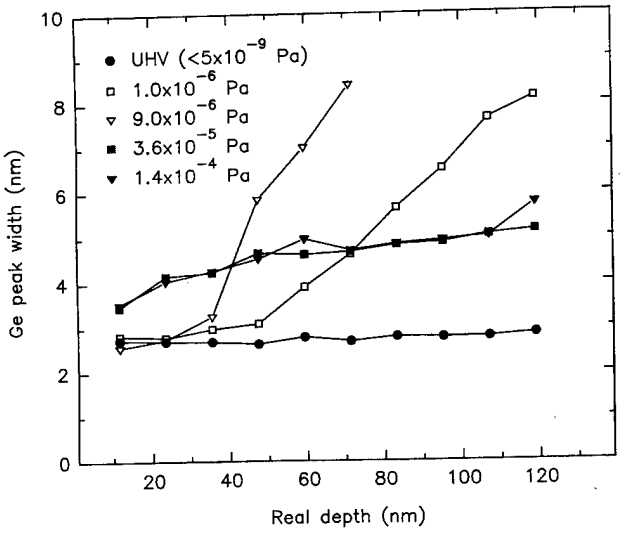


FIG. 14. Ge peak width vs real erosion depth for various oxygen pressures [sample Si(Geδ)].

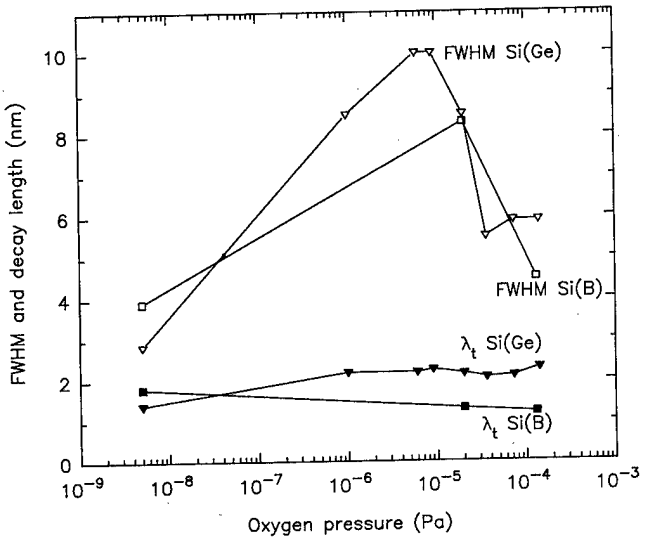


FIG. 16. Peak width, FWHM, and decay length λ_t of Ge and B as functions of oxygen pressure.

deltas in Si bombarded by 1.5, 3, and 6 keV O_2^+, oxygen flooding enlarged both the peak widths and the decay lengths as compared to UHV. The same holds for B deltas bombarded by 1.5 keV O_2^+; only for B analyzed with 3 and 6 keV O_2^+, the decay length was improved with flooding. Using 3 and 8 keV O_2^+ beams, also Erickson and Brigham found shorter decay lengths for B in Si with oxygen flooding.[13] However, for our 1 keV 60° O_2^+ beam, oxygen flooding always resulted in broader peaks of both B and Ge. Figures 14 and 15 show the measured real peak widths (FWHM) of the Ge and B deltas vs the real erosion depth for various oxygen pressures. At UHV, the peak widths remained nearly constant with increasing erosion depth. Although the surface roughness was already visible at a depth of 50 nm (Fig. 9), it had a negligible influence on the peak width. Indeed, the rms value of the roughness (Fig. 12) is too low (~0.4 nm) to expect any effect. For the intermediate pressures, it looks evident that solely the surface roughness determined the measured peak widths. For the saturation pressures, the Ge width increased first rapidly and then slower. At ~100 nm it was already 3 nm larger than under UHV. The B peak width, however, increased only slightly with the erosion depth. In the shallow range studied, it was only ~0.5 nm larger than under UHV. In any case, the roughness (~1 nm rms) is too small to explain the measured B peak width of 4.4 nm (FWHM).

Figure 16 shows the peak width (FWHM) and the decay length, λ_t, for both B and Ge as a function of the oxygen pressure. The data were deduced from the deepest Ge delta (at 120 nm) and the deepest B delta (at 53 nm). The B and Ge peak widths increased with the oxygen pressure up to 10^{-5} Pa and then decreased, roughly in agreement with the AFM results of Fig. 8. It is clear that at the intermediate pressures, the measured B and Ge peak widths were determined by the surface roughness. But for UHV and the saturation pressures, other effects dominated. At UHV, the measured peak widths — 3.9 nm for B and 2.8 nm for Ge — were mainly governed by beam-induced atomic mixing and the quality of the sample. Especially the relative high temperature during growth of the Si(Bδ) sample caused some broadening by diffusion beforehand. At saturation, the B peak width was slightly larger (4.4 nm), but still ~1 nm smaller than that of Ge. Apparently, B suffered only a little from the surface roughening at saturation; Ge, on the other hand, suffered also from another factor, very likely segregation.

The decay length for Ge remained nearly unchanged when the oxygen pressure was increased from 1×10^{-6} to 1.4×10^{-4} Pa, but it was in general 0.8 nm larger than at UHV. The surface roughness had no apparent influence, in agreement with the data by Smirnov *et al.*[37] In contrast, the decay length of B dropped continuously when the pressure

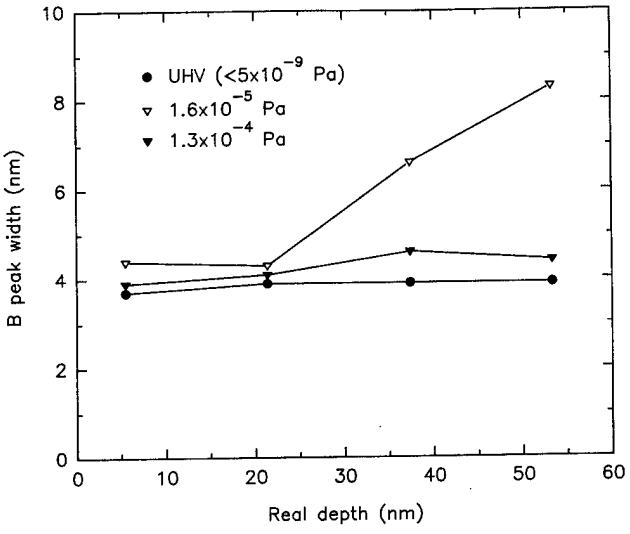


FIG. 15. B peak width vs real erosion depth for various oxygen pressures [sample Si(Bδ)].

was increased. Since the topographies are the same, we conclude that the opposite changes in their decay lengths with oxygen flooding are due to chemical effects. Note that under the saturated oxygen flooding, the surface was fully oxidized. Also for the intermediate pressures, where roughening was most pronounced, parts of the bombarded surface were very likely fully oxidized.[16,17,30] Since Ge segregates towards the oxide/substrate interface,[25] its sputter removal is hindered. This explains the observed larger Ge decay length under all flooding conditions. In contrast, B atoms in Si behave similar to Si atoms.[38] Under flooding conditions, the surface oxidation caused swelling of the sample, which resulted in the shorter decay length of B.[39]

It might look surprising that the decay length of B under flooding conditions was small in spite of the rough surface. The effect of roughness on the exponential tail of a SIMS profile is, however, limited. It can be estimated by convoluting the exponential tail with a function that describes the roughness. It is easy to show mathematically that the exponential tail of the convoluted profile is shifted with respect to that of the original profile, but it has the same decay length.[37] Hence, a good decay length alone is not a sufficient indication of a good depth resolution.

V. SUMMARY AND CONCLUSIONS

Oxygen flooding during low-energy (1 keV) and oblique (60°) O_2^+ bombardment of Si induces surface roughening and erosion rate changes. These phenomena are investigated in this work using MBE and CVD grown Si with multiple B and Ge delta layers. The deepest delta is at 120 nm. The surface roughening and erosion rate change are not related to the presence of dopants. The effects are most pronounced at intermediate flooding pressures. At saturation pressures, a drop of 25%–30% in the erosion rate in the outermost ~25 nm occurs. We relate these drops to the onset of surface roughening, that we observed by atomic force microscopy. However, they may also be related to the gradual build-up of a thick oxide layer.[32] Obviously, the erosion rate change complicates depth calibration. The common but simple procedure, i.e., the use of a constant erosion rate, results under flooding conditions in too shallow and too narrow profiles for shallow dopants. Beam-induced roughness and chemical and physical effects related to oxygen incorporation in the surface layer under oxygen flooding influence the depth resolution for B and Ge in Si in different ways. At intermediate pressures, the large surface roughness determines the peak widths of B and Ge deltas. At saturation pressures, roughening, atomic mixing and chemical effects determine the peak widths, while only the latter two are important at UHV. However, with oxygen flooding, the Ge decay length increases (probably due to oxide enhanced segregation) and the B decay length decreases (probably by swelling of the sample). It is evident from our detailed AFM and SIMS studies with delta layers in Si that a good decay length alone does not guarantee good depth resolution. Regarding the still limited knowledge of surface roughening and the stringent demands for ultrashallow junctions in future generation CMOS devices, the depth resolution achievable with SIMS using low-energy primary beams with and without oxygen flooding should be investigated further. As long as more data for low-energy and oblique O_2^+ beams are not available, we dissuade oxygen flooding for ultrashallow SIMS analysis.

ACKNOWLEDGMENTS

The authors are grateful to B. G. M. de Lange and C. C. G. Visser from the Delft University of Technology for growing the samples.

[1] *The National Technology Roadmap for Semiconductors* (Semiconductor Industry, San Jose, CA, 1994).
[2]K. Wittmaack, Philos. Trans. R. Soc. London, Ser. A **354**, 2731 (1996).
[3]W. Vandervorst and F. R. Shepherd, Appl. Surf. Sci. **21**, 230 (1985).
[4]W. Vandervorst, F. R. Shepherd, J. Newman, B. F. Philips, and J. Remmerie, J. Vac. Sci. Technol. A **3**, 1359 (1985).
[5]W. Vandervorst and F. R. Shepherd, J. Vac. Sci. Technol. A **5**, 313 (1987).
[6]P. F. A. Alkemade, Z. X. Jiang, C. C. G. Visser, S. Radelaar, and W. M. Arnoldbik, J. Vac. Sci. Technol. B **16**, 373 (1998).
[7]M. G. Dowsett, N. S. Smith, R. Bridgeland, D. Richards, A. C. Lovejoy, and P. Pedrick, in *Secondary Ion Mass Spectrometry, SIMS 10*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, 1997), p. 367.
[8]Z. X. Jiang, P. F. A. Alkemade, E. Algra, and S. Radelaar, Surf. Interface Anal. **25**, 285 (1997).
[9]K. Iltgen, C. Bendel, E. Niehuis and A. Benninghoven, in *Secondary Ion Mass Spectrometry, SIMS 10*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, 1997), p. 375.
[10]K. Wittmaack, Nucl. Instrum. Methods **168**, 343 (1980).
[11]M. L. Yu and W. Reuter, J. Appl. Phys. **52**, 1489 (1981).
[12]P. C. Zalm and C. J. Vriezema, Nucl. Instrum. Methods Phys. Res. B **64**, 626 (1992).
[13]J. W. Erickson and R. Brigham, J. Vac. Sci. Technol. B **14**, 353 (1996).
[14]K. Wittmaack and S.F. Corcoran, J. Vac. Sci. Technol. B **16**, 272 (1998).
[15]Z. X. Jiang and P. F. A. Alkemade, in *Secondary Ion Mass Spectrometry, SIMS 11*, edited by G. Gillen and R. Lareau (Wiley, Chichester, 1998), p. 431.
[16]K. Elst, W. Vandervorst, J. Alay, J. Snauwaert, and L. Hellemans, J. Vac. Sci. Technol. B **11**, 1968 (1993).
[17]K. Elst and W. Vandervorst, J. Vac. Sci. Technol. A **12**, 3205 (1994).
[18]Z. X. Jiang and P. F. A. Alkemade, Surf. Interface Anal. **25**, 817 (1997).
[19]Z. X. Jiang and P. F. A. Alkemade (unpublished).
[20]T. M. Mayer, E. Chason, and A. J. Howard, J. Appl. Phys. **76**, 1633 (1994).
[21]D. Matron and J. Fine, Thin Solid Films **151**, 433 (1987).
[22]E. H. Cirlin, J. J. Vajo, R. E. Doty, and T. C. Hasenberg, J. Vac. Sci. Technol. A **9**, 1395 (1991).
[23]L. Cha, W. Lin, Z. Jiang, Y. Tong, Z. Mo, and Q. Zou, in *Secondary Ion Mass Spectrometry, SIMS 8*, edited by A. Benninghoven, K. T. F. Janssen, J. Tumpner, and H. W. Werner (Wiley, Chichester, 1992), p. 53.
[24]D. S. Simons, in *Secondary Ion Mass Spectrometry, SIMS 10*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, 1997), p. 435.
[25]K. Wittmaack, Surf. Interface Anal. **21**, 323 (1994).
[26]R. M. Bradley and J. M. E. Harper, J. Vac. Sci. Technol. A **6**, 2390 (1988).
[27]F. A. Stevie, P. M. Kahora, D. S. Simons, and P. Chi, J. Vac. Sci. Technol. A **6**, 76 (1988).
[28]K. Wittmaack, J. Vac. Sci. Technol. A **8**, 2246 (1990).
[29]D. W. Moon and K. J. Kim, J. Vac. Sci. Technol. A **14**, 2744 (1996).
[30]A. Karen, K. Okuno, F. Soeda, and A. Ishitani, J. Vac. Sci. Technol. A **9**, 2247 (1991).
[31]S. W. MacLaren, J. E. Baker, N. L. Finnegan, and C. M. Loxton, J. Vac. Sci. Technol. A **10**, 468 (1992).
[32]C. Tian and W. Vandervorst, J. Vac. Sci. Technol. A **15**, 452 (1997).
[33]S. Hoffmann, J. Vac. Sci. Technol. A **9**, 1466 (1991).

[34]S. Hoffmann, Appl. Surf. Sci. **70/71**, 9 (1993).
[35]P. C. Zalm, in *Secondary Ion Mass Spectrometry, SIMS 8*, edited by A. Benninghoven, K. T. F. Janssen, J. Tumpner, and H. W. Werner (Wiley, Chichester, 1992), p. 307.
[36]A. Mesarwi and A. Ignatiev, J. Appl. Phys. **71**, 1943 (1992).
[37]V. K. Smirnov, S. G. Simakin, E. V. Potapov, and V. V. Makarov, Surf. Interface Anal. **24**, 469 (1996).
[38]K. Wittmaack, Surf. Interface Anal. **24**, 389 (1996).
[39]M. Meuris, W. Vandervorst, and J. Jackman, J. Vac. Sci. Technol. A **9**, 1482 (1991).

Novel approach to atomic force lithography

S. Hu,[a] S. Altmeyer, A. Hamidi, B. Spangenberg, and H. Kurz
Institute of Semiconductor Electronics II, Sommerfeldstrasse 24, 52074 Aachen, Germany

(Received 7 January 1998; accepted 22 May 1998)

A two-layer resist system consisting of 3 nm titanium and 65 nm polymethylmethacrylate (PMMA) has been developed to expand the potential of mechanical atomic force microscope nanolithography. Approximately 20 nm wide structures have been grooved in an ultrathin Ti film. The realized Ti patterns were transferred into the PMMA bottom layer by an reactive ion etching (RIE) process in oxygen. Finally, 30 nm wide and 45 nm deep grooves, and arrays of holes with a period of 55 nm, which are 20–30 nm in diameter and 45 nm in depth, have been fabricated in silicon by a second RIE step in a SF_6+O_2 gas mixture. © *1998 American Vacuum Society.* [S0734-211X(98)10004-5]

I. INTRODUCTION

The observation of quantum effects in mesoscopic systems, such as conductance quantization[1] or Coulomb blockade,[2] requires structures on nanometer scale. The fabrication of single electron transistors (SETs) working at elevated temperatures[3–7] is an especially major challenge, because nanostructures far below 100 nm must be prepared. Various pattern definition techniques such as electron-beam lithography, x-ray lithography, scanning probe microscope (SPM) lithography and nanoimprint lithography are used to fabricate these nanostructures. Although 10 nm scale fabrication has been demonstrated with both e-beam lithography[8] and nanoimprint technique,[9] it is well known that such methods are still under development and have not reached maturity yet. In addition, both techniques are very expensive. SPM has provided a very suggestive and simple alternative for nanofabrication because of its potential to reach a resolution that may exceed the standard of other lithography processes. In this area of research a variety of techniques has been developed, including the local oxidation of silicon and metals,[4,10,11] a super fine e-beam lithography using the tip of a scanning tunneling microscope (STM) as a low energy electron source,[12–15] the direct movement of atoms and molecules[16,17] as well as the direct mechanical modification of structures.[18–21]

The proximal mechanical interaction and near field electrical interaction between the tips and samples require the use of ultrathin resists or resistless processes to obtain a high resolution in SPM lithography. As a resistless process the local oxidation of metals has shown to be most successful for the fabrication of metallic SETs based on titanium and niobium.[4] Although very fine silicon oxide structures have been fabricated by SPM oxidation of silicon, it is very difficult to transfer these structures into the underlying silicon because of the low resistance of such fine oxide structures to wet KOH etching and appropriate reactive ion etching (RIE).[10,11] The main difficulty in SPM lithography is the inherently low resistance of ultrathin resists to the following dry or wet etching process, which causes problems with the pattern transfer. In the last few years mechanical patterning of materials has been done mostly by an atomic force microscope (AFM) followed by a pattern transfer into underlying bottom resist layers.[18–20] The advantages of this technique are: (1) There is no dependence on the electrical and chemical properties of the materials. (2) The lateral scale of the structures depends only on the geometry of the tip, so there is no proximity effect as in e-beam lithography. (3) It provides the possibility to pattern various materials using a multilayer resist system and anisotropic etching processes. This technique has also been used to pattern III-V semiconductors directly.[18,21]

In principle the mechanical AFM lithography is limited by several technical difficulties. First, the tip is nanometer sized at the very top only and does not have rectangular sidewalls. As a consequence the structures become inevitably large, if the tip moves deeply into the resist. Second, the cantilever can bear very small forces only. Therefore usually very thin (5–10 nm) and soft photoresists are used for the mechanical AFM lithography. The resist in the structured areas has to be removed completely by the cantilever in order to transfer the pattern precisely into the underlying layers. To ensure a complete removal of the resist in spite of a certain surface roughness, which is inevitably present, the tip must be forced into the resist to a depth exceeding the thickness of the resist. If the underlying layer is very hard, however, the tip suffers damage resulting in increased linewidths. In order to avoid such damage, thicker resists are sometimes used.[20] In such a case, however, the residual resist has to be removed from the bottom of the patterned grooves by an etching process, which in turn deteriorates the lithographic precision. Third, upheaping material along the edge of the patterned structures limits their minimum spacing.

II. EXPERIMENT AND RESULTS

In this article we report on a novel concept for the mechanical AFM lithography: The small forces applicable to the cantilever allow the modification of either a soft material[18–21] or of an extremely thin film of a hard material. The latter is the focus of our work. With this idea in mind we prepared a bilayer resist system consisting of a 3 nm thick Ti top layer and a 65 nm polymethylmethacrylate (PMMA) bottom layer on silicon, as shown in Fig. 1. The ultrathin Ti

[a]Electronic mail: hu@iht-ii.rwth-aachen.de

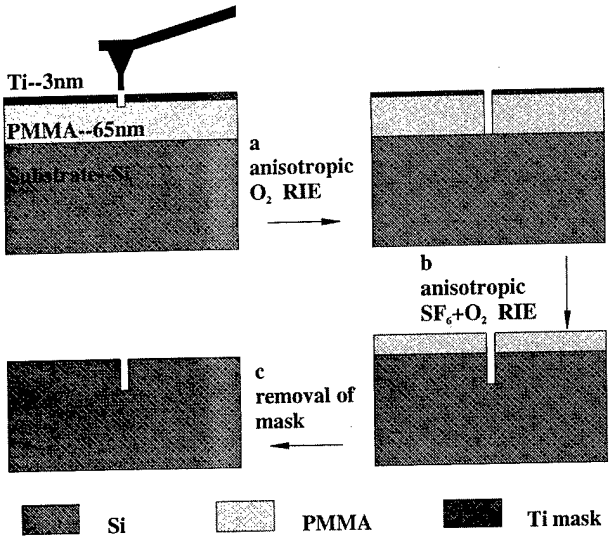


FIG. 1. Process sequence of mechanical AFM lithography and pattern transfer.

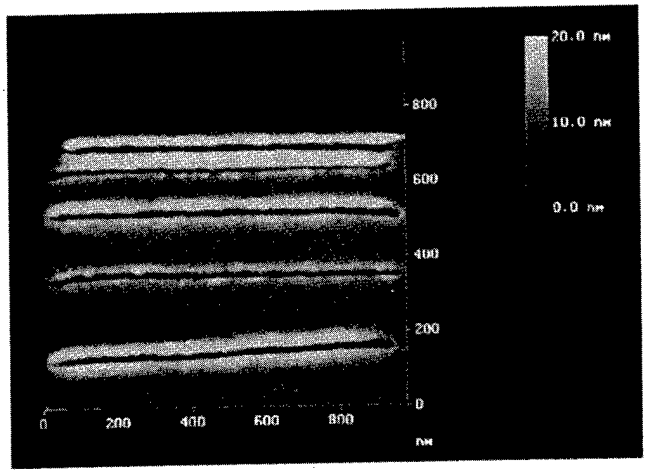


FIG. 2. AFM image after the patterning of the 3 nm Ti image layer.

film, which is more resistive to RIE than most other materials, was structured with conventional Si cantilevers using a Multimode SPM of Digital Instruments with a Nanoscope IIIa controller in air ambient. Since the underlying PMMA bottom layer is very soft, the sharpness of the cantilevers was not affected. The difficulties caused by the mechanical hardness of the resist and the geometry of the cantilevers were overcome very well by applying only a small force yielding a small penetration depth. The structures in the Ti image layer were transferred into the underlying PMMA by an anisotropic RIE process in oxygen and subsequently into the silicon substrate by an SF_6+O_2 RIE process. The complete technological approach for this subtractive pattern transfer is illustrated in Fig. 1.

First, gold marks were patterned on silicon substrates by optical lithography for a proper alignment of the AFM. The thickness of the gold marks was chosen to 30 nm to make sure that the cantilever is not damaged during the alignment procedure. In a second step, 65 nm PMMA was spun on the wafer and baked at a temperature of 165 °C for 35 min. Finally, a 3 nm Ti film was deposited by e-gun evaporation at room temperature. A deposition rate of 0.2 Å per second was used to achieve smooth and continuous films. AFM investigations of the Ti films confirm a good film quality with a root mean square roughness of 0.4 nm.

In order to avoid lateral forces damaging the cantilever, the AFM was used in tapping mode with a conventional Si tip, which has a length of 125 μm and an elastic constant of 20–100 N/m. After the sample alignment the pattern field was scanned precisely as a reference. With respect to this reference the tip was pushed through the Ti film and slightly into the PMMA by applying a vertical force of several μN. The speed of lithography was typically less than 200 nm per minute, which is much lower than in e-beam lithography. After modifying twenty grooves with a length of 1 μm, no change in the shape of the cantilever was observed. Figure 2 shows the AFM image of a sequence of lines whose distance was decreased from 200 to 50 nm in 50 nm steps, to find out the resolution. The picture was taken immediately after patterning the Ti film with the same cantilever in tapping mode. The material heaped up from patterning the last line covered partially the second last and disrupted it. As a consequence this damaged mask will be transferred into the underlying resist and finally into the substrate.

After AFM patterning, the nanostructures primarily grooved in Ti were transferred into the PMMA. To obtain steep sidewalls an anisotropic oxygen RIE process was used. For structuring 65 nm PMMA etching for 60 s is sufficient, but an overetching for 30 s has been performed to ensure a complete removal of the PMMA in such narrow grooves. In a series of tests it has been confirmed that the 3 nm Ti mask sustains the exposure to the oxygen plasma and only half of the Ti mask is removed by the RIE process. With this Ti +PMMA bilayer mask the nanopatterns were subsequently transferred into Si by a second anisotropic RIE process using SF_6+O_2 in a 85:15 mixture. Finally, the residual PMMA was removed in an isotropic high pressure oxygen plasma. In order to observe the exact influence of the individual pattern transfer processes, an AFM analysis was made after each step. Since the conventional cantilevers, which were used for structuring, are not fine enough to scan the grooves down to the bottom, sharpened e-beam deposited (EBD) tips supplied by Nanotools GmbH have been used alternatively to analyze the grooves. In Fig. 3(a) the AFM image of the completely structured Ti+PMMA bilayer mask is shown. The trenches are approximately 22 nm wide. Even with the sharpened tips the trenches appear differently in depth in the cross section profiles: The wider the trench, the higher the value measured for the depth. Because of these tip artifacts, the linewidth is measured at the position of 2/3 width and not, as usually done, at half height. In Fig. 3(b) the structures are shown after the second RIE (SF_6+O_2) process, and the trenches have a width of approximately 24 nm. The comparison of Figs. 3(a) and 3(b) allows the conclusion, that after the

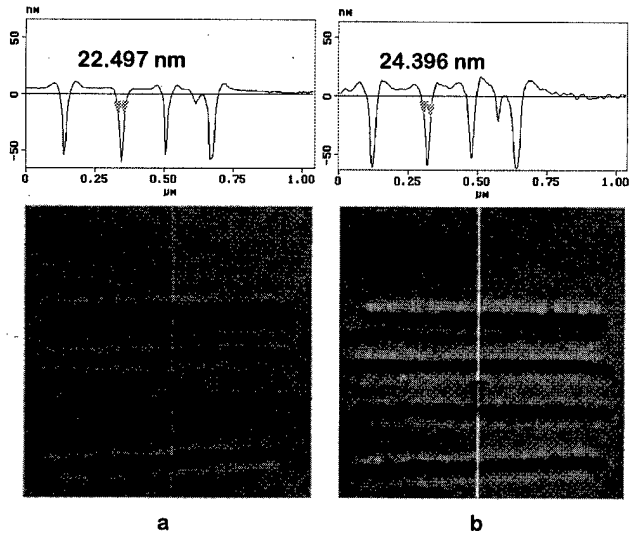


FIG. 3. AFM images of (a) structured PMMA and (b) structured Si before removing the residual PMMA.

SF_6+O_2 RIE process the residual PMMA has not suffered a significant mask shift. This proves the high anisotropy of our O_2 RIE process. Figure 4 shows the AFM image of the structures in the Si substrate after the removal of the residual PMMA. The grooves are approximately 30 nm wide and 45 nm deep. The comparison of Fig. 3(b) and Fig. 4 shows that the Si structures widened approximately 10 nm during the SF_6+O_2 RIE. In order to characterize the grooves better, an inverted three-dimensional AFM image is shown in Fig. 5, and the structures are very uniform except for the second last one: The disruption of the Ti mask, caused by upheaped material, is exactly transferred also into the Si substrate. Finally, to confirm the AFM images and linewidth measurements, a high resolution scanning electron microscope image of the structures was taken (Fig. 6). The comparison of Fig. 6 and Fig. 4 reveals, that the AFM analyses are nearly free of tip artifacts. The minimal period for reproducible lines and spaces was found to be approximately 80 nm.

With the same process described above, arrays of holes were structured in silicon. Figure 7 shows such a pattern. The period is 55 nm, the diameter is 20–40 nm and the depth is 45 nm. For better visibility, this image is inverted also, i.e., the holes appear as dots.

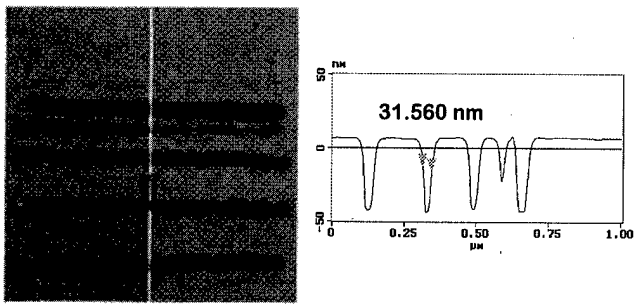


FIG. 4. AFM image of the grooves in the Si substrate. The width is approximately 30 nm and the depth is 45 nm.

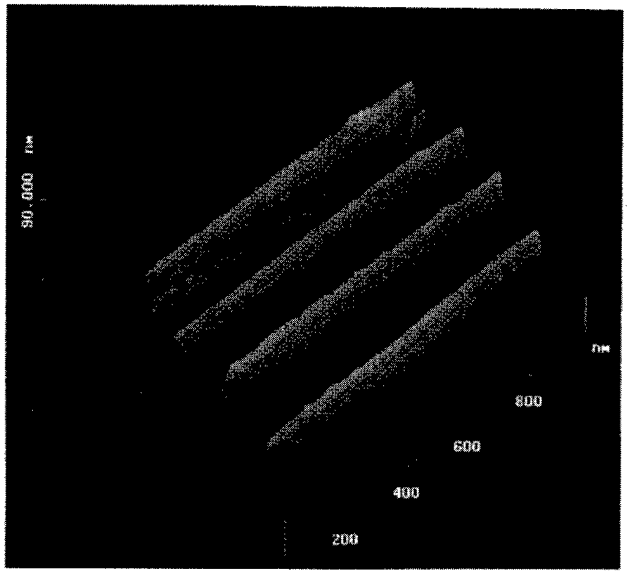


FIG. 5. Three-dimensional AFM image of grooves in Si. For better visibility of the shape of the grooves, the AFM image was inverted, i.e., the grooves appear as wires.

With the aim of patterning smaller structures, we have also tried to modify the resist with sharper non-standard cantilevers, but it was not successful. The super sharp cantilevers were either bent or worn, so that the results did not differ very much from those achieved with conventional tips.

III. CONCLUSION

Using a resist system consisting of 3 nm Ti and 65 nm PMMA a promising progress was brought to mechanical AFM lithography. Grooves of approximately 30 nm width and 45 nm depth, and array of holes with a period of 55 nm, diameters of 20–40 nm and a depth of 45 nm have been achieved in silicon. This technique may be used to pattern other materials on the same scale, if suitable mask materials with respect to the substrate are chosen and an appropriate

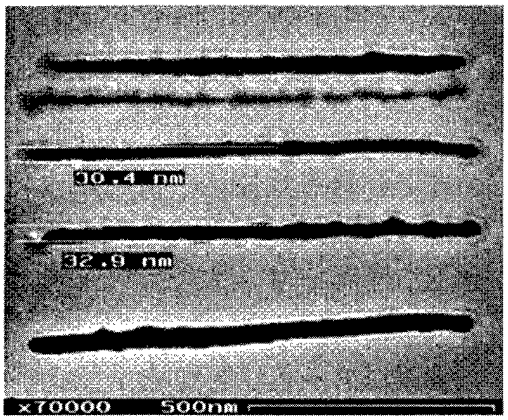


FIG. 6. High resolution SEM picture of the grooves in Si.

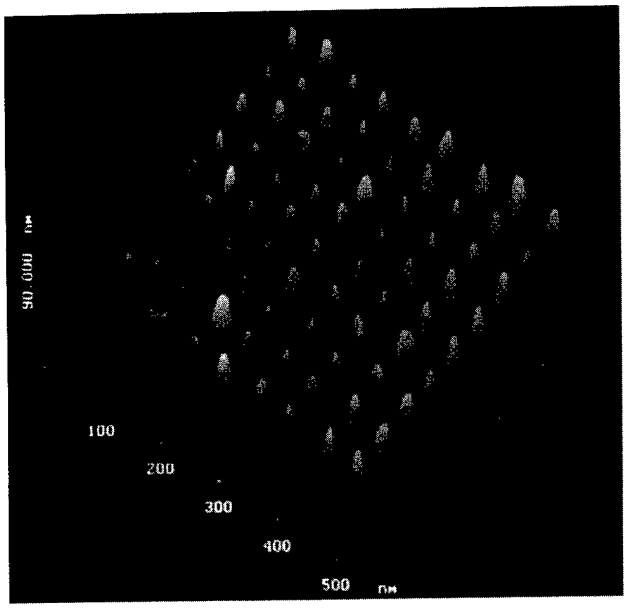


FIG. 7. Three-dimensional AFM image of an array of holes in Si with a period of 55 nm and a depth of 45 nm. The diameter of the holes is 20–40 nm. For better visibility of the shape of the holes, the AFM image was inverted, i.e., the holes appear as dots.

RIE process is applied. The application of this technique for the fabrication of metallic or semiconductor based SETs is under development.

ACKNOWLEDGMENTS

One of the authors (S. Hu) would like to thank the German Academic Exchange Service (DAAD) for a Ph.D student fellowship. The authors would also like to thank B. Vratzov and C. Moormann in the Advanced Microelectronics Center Aachen (AMICA) for the high resolution scanning electron microscope images. This work was supported by the BMBF under Contract No. 13N6256 and No. 13N7166.

[1]R. A. Smith and H. Ahmed, J. Appl. Phys. **81**, 2699 (1997).
[2]D. V. Averin and K. K. Likharev, J. Low Temp. Phys. **62**, 345 (1986).
[3]K. Kurihara, H. Namatsu, M. Nagase, and T. Makino, Microelectron. Eng. **35**, 261 (1997).
[4]K. Matsumoto, Proc. IEEE **85**, 612 (1997).
[5]L. Guo, E. Leobandung, and S. Y. Chou, Appl. Phys. Lett. **70**, 850 (1997).
[6]S. Altmeyer, A. Hamidi, B. Spangenberg, and H. Kurz, J. Appl. Phys. **81**, 8818 (1997).
[7]T. Köster, B. Hadam, K. Hofmann, J. Gondermann, J. Stein, S. Altmeyer, B. Spangenberg, and H. Kurz, J. Vac. Sci. Technol. B **15**, 2836 (1997).
[8]K. Kurihara, K. Iwadate, H. Namatsu, M. Nafase, H. Takenaka, and K. Murase, Jpn. J. Appl. Phys., Part 1 **34**, 6940 (1995).
[9]S. Y. Chou and P. R. Krauss, Microelectron. Eng. **35**, 237 (1997).
[10]P. Avouris, T. Hertel, and R. Martel, Appl. Phys. Lett. **71**, 285 (1997).
[11]J. Servat, P. Gorostiza, and F. Sanz, J. Vac. Sci. Technol. A **14**, 1208 (1996).
[12]A. Majumdar, P. I. Oden, J. P. Carrejo, L. A. Nagahara, J. J. Graham, and J. Alexander, Appl. Phys. Lett. **61**, 2239 (1992).
[13]S. C. Minne, Ph. Flueckiger, H. T. Soh, and C. F. Quate, J. Vac. Sci. Technol. B **13**, 1380 (1995).
[14]M. Amman, J. W. Sleight, D. R. Lombardi, R. E. Welser, M. R. Deshpande, M. A. Reed, and L. J. Guido, J. Vac. Sci. Technol. B **14**, 54 (1996).
[15]M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **6**, 293 (1988).
[16]L. Bartels, G. Meyer, and K.-H. Rieder, Phys. Rev. Lett. **79**, 697 (1997).
[17]A. A. Gorbunov, J. Richter, W. Pompe, A. Sewing, and J. Thomas, Surf. Interface Anal. **25**, 596 (1997).
[18]M. Wendel, B. Irmer, J. Cortes, R. Kaiser, H. Lorenz, J. P. Kotthaus, A. Lorke, and E. Williams, Superlattice Microstruct. **20**, 349 (1996).
[19]L. L. Sohn and R. L. Willett, Appl. Phys. Lett. **67**, 1552 (1995).
[20]V. Bouchiat and D. Esteve, Appl. Phys. Lett. **69**, 3098 (1996).
[21]R. Magno and B. R. Bennett, Appl. Phys. Lett. **70**, 1855 (1997).

Dry-etching development characteristics of $Se_{75}Ge_{25}$ resist for focused-ion-beam lithography

Hyun-Young Lee and Hong-Bay Chung[a)]
Department of Electronic Materials Engineering and Institute of New Technology, Kwangwoon University, 447-1, Wolgyedong, Nowonku, Seoul 139-701, Korea

(Received 10 October 1997; accepted 17 April 1998)

We have investigated the focused-ion-beam (FIB) lithographic mechanism of $Se_{75}Ge_{25}$ resist exhibiting dual-type resist characteristics according to development methods. That is the Ga-FIB exposed $Se_{75}Ge_{25}$ films act as positive and negative type resists for wet-etching development using a solution of HNO_3:HCl:H_2O (1:1:3) and dry development by CF_4 reactive-ion etching (RIE), respectively. With increasing incident energy, both sensitivity S and contrast Υ are enhanced in the case of wet development, while both are deteriorated in the case of dry development. Considering the deposited energy density and implanted ion-concentration distributions obtained by our Monte Carlo simulation, FIB lithography for wet development depends primarily on the deposited energy and for dry development is dominated by the implanted ions themselves. For a 580 Å thick $Se_{75}Ge_{25}$ resist exposed by 30 keV Ga^+ FIB and developed by CF_4 RIE for 5 s under the gas pressure of 200 mTorr, S and Υ are about 1.7×10^{15} ions/cm^2 and 4.44, respectively. In this case the ion concentration exceeds approximately 3.7×10^{20} ions/cm^3, Ga-exposed region is estimated to have a resistance against CF_4-reactive gas. A 0.125 μm line pattern, with very small proximity effect, has been successfully fabricated by RIE development after 40 keV FIB exposure with a diameter of 0.120 μm. © *1998 American Vacuum Society.* [S0734-211X(98)02604-3]

I. INTRODUCTION

High resolution lithography is a key technique in the fabrication of sub 0.1 μm-scale structures and devices.[1,2] Focused-ion-beam (FIB) has been developed for high-resolution lithography and utilized conventional organic polymers as resist materials. However, it is difficult to obtain a fine pattern using organic resists, because of inherent volatile solvents and water as well as the limited capabilities of contrast and thermal property.[3]

We have recently reported[4–6] the results of three-dimensional (3D) Monte Carlo (MC) simulation for Ga^+ ion penetration in an amorphous germanium selenide thin film and pattern profiling of FIB lithography using the simplified exposure and development models, in which the minimum resist thickness Z_{min} is selected to minimize a substrate damage and absorb the ion beam sufficiently at the same time. Also, in previous papers,[7] we presented the experimental results containing ion-induced changes, exposure response characteristics, and development patterns of positive-type $Se_{75}Ge_{25}$ resist by relatively low-energy FIB exposure and wet etching development, and then concluded that the FIB lithography depended strongly on the energy density deposited by the collisions of energetic ions rather than the implanted ions themselves.

The constituents of $Se_{75}Ge_{25}$ resist react upon F-doped gases and form volatile compounds, such as SeF_6 and GeF_4. Nonvolatile compound GaF_x is formed in the surface of Ga^+-exposed region and eventually this film acts as a negative type resist. Therefore, FIB lithography mechanism by CF_4 dry-etching development could be expected to differ from that by wet development.

In order to understand the FIB lithographic mechanism of $Se_{75}Ge_{25}$ resist exhibiting dual-type resist characteristics according to development methods, we investigate here the exposure response and patterning characteristics of the resist by CF_4 dry-etching development, compare to those by wet-etching developments, and analyze using the modified Lindhard-Scharff-Schiott (LSS) theory[8,9] and our MC simulation.[5,6]

II. EXPERIMENT

The use of thinner ion-beam resist leads to the fine pattern formation.[10] It is necessary to select an optimum resist thickness to absorb the ion energy or the ion itself sufficiently and to minimize ion penetration into the substrate at the same time.

In previous papers,[5,7] we presented the minimum resist thickness Z_{min} and the ion range parameters as a function of incident energy using MC simulation, and then the ranges of Z_{min} were 546–590 Å and 685–717 Å for 30 and 40 keV Ga^+-FIBs, respectively. Most stop ions were distributed relatively uniformly in the entire depth of the resist and a few ions per 10^4 only penetrated into the substrate. If the thickness of $Se_{75}Ge_{25}$ resist is much more than the Z_{min} and dry development is performed under isotropic etching conditions, the undercutting from the lower part of the resist appears easily since the nonvolatile layer is formed within Z_{min} , i.e., very shallow regions of the $Se_{75}Ge_{25}$ resist. Therefore, the Z_{min} condition is necessary for obtaining the minimum substrate damage and performing reactive-ion etching (RIE) development without undercutting—independently on

a)Electronic mail: hbchung@daisy.kwangwoon.ac.kr

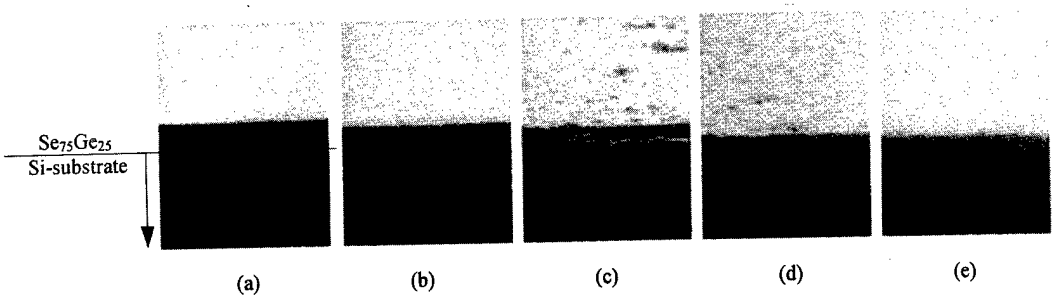


FIG. 1. SEM cross sections of unexposed $Se_{75}Ga_{25}$ resist as a function of RIE-development time: (a) as-deposited, (b) 3 s, (c) 6 s, (d) 10 s, and (e) 12 s. Thickness of as-deposited thin film is about 1000 Å. The RIE was carried out under the CF_4 gas pressure of 200 mTorr and the rf power of 200 W.

the pressure of reactive gases—as well. In our study, the density of a-$Se_{75}Ge_{25}$ is 4.42 g/cm^3, which corresponds to 3.39×10^{22} molecules/cm^3 of the atomic density.

A $Se_{75}Ge_{25}$ bulk glass was prepared by a conventional melt quenching technique. The constituent elements Se and Ge weighed in the appropriate ratio were sealed in an evacuated quartz ampoule, which was then placed in a furnace and heated at 220, 650, and 1000 °C for 2, 2, and 24 h, respectively. The ampoule was constantly stirred during these periods to achieve complete homogenization of the constituents in the melt and quenched in water. An amorphous $Se_{75}Ge_{25}$ thin film was prepared, on p-type Si (100) kept at room temperature, by thermal evaporation of the bulk in vacuum $\sim1\times10^{-5}$ Torr. The deposition rate of films was about 5 Å/s. Films were deposited in four thickness types: less than the Z_{min}, the Z_{min} and more than the Z_{min} to observe the thickness dependence, and a 1000 Å thick selected to evaluate the etching rate R_0 of unexposed resist. The thicknesses prepared for 30 and 40 keV-Z_{min} conditions were about 580 and 700 Å, respectively.

Ga^+ ion exposure was performed using the FIB machine (SMI-8100) at 30 and 40 keV incident energies and then the exposure dose was controlled by changing the exposing time. Note 30 and 40 keV FIB multiscanning with squares of dimension 425 μm$\times$0.20 μm and exposing time of 5–50 s caused the dose range of 7.3×10^{14}–7.3×10^{15} ions/cm^2 for a beam current of 20 pA. Also, 40 keV FIB exposure was performed by a single-raster scanning with the beam diameter of 0.12 μm.

The development before and after the exposure of the resist was carried out using RIE, where the CF_4 gas pressure and the radio frequency (rf) power were about 140–200 mTorr and 200 W, respectively. The development time, T_{RIE}, to develop completely the unexposed resist is determined by Z/R_0, where Z is the fabricated resist thickness. Fabricated linewidth and remaining height after RIE development were measured using a profilometer and a scanning electron microscope (SEM) with which 40°-tilted lithographic images were observed.

III. RESULTS AND DISCUSSION

RIE-development characteristics of unexposed 1000 Å thick $Se_{75}Ge_{25}$ resist are investigated in order to select the optimized development time, T_{RIE}, for various thicknesses and the SEM cross sections as a function of etching time are shown in Fig. 1, where relatively high CF_4 gas pressure of 200 mTorr is used. The resist is very sensitive to CF_4-reactive gases and the entire thickness is completely etched off for a very short time. Also, the surface and the resist / substrate interface of as-deposited film seem to be very uniform and the surface of the resist developed for 6 s appears relatively rough.

The selective etching characteristics of 1000 Å thick $Se_{75}Ge_{25}$ resist exposed to Ga^+ doses of 1.46×10^{15} and 2.19×10^{15} ions/cm^2 at 30 keV, after CF_4-RIE development, is shown in Fig. 2.

The etching rate of unexposed area R_0 and that of exposed area R— a dose of 1.46×10^{15} ions/cm^2 are approximately 80.0 and 72.0 Å/s, respectively. The R of $Se_{75}Ge_{25}$ resist exposed to 2.19×10^{15} ions/cm^2 are separated into two cases; $R=\sim36.7$ Å/s for 0–3 s and $R=\sim2.5$ Å/s up to 9 s.

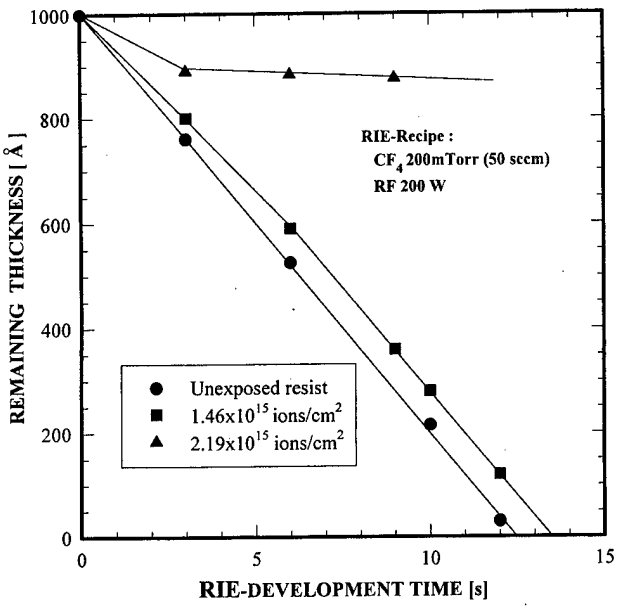


FIG. 2. Etching characteristics of $Se_{75}Ga_{25}$ thin film unexposed and exposed to doses of 1.146 and 2.19×10^{15} ions/cm^2 at 30 keV and developed under the CF_4 pressure of 200 mTorr. The etching rate of unexposed resist R_0 is approximately 80.0 Å/s.

That is, R of the resist exposed to a dose of 1.46×10^{15} ions/cm^2 is similar to R_0 and is much larger than that exposed to 2.19×10^{15} ions/cm^2. This result means that a critical dose lies between two doses. In the resist exposed to 2.19×10^{15} ions/cm^2, particularly, though the undercutting in the lower part of the resist pattern is produced, we have not observed any more thickness reduction by etching—only ~115 Å thickness is dry etched for this dose, which can be well explained considering the concentration distribution of Ga ions penetrated into the $Se_{75}Ge_{25}$ resist. The implanted Ga concentration exceeds a critical level N_c in this shallow region, the implanted Ga ions themselves react with F-doped reactive gases and eventually a protection layer with sufficient nonvolatile compound GaF_x is formed. But in the case of a dose of 1.46×10^{15} ions/cm^2, R become very large because the peak concentration dose not even reach the critical level.

Figure 3(a) shows the implanted ion concentration distributions obtained by LSS theory and MC simulation for the combination of $Se_{75}Ge_{25}$ resist and 30 keV Ga^+ ion beam and Fig. 3(b) is a three-dimensional ion distribution for a dose of 4.38×10^{15} ions/cm^2, where the projected range R_p is about 190 and 210 Å for MC and LSS results, respectively.[7]

From Fig. 3(a), therefore, the ion concentration at depth of 115 Å for a dose of 2.19×10^{15} ions/cm^2 is evaluated as the N_c for CF_4 RIE development, which is approximately 3.7×10^{20} ions/cm^3.

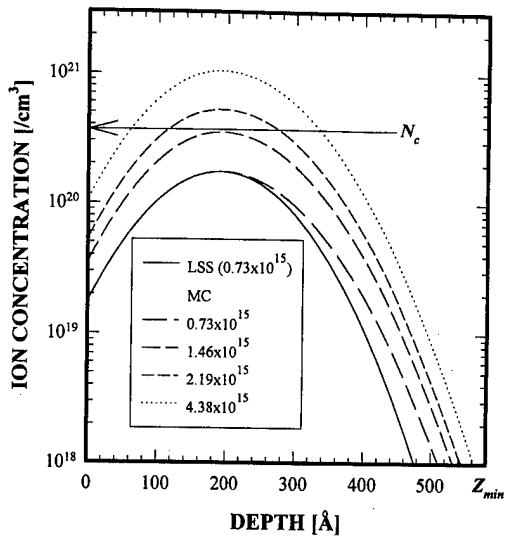
Generally, etching under relatively high CF_4 gas pressure leads to isotropic resist profiles with the undercutting. We in fact have observed that the undercutting was produced under the gas pressure of 140 mTorr in the case of a 1000 Å thick $Se_{75}Ge_{25}$ exposed to 2.19×10^{15} ions/cm^2. For the Z_{min} thick resist, however, this undercutting is not observed for CF_4 gas pressure of 140–200 mTorr.

The three-dimensional MC-simulated profiles to be fabricated after CF_4 RIE development for the $Se_{75}Ge_{25}$ resist exposed to a dose of 4.38×10^{15} ions/cm^2 are shown in Fig. 4, in which the three 30 keV FIBs with 0.1 μm space, 0.2 μm diam and 1.2 μm scan length,[6] and the development recipes in Fig. 1 are utilized. The resist thickness and the development time are $Z_{min}(=580)$ Å and $T_{RIE}(=Z_{min}/R_0=7.25$ s), respectively.

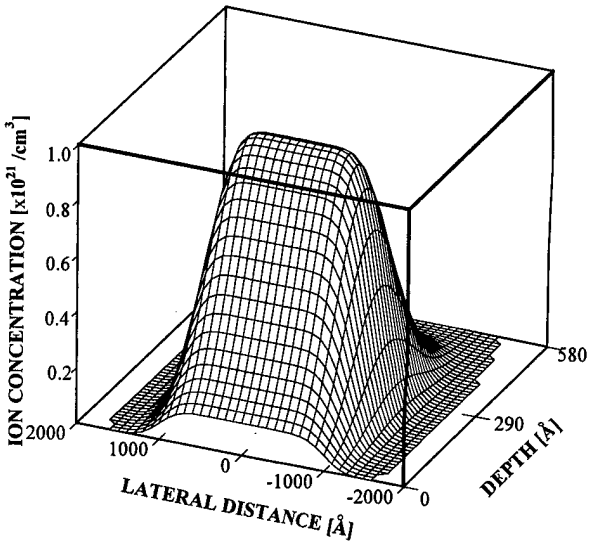
The simulated step height and linewidth are 520 Å and 0.205 μm, respectively. Because the lateral spreading is about 250 Å for both left and right directions, the proximity effect under the above-mentioned conditions can be ignored completely.

Figure 5 shows SEM photographs of lithographic patterns fabricated in 580 Å thick $Se_{75}Ge_{25}$ resists, where exposure was performed by 30 keV Ga^+ FIB with a multiscan beam diameter of 0.20 μm and RIE development was carried out for 5 s under the CF_4 gas pressure of 200 mTorr and the rf power of 200 W.

The pattern is formed initially at the dose range of about $0.73-1.46\times10^{15}$ ions/cm^2 and a linewidth corresponding to the beam diameter is obtained at the range of $219-2.92\times10^{15}$ ions/cm^2. While the linewidth increases with increasing ion dose for the most resists—Si_3N_4,[3] WO_3,[11] PMMA,[12] etc., this resist maintains almost constant linewidth for a dose above 2.19×10^{15} ions/cm^2. The pattern of Fig. 5(e) is in good agreement with the MC result of Fig. 4, although development time is different from each other.



(a)



(b)

FIG. 3. (a) Two-dimensional ion-concentration distributions as a function of dose and (b) three-dimensional 4.38×10^{15} ions/cm^2 obtained by MC simulation for the combination of $Se_{75}Ge_{25}$ resist and 30 keV Ga^+ ion beam. A solid line of (a) presents a symmetrical Gaussian-profiling obtained by LSS-based calculation.

Figure 6 shows typical exposure response curve of Z_{min} thick $Se_{75}Ge_{25}$ resist exposed to 30 Ga^+ FIB. As seen from this curve, sensitivity S, i.e., the threshold ion dose required to correspond to a 50% remaining thickness and the imaging contrast γ, i.e., the slope around the threshold dose, are about 1.7×10^{15} ions/cm^2 and 4.44, respectively. In the case of 40 keV, we have obtained S and γ of 2.1×10^{15} ions/cm^2 and 4.0, respectively.

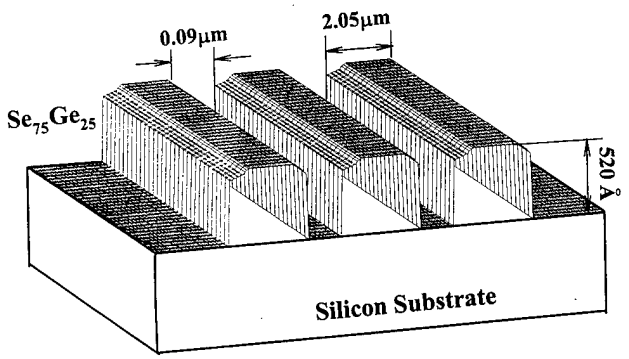


FIG. 4. Three-dimensional MC-simulated patterns of Z_{min} thick $Se_{75}Ge_{25}$ resist exposed by Ga^+-Gaussian FIB and developed by CF_4 RIE for T_{RIE}(FIB exposure conditions: dose=4.38×10^{15} ions/cm^2, incident energy =30 keV, diameter=0.2 μm, spacing=0.1, μm, scan length=1.2 μm).

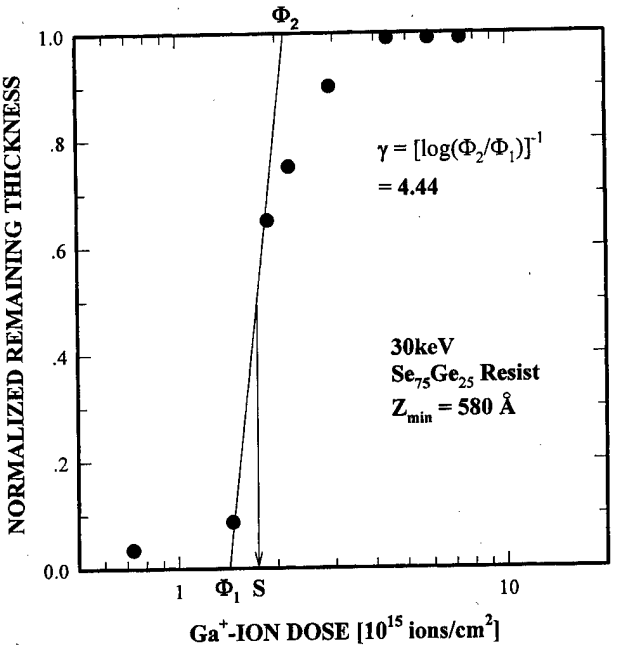


FIG. 6. Exposure response curve of Z_{min}-thick $Se_{75}Ge_{25}$ resist for 30 keV Ga^+ FIB exposure and CF_4 RIE-development for 5 s, where normalized remaining thickness means pattern height/($R_0\times5$).

As the incident energy increases from 30 to 40 keV, both S and γ are deteriorated. As is remarked in Figs. 2 and 3, this reason is because FIB lithography accompanied by CF_4 RIE development depends strongly on the implanted ions themselves rather than the energy density deposited by the collisions of energetic ions. If the deposited energy density is a predominant factor in the exposure characteristics, like a previous paper about wet-etching development,[7] the S should be enhanced with increasing incident energy. Especially, the reduction of can be easily understood from the increase of range straggling ΔR_p with increasing incident energy.

Therefore, that $Se_{75}Ge_{25}$ resist exhibits dual-type resist characteristics — a negative type for dry development and a positive type for wet development — is due to the difference of FIB lithographic mechanism.

Figure 7 shows a SEM photograph of the pattern fabricated in Z_{min}(700 Å) thick $Se_{75}Ge_{25}$ resist obtained from 40 keV-FIB exposure (4.3×10^{15} ions/cm^2) with a single raster-scanning diameter of 0.120 μm and RIE development for 5 s under the gas pressure of 200 mTorr. The fabricated linewidth is estimated to be about 0.125; fabricated height was about 370 Å. A rough surface beside the development line is due to employing insufficient development time (~5 s) for which approximately 300 Å thick $Se_{75}Ge_{25}$ remains in this unexposed area [see Fig. 1(c)]. It can be eliminated if sufficient development time (~10 s in this case) is employed.

IV. CONCLUSION

The FIB lithographic mechanism of $Se_{75}Ge_{25}$ resist exhibiting dual-type resist characteristics according to development methods has been investigated. The Ga-FIB exposed $Se_{75}Ge_{25}$ films act as positive and negative-type resists for wet-etching development, using a solution of HNO_3:HCl:H_2O (1:1:3) and dry development by CF_4 RIE, respectively. With increasing incident energy, both sensitivity S and contrast Υ are deteriorated in the case of dry development. Through MC simulation and experiment procedure,

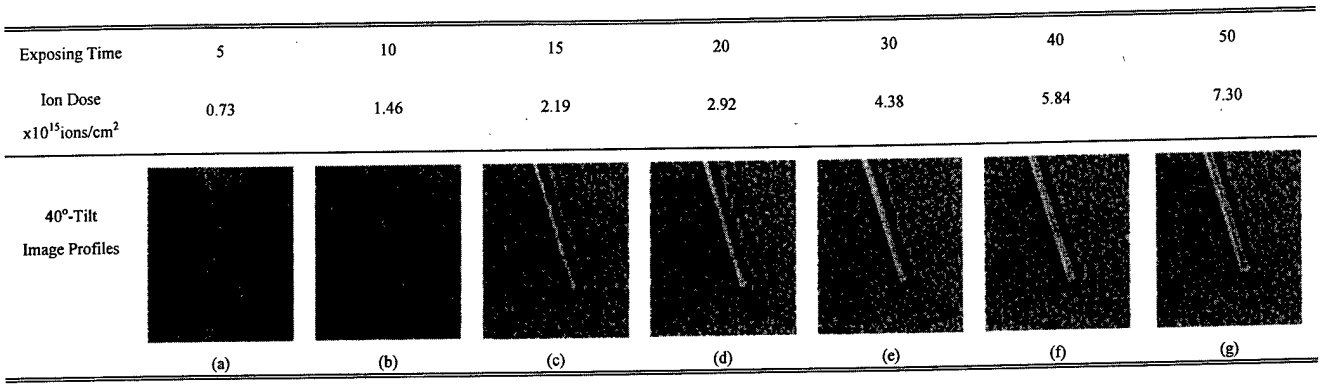


FIG. 5. 40°-tilted SEM image profiles of lithographic patterns fabricated in 580 Å thick $Se_{75}Ge_{25}$ resists. FIB exposure conditions are an incident energy of 30 keV, a multiscan-beam diameter of 0.20 μm, scan length of 425 μm, and beam current of 20 pA, and RIE development conditions are the CF_4 gas pressure of 200 mTorr, the rf power of 200 W, and development time of 5 s which etches off about 400 Å thick unexposed-resist region. Linewidths of (a), (c), (e), and g) are 0.050, 0.20, 0.205, and 0.205, respectively. Step height of (e) is evaluated to be about 340 Å.

FIG. 7. SEM photographs of the pattern fabricated in 700 Å thick $Se_{75}Ge_{25}$ resists. FIB exposure is performed by an incident energy of 40 keV with a single raster scanning of a diameter 0.120 μm and RIE development is performed for 5 s under the gas pressure of 200 mTorr.

FIB lithography for wet development depends primarily on the deposited energy and for dry development is dominated by the implanted ions themselves. For the Z_{min} thick $Se_{75}Ge_{25}$ resists exposed by 30 keV Ga^+-FIB and developed by CF_4 RIE under the gas pressure of 200 mTorr, S and Y are about 1.7×10^{15} ions/cm^2 and 4.44, respectively. In this case the ion concentration exceeds approximately 3.7×10^{20} ions/cm^3 — the critical ion concentration — Ga-exposed region is estimated to have a resistance against CF_4-reactive gas. A 0.125 μm line pattern, with very small proximity effect, is successfully fabricated by RIE development after 40 keV FIB exposure with a diameter of 0.120 μm. Since the resolution of this lithography is a little larger than the beam diameter and lateral spreading is very small, high-contrast patterns of the sub-0.1 μm can be obtained easily from the reduction of FIB diameter and the control of dose and thickness. Though $Se_{75}Ge_{25}$ inorganic resist has several advantages such as physical stability, compatibility with *in situ* process, high contrast, and great ion-stopping power to protect a substrate from ion-induced damage, its sensitivity is 1–2 orders lower than that of conventional polymer resist. Therefore, this problem should be solved, especially in order to employ it as the resist of ion-projection lithography. To obtain the sensitivity of $\sim 10^{14}$ ions/cm^2, a study on FIB lithography characteristics using about 80°-columnar structural Se-Ge resist is now in progress.

ACKNOWLEDGMENTS

The authors are grateful to B.-J. Chang for his help on the experiment of FIB exposure. This study was supported by the academic research fund of Ministry of Education, Republic of Korea, through Inter-University Semiconductor Research Center (ISRC 96-E-4036 and ISRC 97-E-1032) in Seoul National University.

[1]I. Brodie and J. J. Muray, *The Physics of Micro/NanoFabrication* (Plenum, New York, 1992).
[2]H. I. Smith, *Submicron- and Nanometer-Structures Technology* (Nanostructures, Sudbury, 1994).
[3]N. Koshida, K. Yoshida, S. Watanuki, M. Komura, and N. Atoda, Jpn. J. Appl. Phys., Part 1 **30**, 3246 (1991).
[4]H. Y. Lee and H. B. Chung, Synth. Met. **71**, 2047 (1995).
[5]H. Y. Lee and H. B. Chung, J. Appl. Phys. **78**, 5975 (1995).
[6]H. Y. Lee and H. B. Chung, Jpn. J. Appl. Phys., Part 1 **36**, 2409 (1997).
[7]H. Y. Lee and H. B. Chung, J. Vac. Sci. Technol. B **15**, 818 (1997).
[8]J. Lindhard, M. Scharaff, and H. E. Schiott, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **33**, 14 (1963).
[9]J. F. Gibbons, W. S. Johnson, and S. W. Mylroie, *Projected Range Statistics, Semiconductors and Related Materials*, 2nd ed. (Dowden, Hutchinson, and Ross, Pennsylvania, 1975).
[10]W. M. Moreau, *Semiconductor Lithography; Principle, Practices, and Materials; Microdevice Physics and Fabrication Technologies* (Plenum, New York, 1988).
[11]S. Takahashi, M. Ohashi, S. Fukatsu, Y. Shiraki, and R. Ito, J. Vac. Sci. Technol. B **11**, 268 (1993).
[12]T. Shiokawa, Y. Aouago, P. H. Kim, K. Toyoda, and S. Namba, Jpn. J. Appl. Phys., Part 2 **23**, L232 (1984).

Blurring effect analysis of an x-ray mask for synchrotron radiation lithography

Il Yong Kim and Byung Man Kwak[a)]
Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Yusong-Gu, Taejon 305-701, Korea

Young Jin Jeon and Sang Soo Choi
Advanced Research Department, Semiconductor Technology Division, Electronics and Telecommunications Research Institute, Yusong-Gu, Taejon 305-350, Korea

(Received 12 June 1997; accepted 1 May 1998)

During the process of synchrotron radiation lithography, blurring occurs largely because the mask and patterns are distorted under transient thermal loads. An analysis model of blurring has been proposed in this article. Irradiated energy is calculated by summing the moving x-ray power incident to a wafer. The blur is caused mainly by two sources: displacement of the pattern and pattern width change. A transient analysis has been made. In a sample calculation, the distortion of a printed pattern after the process was 1.2 nm while the maximum in-plane distortion during the process was 3.5 nm and width change −3.5 nm for a 1 Hz scanning frequency. Comparison between distortions without considering blurring effect and distortions considering blurring effect was offered. The effect of scanning frequency on printed pattern distortion and width change was also presented.
© *1998 American Vacuum Society.* [S0734-211X(98)03504-5]

I. INTRODUCTION

The distortion of a membrane in an x-ray lithography is increasingly important as the integration density becomes higher. The membrane distorts during fabrication processes such as cooling or etching. It can also be distorted while it is under use. Fixturing, gravity, heating, and thermal damage by an x-ray exposure cause the mask to be distorted. Distortions by gravity and fixturing were a topic of intensive study especially by using computer simulation.

Since the width of synchrotron radiation is too small to expose the entire membrane at one time, the beam is scanned on the mask. The x-ray scanning causes the membrane to be distorted dynamically, and thus the distortions depend on the position of the beam. They were studied by using a quasi-static model,[1,2] or a dynamic model[3,4] considering only the irradiated energy. In this article, blurring due to transient thermal distortions with a resist sensitivity is simulated and the resulting printed pattern distortion on a wafer obtained.

II. MODELING AND A CASE CALCULATION

The mask model used in this study consists of a 2 μm thick Si_3N_4 membrane, a Si wafer, and a supporting Pyrex ring. The wafer is 625 μm thick and the window size 40 mm×40 mm. There is a device wafer under the mask. Energy intensity from the synchrotron is assumed to have a Gaussian distribution along the scanning direction and a uniform distribution along the direction perpendicular to it, as shown in Fig. 1. Part of the incident x-ray energy is absorbed by the membrane and absorbers, a small portion of the energy is reflected by the membrane, and the remaining portion passes through it reaching the wafer. Energy absorbed by the membrane varies according to time and position of the scanning beam. The moving beam causes the membrane to be dynamically distorted.

The x-ray power absorbed by the membrane is emitted by radiation and convected at the outer surface. A thermal conduction occurs through the mask membrane and the proximity gap is filled with helium. Previous studies show that the thermal energy transferred by the natural convection is negligibly small when the proximity gap is smaller than about 100 μm.[2,4] Therefore it was neglected in the present study. The analysis in our study has been made by the commercial code ANSYS® and the method to calculate temperature distribution and thermal distortion is similar to those in the study.[5] A two-dimensional model including x-ray mask, proximity gap, resist, device wafer, and holder was used for thermal analysis. A three-dimensional model of membrane was used to calculate thermal distortion. We used about 2000 two-dimensional plane elements for thermal analysis. A model of 60×60 four-node linear three-dimensional shell elements were used for structural analysis of the membrane.

The dynamic deformation of a membrane will bring in a dynamic pattern distortion. Figure 2 explains the proposed model of pattern blur by dynamic thermal membrane distortions. A pattern will move to the right when an incident x-ray power is on the left side, and the pattern will be printed at this position. When the peak point of the power is right on the pattern, the pattern will be printed directly below. When the power source moves further to the right, the pattern will move leftwards. The printed image on the wafer due to the distortion is therefore widened or blurred over some area. Figure 2 is for the case of a static thermal analysis. If we consider transient effects, the printed pattern will be distorted even when the power source is right above the pattern, causing blur. In this case, the blurred image will not show sym-

a)Electronic mail: bmkwak@khp.kaist.ac.kr

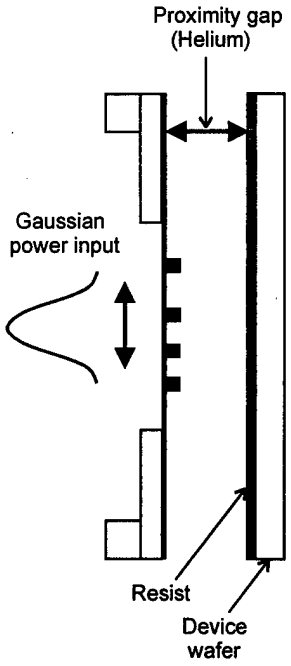


FIG. 1. Schematic of x-ray mask exposure.

metry. In general, a pattern width is about 250 nm or below, and distortion is less than a few tens of nanometers. The drawing in the figure is exaggerated to explain the blurring effects.

Energy passing through the membrane is not always used for printing on the wafer. Figure 3 shows a model of the

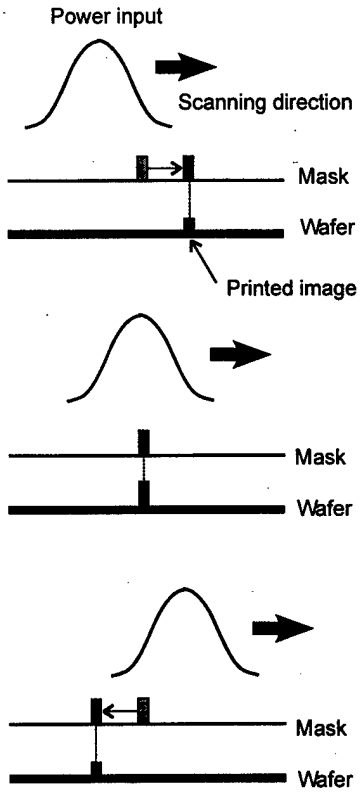


FIG. 2. Blurring process.

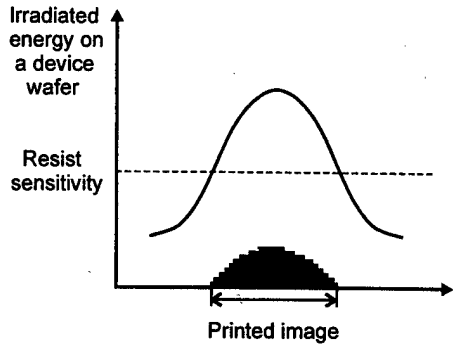


FIG. 3. Model of irradiated energy on a wafer.

irradiated energy that can go to printing. The dashed line is resist sensitivity, which is a threshold for creation of a printed pattern. The part which absorbs more energy than this amount is assumed used for printing and the rest is not. The amount of absorbed energy, its profile, and resist sensitivity are important factors which determine the amount of blur.

The blurring effect consists of two factors: a shift of the center line of a printed line image on a device wafer and a width change of the line image. The maximum shift during one scanning cycle will be denoted by IPD_B. Another quantity of interest for the computational purpose is IPD denoting the maximum displacement of an absorber on the membrane during a scanning cycle. The width change is the difference in width between an absorber on a membrane and a printed image on a device wafer. IPD_B, IPD, and a width change are illustrated in Fig. 4. When the temperature gradient is large, the absorber displacement will become large. If the energy

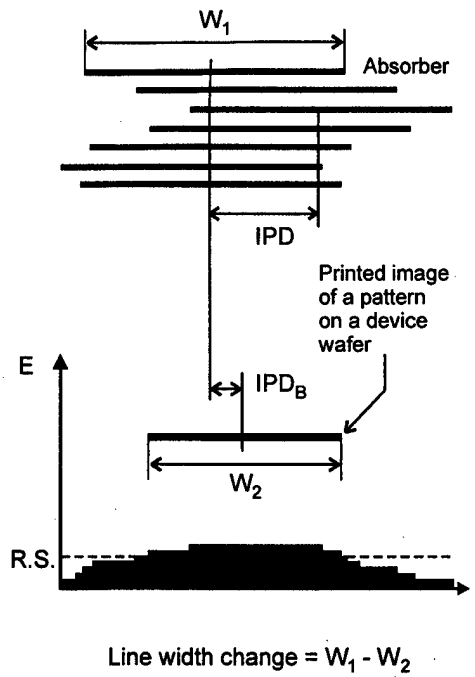


FIG. 4. Illustration of IPD, IPD_B, and width changes.

TABLE I. Base conditions of exposure system parameters.

Maximum x-ray power intensity	150 mJ/cm^2
Gaussian standard deviation	2.5 cm
Proximity gap	40 μm
Scanning frequency	1 Hz
Pattern coverage	50%

irradiated on the device wafer is higher, the lines of the printed pattern will be thicker.

III. RESULTS AND DISCUSSIONS

In this study, the scanning frequency is taken as 1 Hz and the pattern coverage, that is, the ratio of the absorber area to the whole mask area is 50%. The maximum x-ray power intensity assumed is 150 mW/cm^2 and the Gaussian standard deviation of the beam 2.5 mm. A proximity gap of 40 μm is assumed and filled with He gas. Table I shows the base conditions of exposure system parameters. The beam is assumed scanning along the y direction. In this case the maximum displacement occurs along the middle line in the y direction. In the following, the distortions along this middle line are described. The amount of absorber translation (IPD) is compared with the IPD_B during the first scanning period.

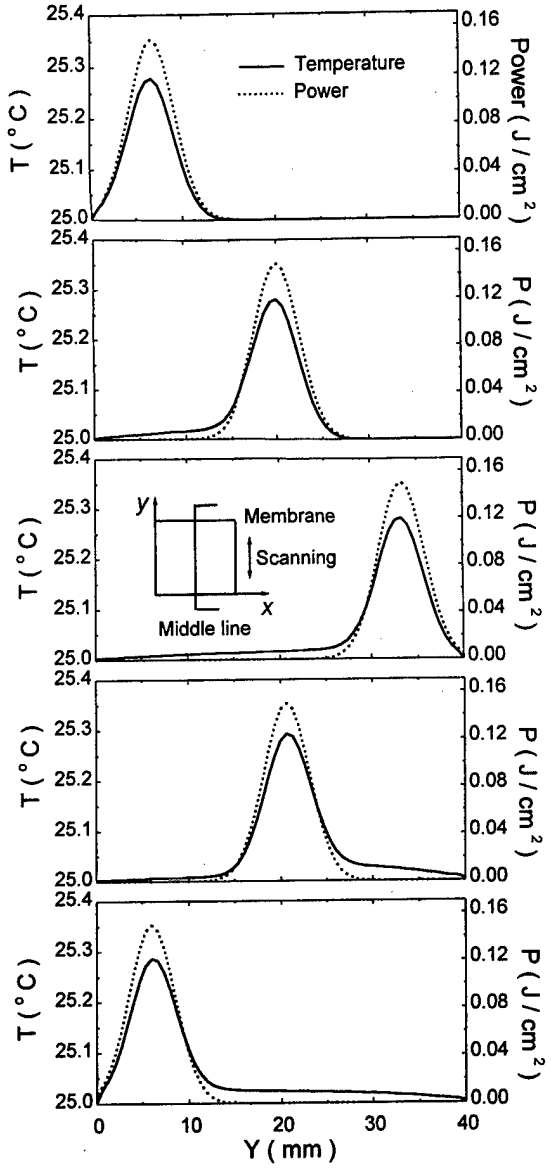


FIG. 5. Power and temperature distribution depending on the beam location.

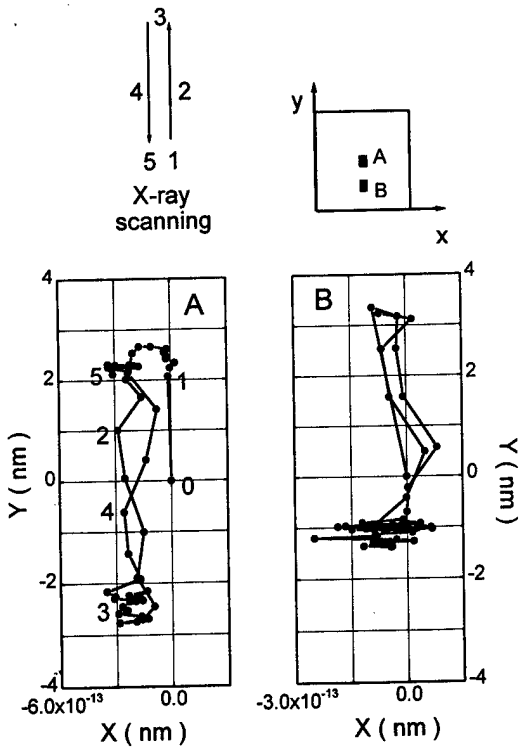


FIG. 6. History of pattern displacements.

A. Temperature rise

Figure 5 shows an x-ray power source and corresponding temperature rise of a membrane with respect to time. The base temperature is 25 °C. The scanning direction is from left to right with a frequency of 1 Hz. The maximum temperature rise is about 0.3 °C near the peak point of power intensity. In this transient thermal analysis case, the temperature profile is seen slightly lagging behind the power profile.

Both of the profiles of the power and temperature are similar in shape to each other. The severe temperature gradient in this profile makes the membrane and the patterns on it be distorted. In the first exposure of x-ray, the membrane temperature is uniform at 25 °C. After one pass of scanning, due to heat accumulation, the temperature profile is changing. The strength of x-ray power source, scanning frequency, absorption coefficient of membrane, and pattern coverage are found to be important factors which determine the temperature field.

B. Trajectory of a pattern

As a typical case, we consider the behaviors of two points A and B in a membrane to study the trajectory of a pattern. For purpose of description, the origin of the coordinate is

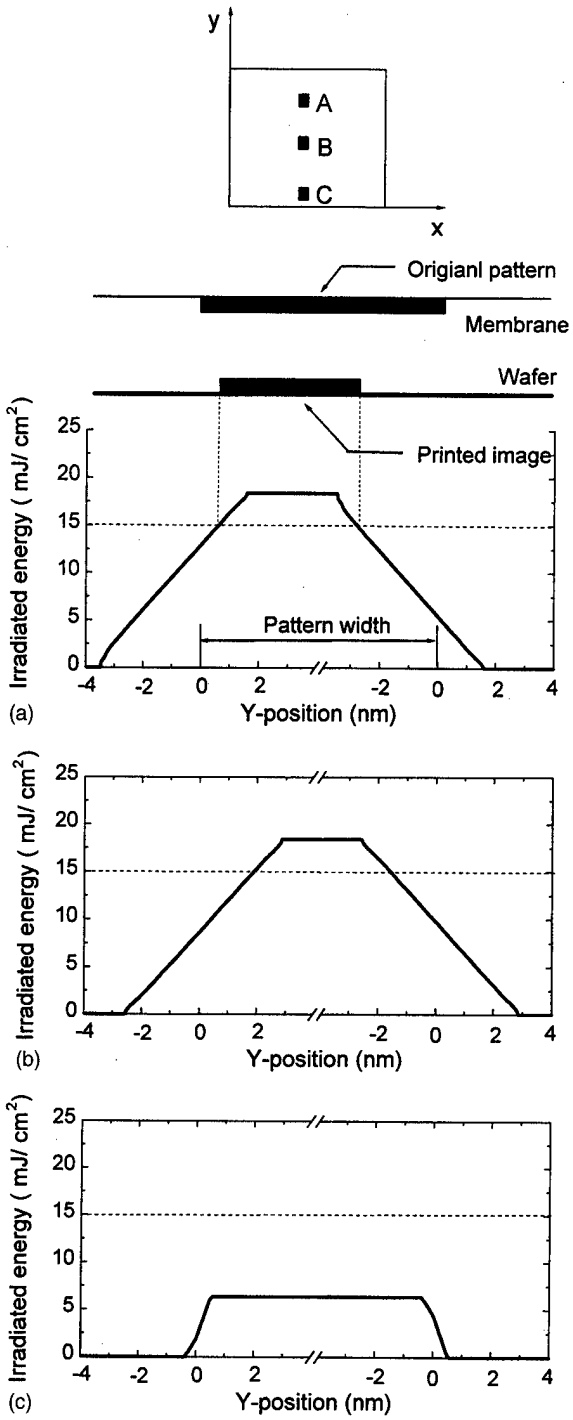


FIG. 7. Irradiated energy on a wafer during a scanning cycle. (a) Point A, (b) point B, (c) point C.

assumed to be at the initial location of each point as in Fig. 6. Point A is the center point of the pattern which is positioned at the center of the membrane. When scanning begins, point A moves upwards rapidly from the center from 0 to 1 in the figure. As the power comes to the upper end (position 3 in the scanning), the point moves down slowly and passes through the center almost at the same time as the power source passes through the center. Even when the beam reaches position 4, point A is still displaced at a downward

TABLE II. IPD_B and width changes (nm).

	Displacement of the left edge (d_L)	Displacement of the right edge (d_R)	IPD_B $\lvert(d_L+d_R)/2\rvert$	Width change d_R-d_L
Point A	0.62	−2.79	1.09	−3.41
Point B	1.91	−1.56	0.18	−3.47

location. As the power comes down further, it moves upwards, passing through the center again. As it can be seen from the figure, movement in the x direction is negligible. Since point A and point B are at the center horizontally, we can expect movements in the x direction would be almost zero.

The history of movements for point B is almost the same. But since point B is located below the center, it remains in the lower part longer than point A. In addition, the displacement of point B is less when the power source is located above point B than below B.

C. Irradiated energy

Figure 7 shows irradiated energy on a device wafer. The scanning direction taken is again along the y direction. Pattern width is about 250 nm or less. Irradiated energy depends on the position of the membrane and the power intensity. It is assumed that there is a threshold to get a print image. Maximum movements of the pattern and width change at three different locations are shown. Almost all parts of the pattern reach the same maximum level of energy intensity with a value of 17 mJ/cm^2. At points A and B, the printed image linewidth is less than the original pattern linewidth,

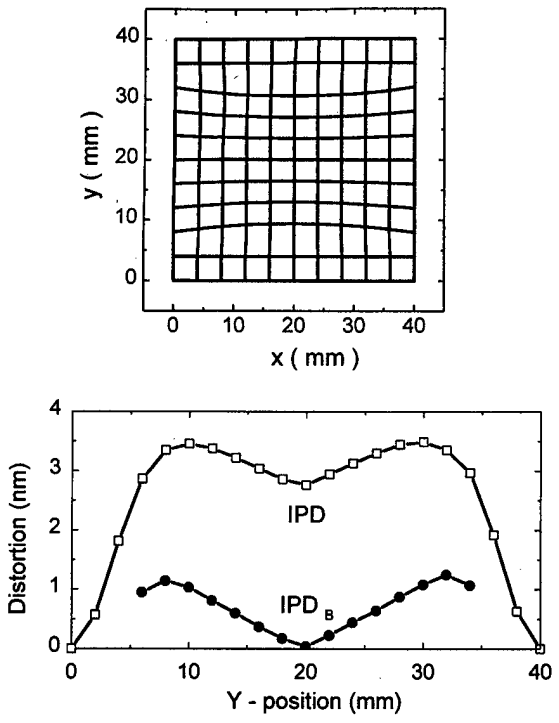


FIG. 8. Pattern distortions.

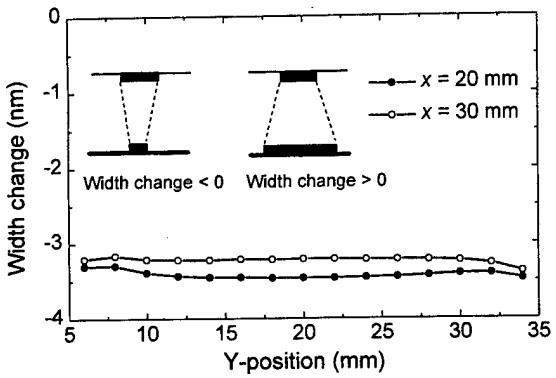


FIG. 9. Pattern width changes.

which shows the width change is about −3.5 nm. As shown in Fig. 7(c), since the irradiated energy at location C near the bottom edge is below the resist sensitivity, no pattern is printed. If the scanning started further below the bottom end, a print image would occur for the pattern at point C. IPD_B and width change are determined by the shift from the expected locations of the left edge and right edge of the printed pattern image. The numerical results from the present model case are summarized in Table II. This shows a width reduction of about 3.5 nm and a shift of about 1 nm.

D. Blur analysis

Figure 8 illustrates two kinds of pattern distortions defined previously, IPD and IPD_B. Both of them show symmetric shape, and IPD is three times lager than IPD_B which should be considered in practical use. Because irradiated energy in the bottom and top part of the membrane are not sufficient to print pattern, there are no IPD_B in these regions.

Figure 9 shows width changes in the y direction and they are about −3.5 nm at the center of the membrane. Since distortions near the edges are smaller than those in the center, the patterns near the edges shrunk less. Because the width changes in the x direction is one order smaller than in the y direction, it was not considered.

A parametric study on the IPD_B and the width change has been made next by changing scanning frequency and x-ray dose. IPD and IPD_B are found not to depend much on the scanning frequency, as shown in Fig. 10(a). This agrees with a previous study.[4] But the width change is greatly influenced by the scanning frequency. For the base conditions which were used in this study, the width change was −3.5 nm. And the linewidth change becomes zero at 0.62 Hz. It can be concluded that we can make width change to be almost zero by adjusting the scanning frequency. Figure 10(b) shows the influence of x-ray dose on distortions and width changes for the base conditions. They are increased almost linearly with x-ray dose. When the x-ray dose is low, no image is printed. The frequency at which there is no linewidth change can be changed by adjusting x-ray dose, as shown in Fig. 10(c). The estimated frequency for three cases are 0.41, 0.62, and 0.75 Hz, respectively, as shown in the figure. It is seen that the influence of the scanning frequency on IPD decreases as the

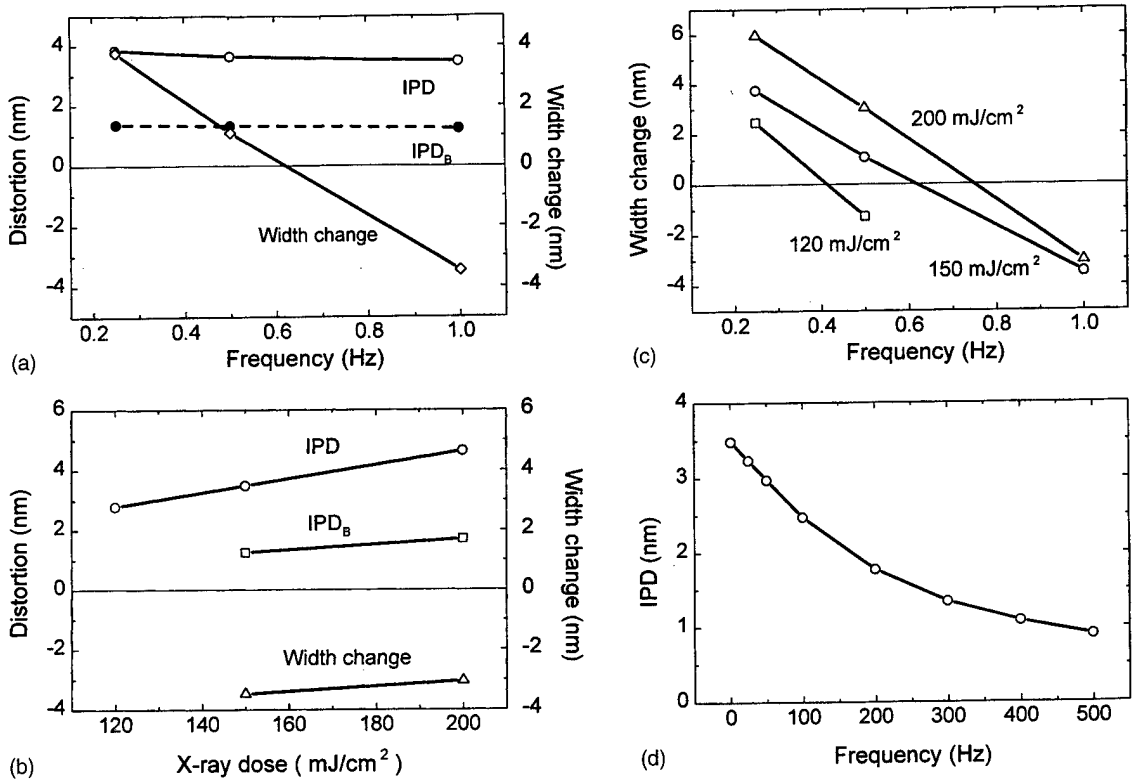


FIG. 10. (a)–(d) Distortions and width changes under various conditions.

frequency gets higher. The IPD, as shown in Fig. 10(d), becomes less, the higher the frequency is.

IV. CONCLUSIONS

In this article, blurring effects due to thermal distortion for a synchrotron radiation lithography were studied extensively. When we study thermal pattern distortions of an x-ray mask, it is necessary to consider blur because patterns are distorted dynamically. Movements of pattern were studied and irradiated energy on a device wafer were calculated considering important factors. Blurring studied has two components: shift and width change of the printed line image. They come from the distortion and displacement of the pattern on the membrane, and irradiation change of the moving x-ray. For the data we used, the displacement of the absorber pattern was 1.2 nm at the center and width change of the printed line image was about -3.5 nm. Distortions considering blurring effect were compared with distortions without considering it. Maximum IPD was about three times greater than IPD_B. It has been shown in this study that the linewidth change is inversely proportional to scanning frequency. By adjusting the frequency and x-ray dose, the width change of an image can be zero or minimized.

ACKNOWLEDGMENT

This work was supported by the Electronics and Telecommunications Research Institute.

[1]A. Chiba, Jpn. J. Appl. Phys., Part 1 **31**, 2949 (1992).
[2]A. Chiba and K. Okada, J. Vac. Sci. Technol. B **9**, 3275 (1991).
[3]Y. Vladimirsky, J. Maldonado, R. Fair, R. Acosta, O. Vladimirsky, R. Viswanathan, H. Voelker, F. Cerrina, G. M. Wells, M. Hansen, and R. Nachman, J. Vac. Sci. Technol. B **7**, 1657 (1989).
[4]K. Yamazaki, F. Satoh, K. Fujii, Y. Tanaka, and T. Yoshihara, J. Vac. Sci. Technol. B **12**, 4028 (1994).
[5]E. Haytcher and R. Engelstad, Proc. SPIE **1671**, 347 (1992).

Photoresist erosion studied in an inductively coupled plasma reactor employing CHF_3

M. F. Doemling, N. R. Rueger, and G. S. Oehrlein[a)]
Department of Physics, University at Albany, State University of New York, Albany, New York 12222

J. M. Cook
Lam Research Corp., Fremont, California 94538-6470

(Received 19 June 1997; accepted 15 May 1998)

The evolution of integrated circuits into the ultralarge scale integrated regime takes today's 0.35 μm circuit design rules to even smaller values of 0.18 μm and beyond. As a consequence, photoresist masks are becoming thinner and even more prone to erosion by etching. For this work an *I*-line novolak resist was used. Etch rates for various process conditions using *in situ* ellipsometry were obtained. Also the fluorocarbon surface layer, present on top of the photoresist during steady state etching was examined with x-ray photoelectron spectroscopy. The investigated pressure range was 6 to 20 mTorr and the inductive power range was 300 to 1400 W. It was found that there are two distinct regimes of etching behavior. At inductive powers below 600 W the etching is energy flux limited, at higher inductive powers the etching is ion energy limited. © *1998 American Vacuum Society.* [S0734-211X(98)07104-2]

I. INTRODUCTION

As integrated circuit feature sizes continue to shrink, the lithography requirements become much more demanding. A consequence of this is photoresists that are thinner and inherently less resistant to erosion by the etching processes. Simultaneously, the selectivity (of SiO_2 to photoresist) needs to increase owing to novel structures designs, approaching 10:1 in some instances. To be able to design an etching process and improve the etching resistivity of resist materials that will meet this requirement, it is important to understand the mechanisms which control the resist erosion, in particular in low pressure, high density plasmas. The etching of this amorphous organic compound is at this point not well understood. Therefore a detailed study of the etching behavior and post plasma surface analysis in dependence on inductive power, pressure, biasing power and self biasing was carried out.

II. EXPERIMENT

For this work, an inductively coupled plasma tool (ICP) was used similar to the apparatus described by Keller *et al.*[1] A planar induction coil, supplied with 0–2000 W rf power at 13.56 MHz, generates the plasma through a 16 mm thick quartz window. The wafer is mounted on an electrostatic chuck which is cooled to a temperature of 10 °C. A pressure of 5 Torr of helium is applied to the back of the wafer in order to achieve a good thermal contact. The distance between wafer and induction coil is 7 cm. For the present experiments, we used CHF_3 plasmas generated using 300 to 1400 W inductive power in the pressure range of 6 to 20 mTorr at a constant gas flow of 40 sccm. Independent rf biasing power was applied to the electrostatic chuck at 3.4 MHz. This rf bias power determines the self bias voltage which is present over the sheath region above the wafer and which largely determines the ion energy. The ion energy is the charge of the ion multiplied by the potential difference between plasma and wafer, which is the plasma potential [around 20 V (Ref. 2)] plus the self bias voltage. The relation between the rf bias power and the corresponding self bias voltage for certain plasma conditions has been established with two different methods. The first of these is by direct measurement of the wafer potential with a voltage probe attached to the wafer surface described by Rueger *et al.*;[3] the second by measuring the ion current density above the wafer with a Langmuir probe.

Assuming that all the power of the rf biasing is going into the acceleration of the ions in the sheath region towards the wafer, the self bias voltage can be expressed by

$$V_{dc}=\frac{P}{\mathrm{ICD}*A},$$

where V_{dc} is the self bias voltage, P the rf bias power, ICD the ion current density and A the area of the wafer. The calculated self bias voltages using the ion current densities agree well with the direct measurements of the self bias voltages using a wafer probe. Figure 1 shows the ion current densities for the plasma conditions used for the experiments presented in this article.

The *I*-line novolak resist material was supplied by SEMATECH. Before each resist etching experiment the process chamber was cleaned with an oxygen plasma. The etch rates were obtained using *in situ* ellipsometry. In some experiments the resist etching was performed for at least 1 min, the process was then stopped for x-ray photoelectron spectroscopy (XPS) analysis. The process was stopped by first switching off the inductive power and within approximately 0.5 s the rf bias power. Then the samples of size 2.5×2.5 cm

[a)]Electronic mail: oehrlein@cnsibm.albany.edu

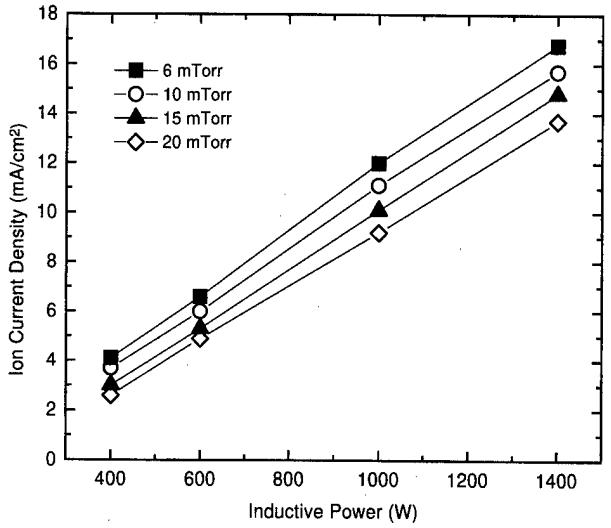


FIG. 1. Ion current density vs inductive power for the pressures 6, 10, 15 and 20 mTorr in CHF_3 at a flow rate of 40 sccm.

were transferred to the XPS analysis chamber. The scanned energy regions for XPS analysis were C(1*s*), F(2*s*), F(1*s*) and O(1*s*) under 15° and 90° with respect to the surface. The pressure in the XPS analysis chamber was $(9\pm2)\times10^{-9}$ mbar. It was observed that a delay of the XPS experiments of less than an hour would lead to a significant change in the results. Therefore, all these experiments were performed carefully in a consistent manner.

III. RESULTS AND DISCUSSION

A. Resist etch rates

An important observation is the time dependent behavior of the photoresist etch rate, and is shown in Fig. 2. All controllable parameters for this experiment were kept constant, except the rf bias power as indicated in the figure. The arrows indicate the time when the rf bias power was changed to a new value. The etch rate drops initially over a time

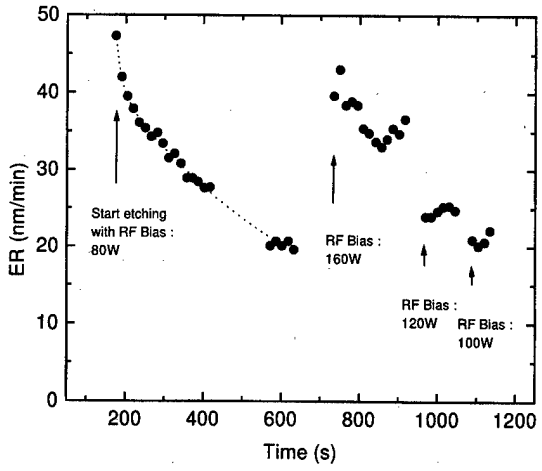


FIG. 2. Time dependence of the resist etch rate. The plasma was started at $t=0$. The arrows indicate the times when the rf bias power was adjusted to new values.

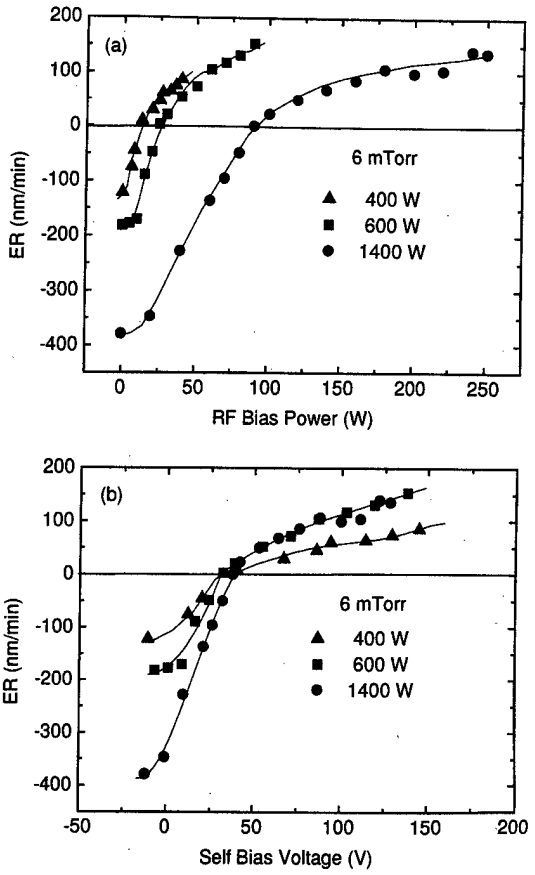


FIG. 3. Etch rate at 6 mTorr for 400, 600 and 1400 W inductive powers (a) vs rf bias power and (b) vs self bias voltage.

interval of about 500 s. After this the etch rate is stable. The change in resist etch rate with time can be explained by a change of the chamber wall conditions and the corresponding change in plasma chemistry[4] and/or an alteration of the resist material by the plasma radiation or chemistry. The duration of the following experiments was such that the time dependent change of the etch rate is negligible.

To get an overview of the etch rate behavior, we started out varying the rf bias power while keeping the other parameters constant. We did this at a pressure of 6 mTorr for inductive powers of 400, 600 and 1400 W and at 20 mTorr using 400, 1000 and 1400 W. The flow rate in these experiments was 40 sccm. The results are shown in Figs. 3(a) and 3(b) and Figs. 4(a) and 4(b), where the rf bias power has also been converted into self bias voltage.

An interesting observation here is in the etch rate behavior with respect to the self bias voltage. At 6 mTorr pressure we see that the etch rates for 600 and 1400 W are the same at the same self bias voltage whereas the etch rate at 400 W is lower. At 20 mTorr we can see a similar behavior, but for different values of inductive powers. Here we get coinciding etch rate curves for 1000 and 1400 W inductive power, but at 600 W the etch rate is much lower.

A comparison of the etch rate at pressures of 6 and 20 mTorr is shown in Figs. 5(a) to 5(d). Again the etch rate is plotted versus rf bias power as well as self bias voltage. In

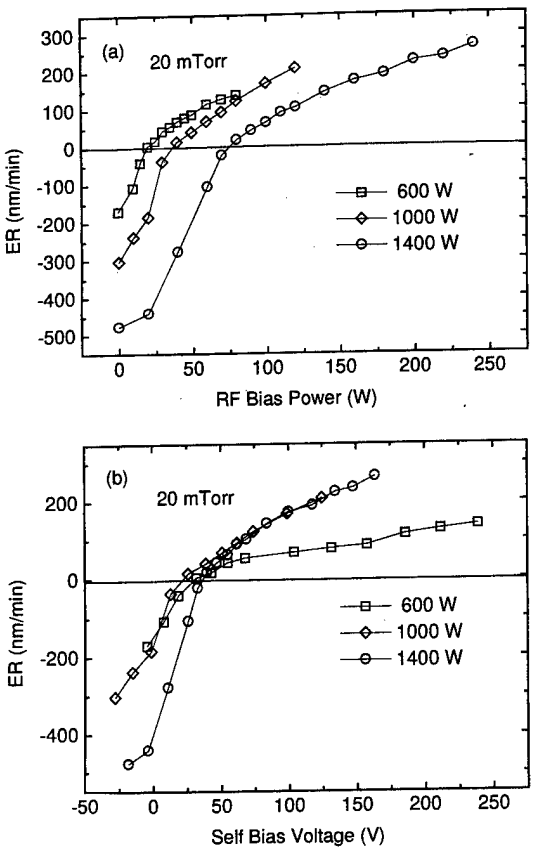


FIG. 4. Etch rate at 20 mTorr for 600, 1000 and 1400 W inductive powers (a) vs rf bias power and (b) vs self bias voltage.

Fig. 5(a) where the data are plotted versus rf bias power, we see that for 600 W the etch rates of both pressures are very close, especially for bias powers well in the etching regime. When plotted versus self bias voltage [Fig. 5(c)] the etch rates at higher self bias voltages are quite different, but at the transition from deposition to etching both curves coincide. The etching at both pressures starts at the same self bias voltage.

This observation seems to indicate that for 600 W inductive power, the onset of the etching is determined by the self bias voltage and therefore a well defined ion energy. For higher values of the self bias voltage, the total energy flux seems to be a limiting factor of the etch rate so that we measure the same etch rate for both pressures at the same rf bias power. Figure 5(b) shows that at 1400 W the rf bias power does not seem to limit the etch rate, but the onset of etching is still determined by the self bias voltage [see Fig. 5(d)].

Since the etch rate at a fixed self bias voltage above a certain inductive power seems to be independent of the inductive power [Figs. 3(b) and 4(b)], we measured the etch rate at a constant self bias voltage, varying the inductive power. In Fig. 6 we see the results for two different self bias voltages, 50 and 80 V at three different pressures 6, 15 and 20 mTorr. It can be seen that the etch rate levels out for higher ion currents. Even a change in pressure does not change significantly the etch rate at higher ion currents for the two given voltages. At lower ion currents the lower pressure yields a higher etch rate.

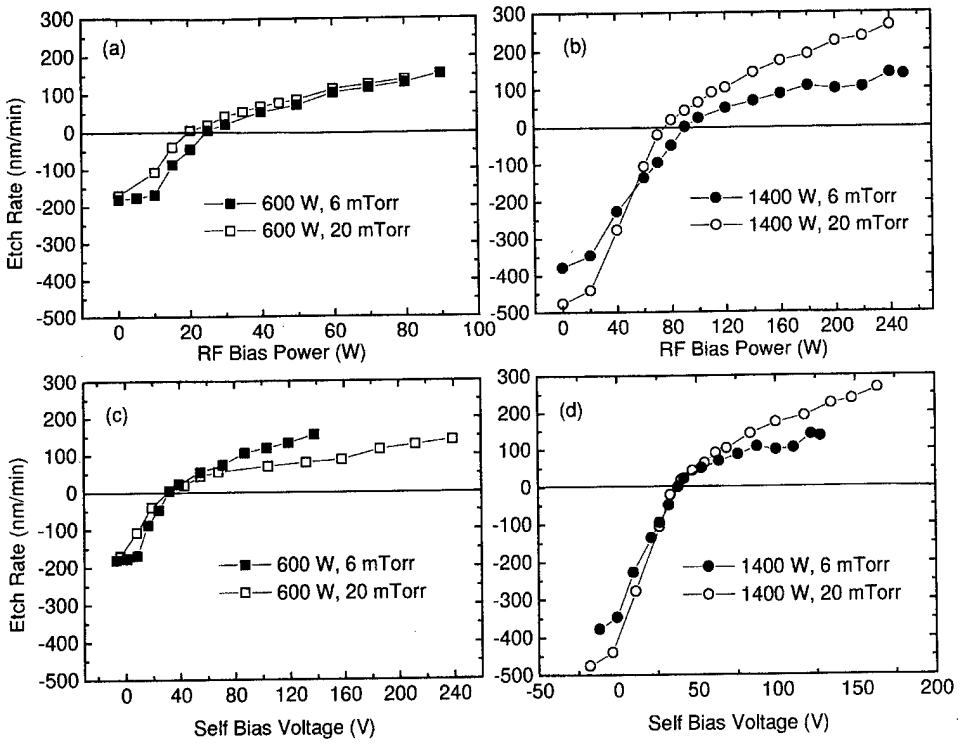


FIG. 5. Comparison of etch rates of two different pressures 6 and 20 mTorr. (a) 600 W vs rf bias power; (b) 600 W vs self bias voltage, (c) 1400 W vs rf bias power, and (d) 1400 W vs self bias voltage.

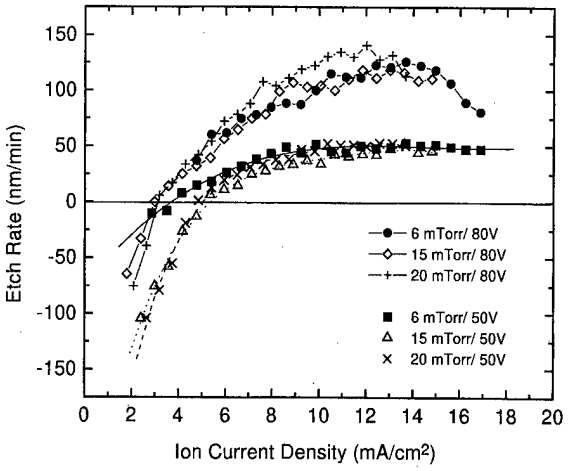


FIG. 6. Etch rate vs ion current density for two constant self bias voltages 50 and 80 V and three different pressures 6, 15 and 20 mTorr.

B. Correlation between resist and fluorocarbon etching

A correlation between the resist etch rate and the etch rate of a passively deposited fluorocarbon layer was observed in the following experiments. Initially, a fluorocarbon film was deposited without applying an rf bias power. After a sufficiently large CF_x layer was grown, the rf bias power was applied, and the CF_x etch rate was measured followed by the resist etch rate, once the fluorocarbon film was removed. This experiment was repeated for each rf bias power investigated. The obtained results for 1400 W inductive power and 6 mTorr pressure are shown in Figs. 7(a) and 7(b). We observe that the etch rates of the passively deposited CF_x and the etch rates of the resist are proportional to each other.

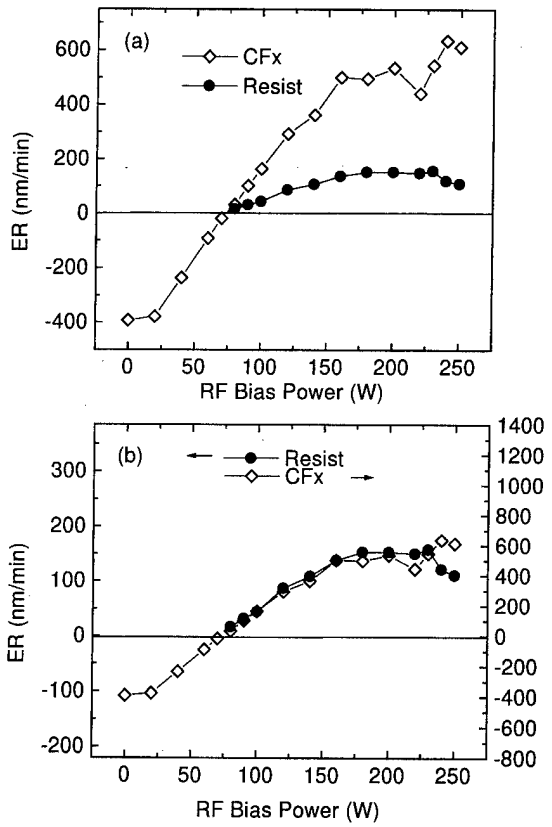


FIG. 7. (a) Etch rate of resist and CF_x at 1400 W and 6 mTorr vs rf bias power. (b) Same as (a) with two different scales for the CF_x etch rate and the resist etch rate.

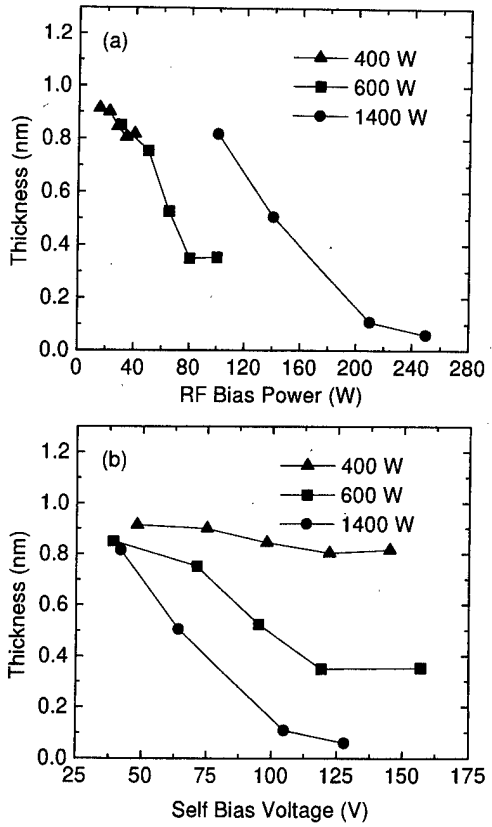


FIG. 8. Thickness of the CF_x top steady state etching layer at 6 mTorr and three different inductive powers 400, 600 and 1400 W (a) vs rf bias power, (b) vs self bias voltage.

C. X-ray photoelectron spectroscopy surface analysis

XPS analysis of the resist surface during etching was performed in order to study properties of the CF_x steady state etching layer on top of the resist as a function of various etching conditions. Unfortunately, the resist itself contains carbon, so that the total carbon signal is the sum of the signals coming from the CF_x and from the underlying resist. The CF_x layer is too thin to neglect the contribution from the resist. Since the resist does not contain fluorine, we know the origin of the C–F_x bond signals and we can use these to make statements about the composition of the CF_x layer.

The resist also contains oxygen which we can use to determine the thickness of the CF_x layer by comparing the intensities emitted under 15° and 90° with respect to the surface using the method outlined by Briggs and Seah.[5] Figures 8(a) and 8(b) show the thickness of the CF_x layer at 6 mTorr for 400, 600 and 1400 W inductive power with respect to rf bias power and self bias voltage. We see that the thickness decreases with self bias voltage as well as with

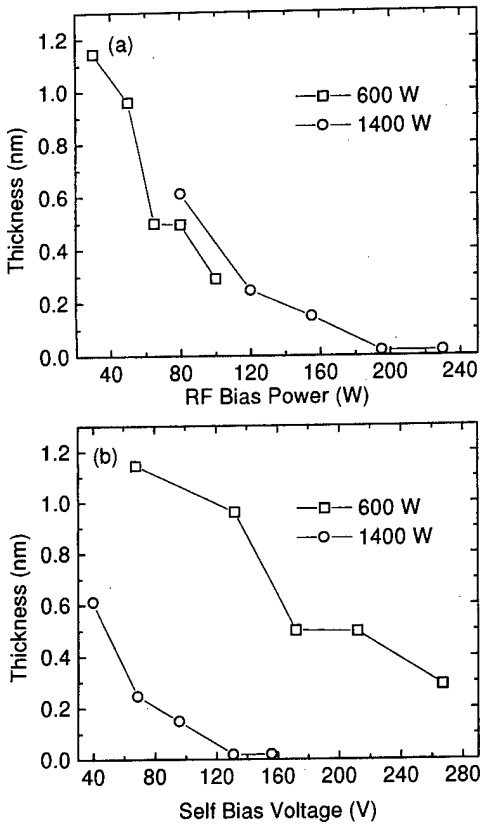


FIG. 9. Thickness of the CF_x top steady state etching layer at 6 mTorr and two different inductive powers 400 and 1400 W (a) vs rf bias power and (b) vs self bias voltage.

inductive power. Therefore we get the thinnest film at the highest self bias voltage and the highest inductive power. Figures 9(a) and 9(b) show thicknesses obtained at 20 mTorr at 600 and 1400 W inductive power. Here we see the same trend. It is interesting to notice that here the thickness seems to be determined by the rf bias power and therefore by the total energy flux, as shown in Fig. 9(a), independently of the inductive power. Also in Fig. 8(a) for the curves of 400 and 600 W inductive power the rf bias power seems to be the determining factor of the thickness.

In Figs. 10(a) to 10(d) a comparison of thicknesses of CF_x layers processed under the two pressures 6 and 20 mTorr is shown. For 600 W, we see that at the lower pressure we find lower thicknesses, whether plotted versus rf bias power [Fig. 10(a)] or self bias voltage [Fig. 10(c)]. For the rf bias powers above 60 W, the film thicknesses for both pressures appear to be very close [Fig. 10(a)]. At the high inductive power 1400 W the behavior is reversed. Here we get the higher thicknesses for the lower pressure 6 mTorr. How does this changing behavior of the CF_x film thicknesses as a function of inductive power, correlate to the etch rates at these conditions? The corresponding etch rate graphs were already shown in Fig. 5. When we compare each etch rate graph with the corresponding thickness graph, we find at 600 W inductive power when plotted versus rf bias power: (1) A thinner film is seen at 6 mTorr [Fig. 10(a)] at the onset of etching at around 30 W [Fig. 5(a)]. However, the etch rate is lower. (2) At higher rf bias powers, the etch rates are almost identical and also the CF_x thicknesses at both pressures are very close.

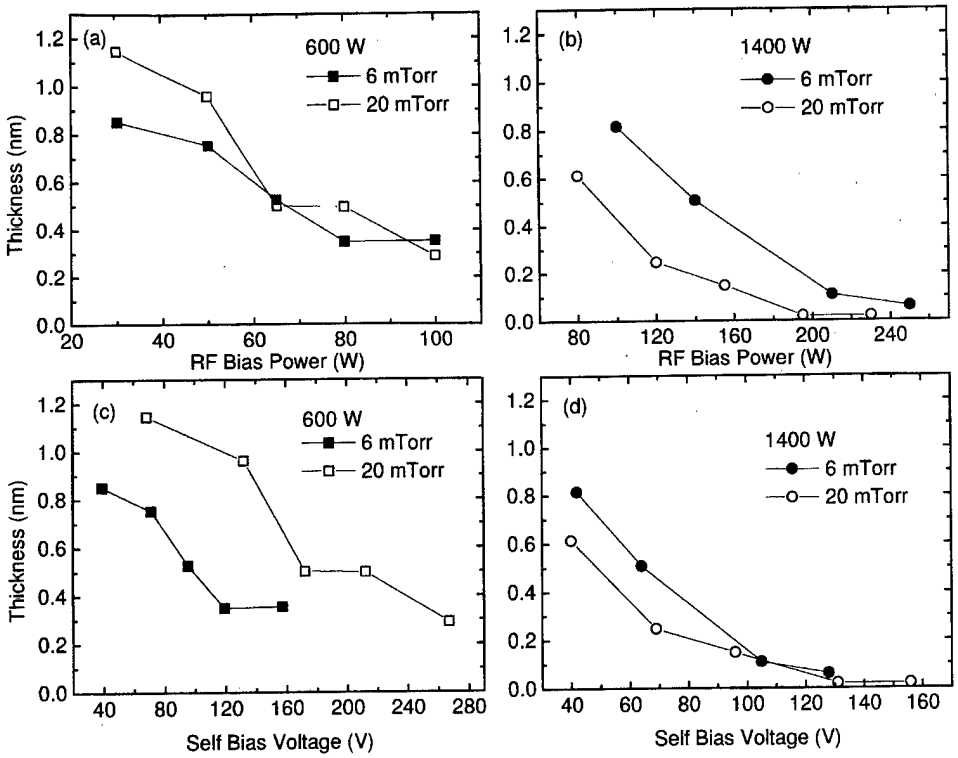


FIG. 10. Thickness of the CF_x top steady state etching layers for 6 and 20 mTorr (a) at 600 W vs rf bias power, (b) at 1400 W vs rf bias power, (c) at 600 W vs self bias voltage, and (d) at 1400 W vs self bias voltage.

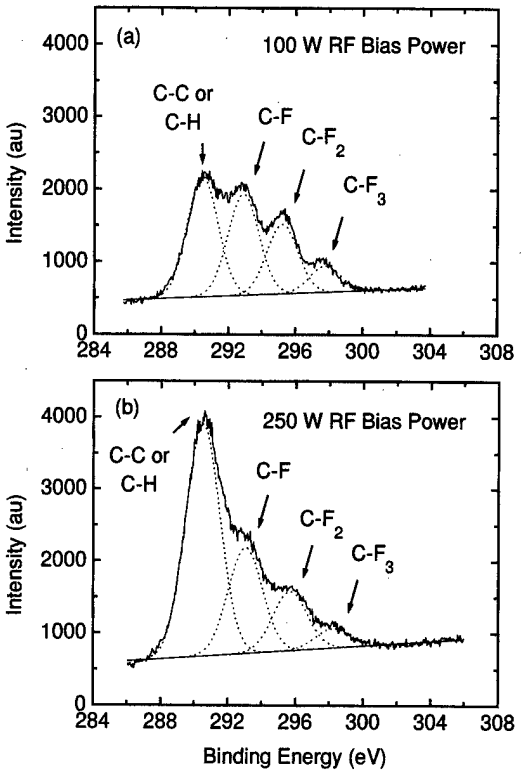


FIG. 11. Carbon spectra of resist samples processed at 1400 W inductive power, 6 mTorr pressure at rf bias powers of (a) 100 W and (b) 250 W. The spectra were obtained at an angle of 15° with respect to the surface.

This means that for higher rf bias powers, we get the same etch rate and the same film thickness at the same rf bias power. As we discussed earlier, the etching in this regime appears to be energy flux limited. Our data can be interpreted in that the energy flux limitation is correlated with the thickness of the CF_x top layer which needs to be removed to enable more effective etching of the resist. This is also consistent with the results shown in the two curves at 1400 W versus rf bias power [Figs. 5(b) and 10(b)]. We find here that the higher CF_x film thickness at 6 mTorr corresponds to a lower etch rate which is the expected behavior.

By plotting the data versus self bias voltage, Figs. 5(c) and 10(c) are obtained. Here thinner film corresponds to a higher etch rate for 1400 W as well as for 600 W. As mentioned earlier the thickness difference between 6 and 20 mTorr at 600 W is reversed for 1400 W. At 600 W and 6 mTorr the CF_x film is thinner than the one at higher pressure and at 1400 W the opposite is the case. This corresponds to the same change of the etch rate difference when going from 600 to 100 W. At 600 W and 6 mTorr the etch rate is higher than for 20 mTorr and at 1400 W the opposite is the case. This indicates that there is an important correlation between the formation of the steady state etching CF_x film, and the etch rate.

Figure 11 shows some typical carbon 1*s* spectra for two different bias powers processed at 1400 W inductive power and 6 mTorr pressure. By fitting these spectra with a combi-

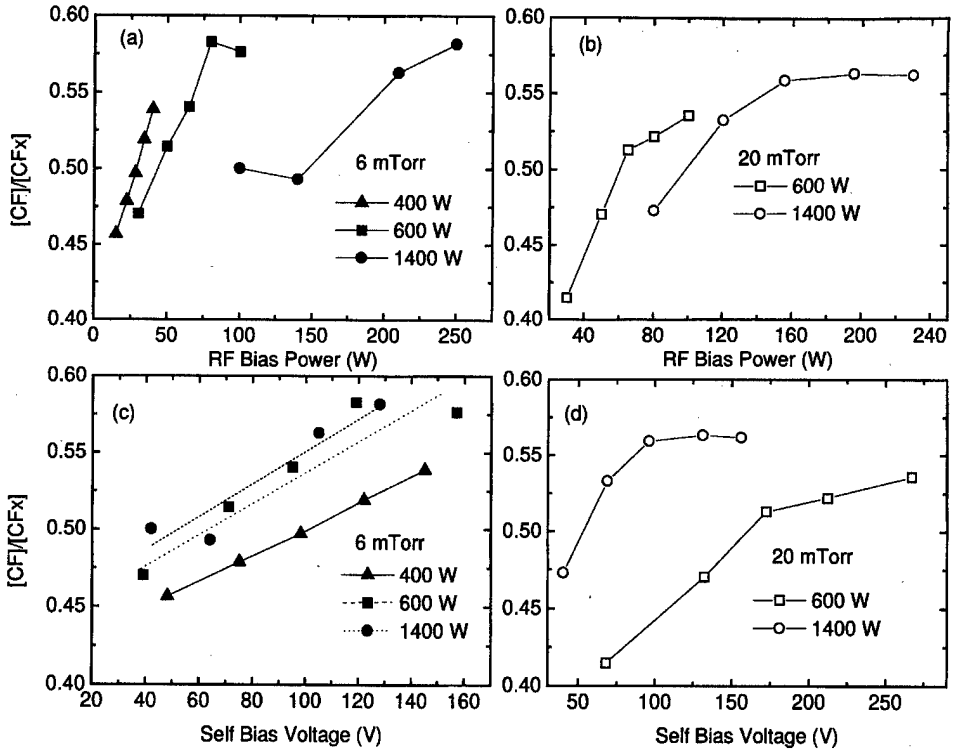


FIG. 12. Ratio of C–F bonds among all fluorine bonds in the steady state etching layer (a) at 6 mTorr for 400, 600 and 1400 W inductive power vs rf bias power, (b) at 20 mTorr for 600 and 1400 W inductive power vs rf bias power, (c) at 6 mTorr for 400, 600 and 1400 W inductive power vs self bias voltage, and (d) at 20 mTorr for 600 and 1400 W inductive power self bias voltage.

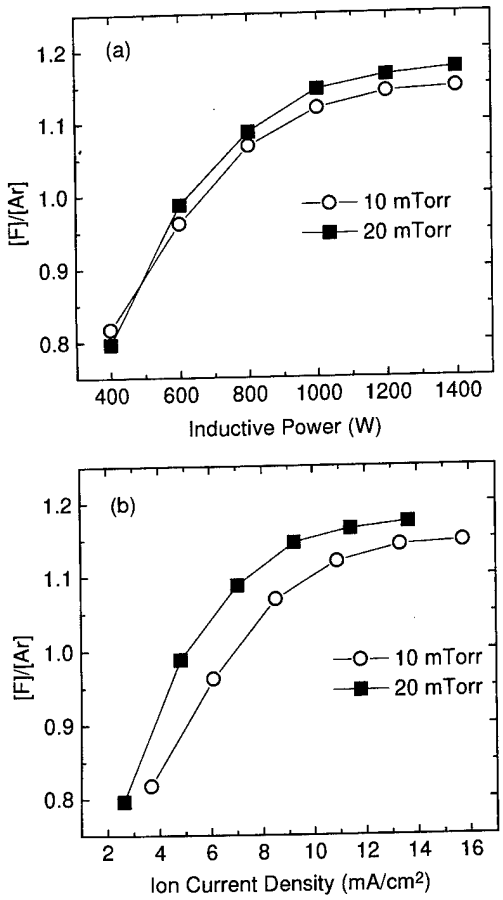


FIG. 13. Fluorine over argon intensity vs (a) inductive power and (b) ion current density at the pressures 10 and 20 mTorr.

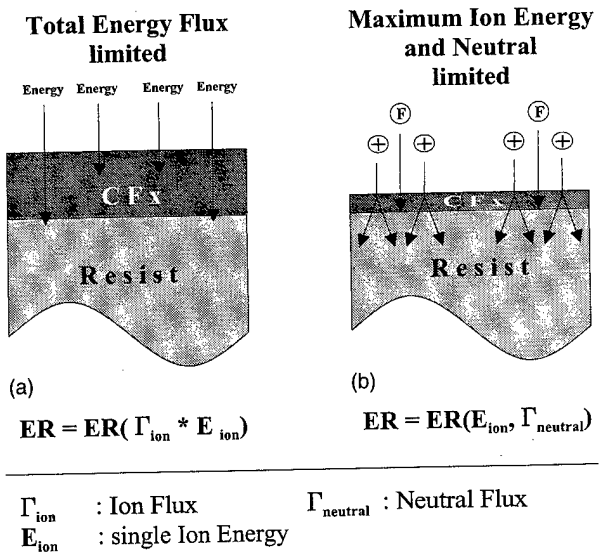


FIG. 14. Schematic picture of the two different etching regimes. (a) In the energy flux limited regime a thick CF_x layer is present. (b) In the ion energy limited regime the amount of ion fragmentation and the penetration depth of the ions is important.

nation of Gaussian curves, the relative ratios of the corresponding bonds can be determined. The carbon spectrum contains four peaks according to the four types of bonds that are possible in a pure CF_x layer, the C–C, C–F, C–F_2 and the C–F_3 bonds. As already mentioned earlier we can only make use of the C–F_x bonds, since the resist contributes to the other signals with its C–C bonds. In addition, C–H bonds that exist in the resist contribute to the C–C peak since their binding energies are practically undistinguishable. All the C–F_x peaks must come from the CF_x layer since the resist does not contain fluorine. Using the peaks arising from a carbon atom with at least one fluorine bond, the relative occurrence of C–F bonds among all the C–F_x bonds were calculated. This was done with the carbon 1*s* signals obtained at a detector angle of 15° with respect to the sample surface. The results are shown in Figs. 12(a) to 12(d). Again we can compare the results plotted against rf bias power with the plots where the results are shown versus self bias voltage. By doing so we find that the ratio of C–F bonds is a function that is more directly connected with the self bias voltage than the rf bias power. We see that with increasing self bias voltage the CF_x film contains a higher amount of C–F bonds. This is to be expected since a higher ion energy introduces a higher damage to the surface, and the ions will on average experience a greater fragmentation.[6] By increasing the inductive power at the same self bias voltage, the amount of C–F bonds can be further increased. This behavior is common for both pressures 6 and 20 mTorr.

A higher relative amount of C–F bonds means that the film becomes less fluorinated and therefore harder to etch, since fluorine is the important etch precursor. Since our goal is to minimize the erosion of photoresist a less fluorinated C–F_x layer is desirable. We also want to maximize the thickness of the CF_x layer since this ultimately controls the etch rate of the underlying resist and a thicker film is a better protection against erosion. As can be seen from the graph the least fluorinated film [highest C–F bond ratio, Figs. 12(c) and 12(d)] is achieved at the highest self bias voltage with the highest inductive power. Unfortunately, this process direction also leads to the thinnest films [Figs. 8(b) and 9(b)]. This indicates that ion bombardment energy control alone may be insufficient to control resist erosion.

IV. ETCHING MODEL

From the etching behavior we can distinguish two regimes of etching. At low inductive powers the etch rate is energy flux limited and the rf bias power is the limiting factor [Fig. 5(a)]. At high inductive powers above a "critical ion flux," the etching is independent of the inductive power at fixed self bias voltage and determined by this self bias voltage [Figs. 3(b) and 4(b)] and the neutral density as discussed below. Since with increasing inductive power at fixed self bias voltage the ion current density and therefore the total energy flux to the wafer surface increases, it is clear that in this regime it is not the energy that controls the etching process. However, an increase in self bias voltage and therefore the maximum ion energy does increase the etch rate, so we conclude that the greater ion fragmentation[6] upon impact,

and the greater penetration depth of the ions, at higher ion energies are responsible for the increase in etch rate.

The independence of the etch rate from the ion current density can be explained when taking the atomic fluorine density in the plasma into consideration. Fluorine is an important etch precursor and the etch rate should always depend on the atomic fluorine density to which the surface is exposed. We observed small fluorocarbon layer thicknesses on the resist (<1 nm), which decrease with increasing inductive power and increasing self bias voltage. For high inductive power and self bias voltage, these films are not present. Therefore, above the critical ion flux, the fluorine density to which the resist surface is exposed, is equal to the fluorine density in the gas phase. Argon actinometry was performed to investigate the fluorine densities present during the etch process. Figure 13 shows the F/Ar ratios from optical emission spectra versus inductive power and ion current density. The fluorine density saturates as the inductive power increases above roughly 1000 W. In the case of CHF_3, hydrogen scavenging of fluorine can prevent a linear increase of the fluorine density with inductive power. This behavior of the fluorine density is similar to the observed etch rates of the resist shown in Fig. 6.

It is concluded that the etch mechanism for ion fluxes above the critical ion flux is an ion enhanced neutral etching and depends on the atomic fluorine density. Observation of the etch rate being independent from the ion current density (inductive power), is only possible when at the same time the atomic fluorine density is also independent of the ion current density (inductive power). Figure 14 shows a schematic picture of the processes in the two different regimes.

V. CONCLUSIONS

To find an optimum process one has to keep in mind that a high selectivity must be achieved. Therefore while minimizing the resist erosion, the etch rate of the SiO_2 has to remain high. When considering conditions under which the resist erosion can be minimized the etching behavior of the SiO_2 has to be taken into account and cannot be treated separately. It can be very helpful to use the presented etching model and the correlations found between etch rate, rf bias power, self bias voltage, CF_x film thickness, etc. in such considerations when comparing them to the corresponding behavior of other materials. For instance, the SiO_2 etch rate is directly proportional to the ion current and proportional to the square root of the ion energy.[7] This different dependence can be used to maximize the SiO_2/resist etch rate ratio.

This study reveals some insights into the behavior of a number of properties that are related to the erosion of photoresist. There still remain many unanswered questions that need to be addressed for a good understanding of the complex mechanisms and factors that determine resist loss. How does temperature change the resist behavior during etching? Does the plasma radiation alter the photoresist? How is the etching of CF_x related to the etching of resist? What are the detailed mechanistic roles of neutrals and ions in the etching of resist? These questions will be addressed in future work.

ACKNOWLEDGMENTS

The authors would like to thank M. Schaepkens, T. E. F. M. Standaert, J. Mirza and H. Sun for stimulating discussions and help with experiments. The work was supported in part by Lam Research and the New York State Science and Technology Foundation.

[1]J. H. Keller, J. C. Forster, and M. S. Barnes, J. Vac. Sci. Technol. A **11**, 2487 (1993).
[2]M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing* (Wiley, New York, 1994).
[3]N. R. Rueger, J. J. Beulens, M. Schaepkens, M. Doemling, J. M. Mirza, T. E. F. M. Standaert, and G. S. Oehrlein, J. Vac. Sci. Technol. A **15**, 1881 (1997).
[4]J. A. O'Neill and J. Singh, J. Appl. Phys. **77**, 497 (1997).
[5]D. Briggs and M. P. Seah, *Practical Surface Analysis: Auger and X-Ray Photoelectron Spectroscopy*, 2nd ed. (Wiley, New York, 1996).
[6]W. H. Chang, I. Bello, and W. M. Lau, J. Vac. Sci. Technol. A **11**, 1221 (1993).
[7]G. S. Oehrlein, D. Zhang, D. Vender, and O. Joubert, J. Vac. Sci. Technol. A **12**, 333 (1994).

Apparent tunnel barrier heights of PtIr–Au interfaces in relation to the Au surface composition

Lionel Boyer,[a] Sophie Noël, and Frédéric Houzé
Laboratoire de Génie Électrique de Paris, URA CNRS 0127, Universités Paris VI et Paris XI, Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette cedex, France

(Received 4 December 1997; accepted 24 April 1998)

In the literature, the phenomenon of barrier height lowering in tunneling experiments is mainly assigned to the presence of contaminants at the interface. In this article we present the results of barrier height measurements made between a PtIr tip and various types of gold surfaces. Vacuum deposited and galvanic gold samples are investigated, as received, after ultraviolet (UV)/ozone treatment and covered with a fluid used to lubricate electrical contacts. It is found that the value of the barrier height depends both on the gold deposition process and on the treatment. Results of surface analyses performed on the nonlubricated samples are also reported. They show that the galvanic gold is more contaminated by carbon than the vapor deposited gold. The effects of the ultraviolet (UV)/ozone treatment are then described. It is observed that for nonlubricated surfaces, the higher the oxygen content, the higher the mean barrier height, and that, for each kind of gold, the lubricated surface exhibits the largest barrier height. All these various results are discussed with reference to the results of the surface analyses and to articles found in the literature. © *1998 American Vacuum Society.* [S0734-211X(98)03204-1]

I. INTRODUCTION

The phenomenon of barrier height lowering in scanning tunneling microscopy has been investigated many times in the past years.[1–7] It has been observed only when surfaces are contaminated, but paradoxically the role of the contaminants in the lowering mechanism is not yet well understood. From a very different point of view, it is recognized that the electrical and tribological behaviors of low level electrical contacts (i.e., contacts without Joule heating or arcing) are improved when liquid lubricants are used to protect their metallic surfaces.[8–16] It has been experimentally demonstrated that lubricants reduce the wear of metallic coatings (gold, in particular) and that they can also protect these metallic surfaces from corrosive atmospheres. However, to our knowledge, no study concerning the electrical role of lubricants in electrical contacts can be found in the literature. When we started this work, we intended to study tunnel junctions involving gold surfaces and a lubricant used for electrical contacts. Indeed, the surfaces of the electrical contacts for high performances applications, are covered by a final layer of galvanic gold. The parameter we considered was the apparent barrier height: we compared the results for lubricated and nonlubricated junctions and observed that the values found with the lubricant were significantly higher than those found without it. Since it is known that galvanic gold is not pure, we carried out for comparison experiments on vapor deposited gold. Finally, we used in some circumstances the UV/ozone process to clean the two types of gold surfaces. We also observed an effect on the barrier height values after this kind of treatment. Surface analyses were performed, on the different nonlubricated gold surfaces, in order to look for the origin of the observed differences of barrier heights. The purpose of this article is to present the results of our electrical and physico-chemical investigations and to propose a qualitative explanation for the observed differences in the values of barrier heights.

II. SAMPLE PREPARATION AND EXPERIMENTS

A. Sample preparation

Two kinds of gold samples have been used in the present study. First, samples with coatings similar to those used for the electrical contacts were used, i.e., a brass substrate covered with a galvanic sulfamate, 3-μm-thick, nickel plating, and with a 1.2-μm-thick galvanic gold terminal coating. Second, electron gun vapor deposited gold films, 1000 Å thick, on glass substrates were prepared. These samples were used as a reference for pure gold.

In order to test the effect of cleaning by the UV/ozone (UVO) process, samples of the two kinds were placed in the drawer of a Jelight™ UVO cleaner for 20 min.

The lubrication of both kinds of gold samples was performed by the dipping and withdrawal method. A solution of Y1800 FOMBLIN® PerFluoroAlkylEther (PFAE) at 1% in freon 113 was used to deposit the PFAE lubricant layer. The structure of the molecule is as follows:

$$CF_3\text{–}[(O\text{–}\underset{\displaystyle CF_3}{\underset{|}{CF}}\text{–}CF_2)_M\text{–}(O\text{–}CF_2)_N]\text{–}O\text{–}CF_3$$

with $M/N \approx 30$. The immersion time was 20 min and the withdrawal speed 5 mm/s. This gave homogeneous thin films.

[a]Electronic mail: boyer@lgep.supelec.fr

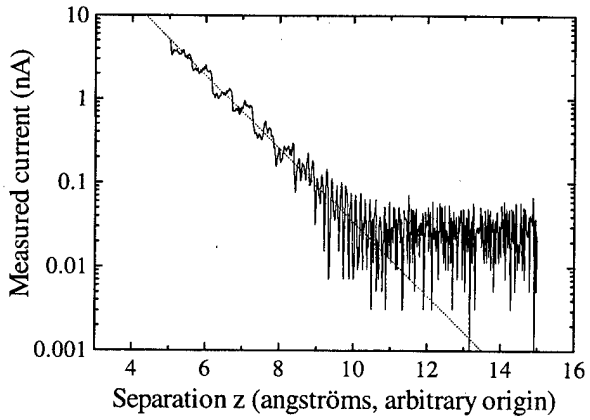


FIG. 1. Example of a plot of the measured current vs the tip surface separation. The origin is arbitrary.

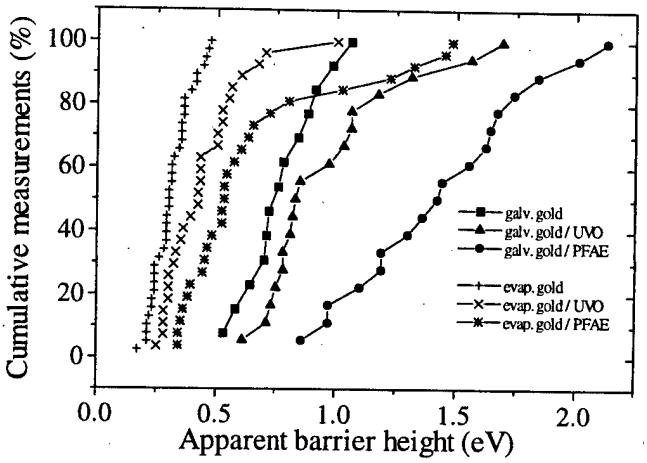


FIG. 2. Values of the apparent barrier height for various gold surfaces.

B. Tunneling measurements and apparent barrier height determination

Tunneling measurements were carried out in room atmosphere with a Nanoscope® III (Digital Instruments) setup in the scanning tunneling microscope (STM) TipView™ configuration. Probes were commercial PtIr 80/20 "NanoTips" resulting from cutting a 0.01-in-diam wire. For each gold sample, the microscope was first operated in the constant current image mode. The main parameters of the system (current setpoint, bias, feedback gains) were optimized to obtain satisfactory images for decreasing scan sizes from 5 μm to 100 nm. Then the microscope was switched to the spectroscopic $I(z)$ mode in order to measure the change in the tunneling current as a function of the tip to sample separation at fixed given (x,y) locations on the sample. It is classically accepted that, when the $\ln(I)$ versus z curve exhibits a substantial linear portion, I is of tunneling type and the apparent barrier height $\bar{\Phi}$, in eV, can be determined by calculating the expression[5]

$$\bar{\Phi}=0.952\left(\frac{\partial \ln(I)}{\partial z}\right)^2 \quad (1)$$

in the linear part of the curve, plotted with z in Å. Each $I(z)$ acquisition cycle on the Nanoscope is composed of two phases during which the bias voltage is held constant. During the first phase the feedback is activated to establish the setpoint current previously chosen. Then the second phase consists in ramping z away from the surface and acquiring the current values at a specified number of points. About 20 to 30 curves of such type were recorded on each sample. All current data were converted from the coded integers of Nanoscope files into ASCII files of real nAmps values and displayed with the Origin™ software in the $\ln(I)$ vs z form. We show in Fig. 1 a typical example of such a curve obtained with the lubricated galvanic gold sample. Finally, for each curve, data were clipped so that only the linear part remained and the desired slope $\partial \ln(I)/\partial z$ was extracted by running the automatic linear fit function of the software. Then $\bar{\Phi}$ values were calculated from expression (1).

C. Surface analysis method

X-ray photoemission spectroscopy (XPS) measurements were performed using a RIBER MAC2™ Auger/XPS system. The samples were excited with the Mg $K\alpha$ nonmonochromated α radiation (1253.6 eV). Measurements were recorded at a constant electronic resolution of 0.5 eV which gave a full width at midheight of 1.37 eV for the copper $2p_{3/2}$ peak. Calibration procedures were performed before any daily acquisition giving the Au $4f_{7/2}$ peak at 84.03 eV. Atomic fractions of the analyzed surfaces were calculated with sensitivity coefficients based on the ionization cross section, the mean free paths, and the analyzer characteristics.

III. RESULTS

A. Barrier heights

In Fig. 2 are plotted the cumulative measurements showing the barrier height distributions for the six types of gold plated coupons. From these data, we have calculated the mean values and the standard deviations reported in Table I. Two facts may be outlined about these results. First, whatever the state of the surface, i.e., without treatment, UVO treated or lubricated, the values for evaporated gold are all lower than those for galvanic gold, exception made for some values relative to the lubricated samples. Second, for each type of gold, the mean apparent barrier heights $\bar{\Phi}_{mean}$ appear in the same order, i.e.,

TABLE I. Values, in eV, of the mean barrier heights $\bar{\Phi}_{mean}$ and of the mean standard deviations σ for the six gold surfaces studied. ED: evaporation deposited, Gal.: galvanic, AR: as-received, UVO: 20 min. UV/ozone treated, Lub.: lubricated.

	ED AR	ED UVO	ED Lub.	Gal. AR	Gal. UVO	Gal. Lub.
$\bar{\Phi}_{mean}$	0.31	0.44	0.66	0.77	0.97	1.44
σ	0.08	0.16	0.34	0.15	0.29	0.36

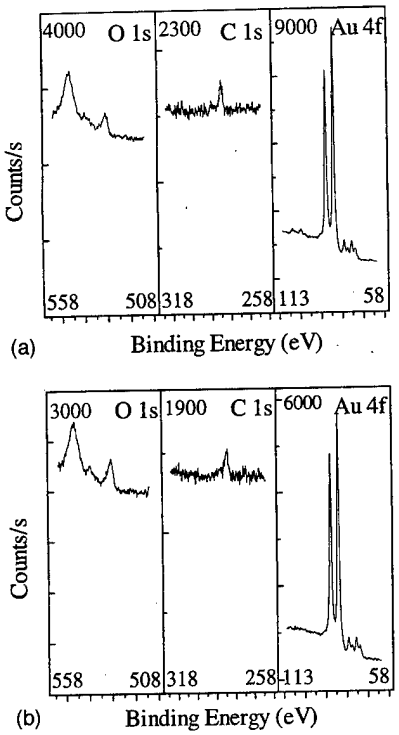


FIG. 3. XPS spectra showing the oxygen, carbon, and gold core levels of evapored gold surfaces: (a) as-deposited and (b) after 20 min UVO.

$$\bar{\Phi}_{\text{mean}}(\text{without treatment}) < \bar{\Phi}_{\text{mean}}(\text{UVO}) < \bar{\Phi}_{\text{mean}}(\text{lubricated}).$$

This result will be discussed later.

B. Surface analysis

The surfaces of deposited and UVO treated evaporated and galvanic gold samples have been analyzed. Those of the lubricated samples have not been analyzed because the lubricant film is too thick. Only carbon, oxygen, and fluorine from the lubricant are detected on such surfaces.[17] We will begin with the results from the evaporated gold samples, since they are the easiest to understand.

TABLE II. Surface composition in % of the evaporated and galvanic gold layers, as-received and after 20 min UVO treatment. The O peak is decomposed into contamination type and into O···Au type oxygen and is given in % of the total oxygen content.

					O		
		C	S	Au	Conta.		O···Au
ED	AR	39.5	···	47.7		12.8	
					51		49
	UVO	25.2	···	60.3		14.5	
					55		45
Gal.	AR	68.4	···	13.0		18.6	
					94		6
	UVO	25.6	6.4	28.0		40.0	
					52		48

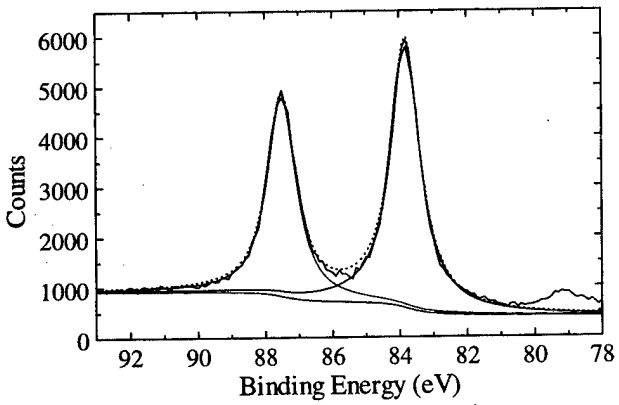


FIG. 4. XPS spectra of the Au 4*f* peaks of the evaporated surface showing the experimental and the fitted spectra.

1. Evaporated gold as received and after 20 min UVO

Figure 3 shows the core levels of oxygen, carbon, and gold for the as-received (a) and the 20 min UVO (b) evaporated samples. No great modifications are recorded: both figures display relatively small amounts of carbon (39.5% before, 25.2% after UVO) and oxygen (12.8% before, 14.5% after UVO) as shown in Table II. Figure 4 shows the Au 4*f* doublet of the UVO treated surface. It displays the experimental curve and a peak reconstruction with Lorentzian/Gaussian components, as well as a tail to account for the asymmetry of the metallic peak. Gold oxide peaks have been shown in various work[18–22] at higher binding energies than metallic gold. Here no peak corresponding to oxide can be observed. This is in agreement with Vig[23] and Worley[24] who also do not report any oxidation of gold after short UVO exposure. Figure 5 compares the oxygen 1*s* peaks of the as-received evaporated gold and the UVO one. Both peaks have similar areas and comprise several unresolved components. The recompositions, after peak fitting, are summarized in Table III. The as-received sample oxygen shows three components: one low energy one at 530.19 eV (49% of the total 1*s* peak) and two higher binding energies ones at 531.36 eV (37%) and 533.23 eV (14%). The O 1*s* of the

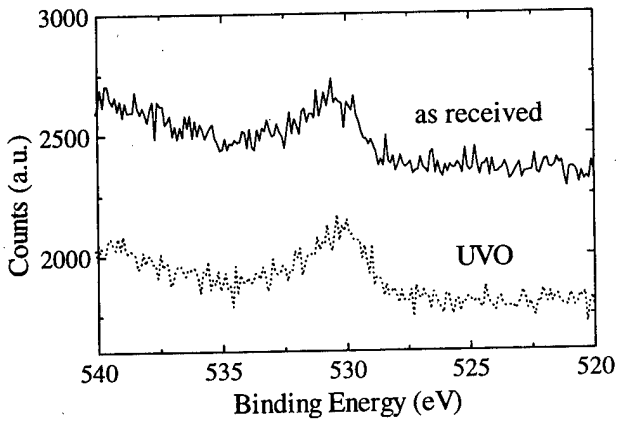


FIG. 5. XPS spectra of the oxygen 1*s* core level for as received and UVO evaporated gold surfaces.

TABLE III. Decomposition of the O 1*s* core level of the evaporated gold as-received and after 20 min UVO treatment: binding energies (eV) and % of the whole peak.

AR		UVO	
530.19	49%	530.07	45%
531.36	37%	531.18	28%
533.23	14%	532.43	17%
		533.84	10%

UVO treated sample shows the same low binding energy component at 530.07 eV (45%) as well as three other ones at 531.18 (28%), 532.43 (17%), and 533.84 eV (10%). In Table II we assign the low energy component to oxygen chemisorbed on gold (O···Au) and the higher energy ones to an overlayer of contamination involving C and H. Krozer *et al.*[21] report a similar low binding energy oxygen component which is not associated to any Au 4*f* gold oxide peak and which they assign to a form of chemisorbed oxygen. It is interesting to note that the carbon energy of the as evaporated sample is 284.6 eV as expected for contamination, while that of the UVO gold is 283.9 eV thus indicating that some carbon also has reacted with the metal.

The UVO (20 min) treatment on evaporated gold layers is seen to have two important effects: to diminish the quantity of carbon on the surface, though this quantity was already small before the treatment (39.5% before and 25.2% after), and to shift its binding energy towards lower values. No indication for the presence of metallic oxide is recorded; nonetheless a low binding energy component of the O 1*s* core level can be attributed to chemisorbed species.

2. Galvanic gold as received and after 20 min UVO

Figures 6(a) and 6(b) show the oxygen, carbon, sulfur, and gold core levels for galvanic gold samples as-received and after 20 min UVO treatment. Table II summarizes the surface compositions. Several features can be observed:

(1) as-received surfaces have a large contamination involving C and O: gold only represents 13.0% of the analyzed surface;
(2) after the UVO there is far less carbon (from 68.4% to 25.6%) but more oxygen (from 18.6% to 40.0%) and some sulfur (6.4%).

Figure 7 shows that, after the UVO treatment, the energy of the global O 1*s* peak is shifted to lower binding energies. This is further investigated by a peak fitting process of the O 1*s* level which allows us to recompose the unresolved peaks with elementary components. The results are summarized in Table IV. Oxygen on the galvanic gold surface is mainly due to C and H type contamination: the component at low binding energy represents only 6% of the total peak. On the other hand, after UVO, the low binding energies peaks at 529.71 and 530.72 eV represent 48% of the total quantity of oxygen. As in the case of the evaporated gold, the Au 4*f* doublet represented Fig. 8, together with the peak fit, does not indicate the presence of an oxide.

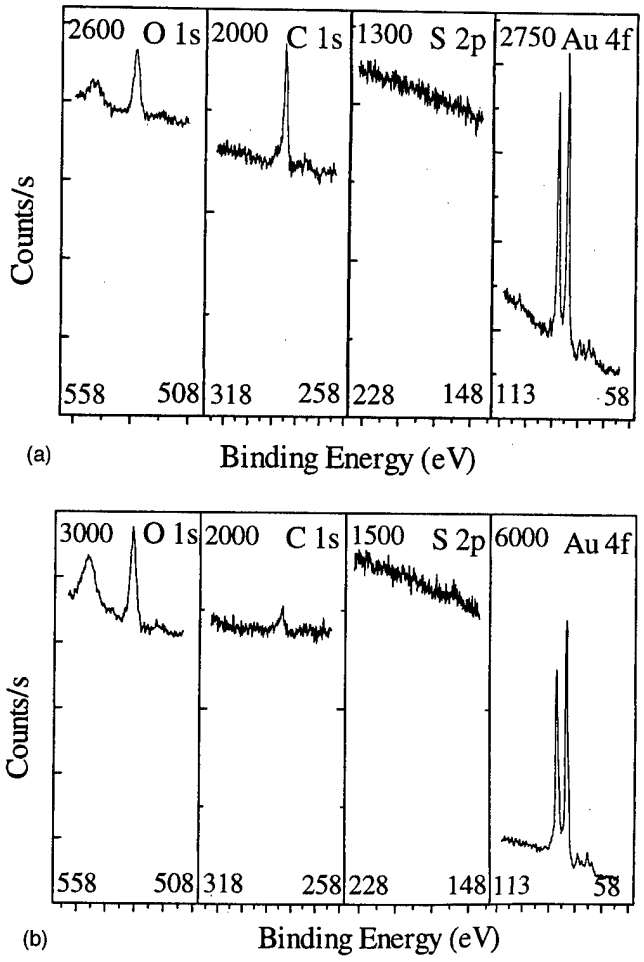


FIG. 6. XPS spectra showing the oxygen, carbon, sulfur, and gold core levels of galvanic gold surfaces: (a) as-received and (b) after 20 min UVO.

UVO (20 min) on galvanic gold removes most of the carbonaceous contamination. The global quantity of oxygen is seen to be increased by the treatment. Moreover, physisorbed oxygen seems to be replaced by some chemisorbed oxygen, which is not associated to a metallic oxide. Some sulfur is also detected probably due to the removal of the

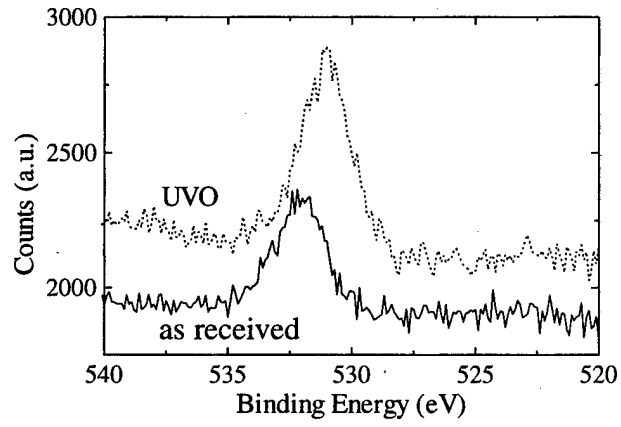


FIG. 7. XPS spectra of the oxygen 1*s* core level for as received and UVO galvanic gold surfaces.

TABLE IV. Decomposition of the O 1*s* core level of the galvanic gold as-received and after 20 min UVO treatment: binding energies (eV) and % of the whole peak.

AR		UVO	
530.17	6%	529.71	12%
531.41	36%	530.72	36%
532.30	41%	531.50	31%
533.40	17%	532.40	16%
		533.84	5%

large carbon contamination allowing a better analysis of the gold surface (the galvanic bath contains sulfur).

C. Synthesis of the experimental results

The main results from the two techniques can be synthesized as follows:

(1) galvanic gold, even after the UVO exposure, is more contaminated than the evaporated gold;
(2) UVO drastically increases the quantity of chemisorbed oxygen on the galvanic gold surface;
(3) the mean barrier height values vary with the total quantity of oxygen on the surface: the more oxygen on the surface, the higher the value of the mean barrier height (see Tables I and II);
(4) surfaces with the PFAE lubricant films have the larger values of the mean barrier height with the larger scatter as well.

Additional atomic force microscopy (AFM) investigations of the topography of the surfaces, before and after UVO, not reported here, have shown little differences. After UVO, the images are better defined, due to the removal of the contaminant, but no change in the general aspect of the morphology is observed.

IV. DISCUSSION

As other authors, we observe that the measured barrier heights are lowered with respect to the expected values for similar systems carefully cleaned and working in UHV (these expected values are about 5 eV). The majority of authors assign the barrier height lowering to surface contamination. However, the way by which surface contaminants lower the barrier height is still an open question.

An attempt to explain this effect by the fact that the stiffness k of the tip-sample system is not infinite, was presented by Meepagala and Real.[5] In other words, they assume that when a repulsive force ΔF—mediated by the contaminant—between tip and sample exists, the increase of the tip-sample distance is $\Delta F/k$. In such conditions, these authors show that the apparent mean barrier height $\bar{\Phi}_a$ is related to the mean barrier height $\bar{\Phi}$ in Eq. (1) by

$$\sqrt{\bar{\Phi}_a}=\sqrt{\bar{\Phi}}[1+(1/k)(dF/dz)], \tag{2}$$

where F is the force between the tip and the sample. When this force is repulsive, dF/dz is generally negative, and therefore, the apparent barrier height that is measured by the STM is lower than the true barrier height. In order to test this model, the authors perform an experiment on a home made STM/AFM. Their conclusion is that their explanation does not hold and that the tip-sample force gradient does not play a significant role in the lowering of the apparent barrier heights in STM measurements on gold in air.

Another interesting contribution to the problem is an article by Salmeron *et al.*[25] concerning the viscoelastic and electrical properties of self-assembled monolayers (SAMs) of alkanethiolates on gold. The authors present the results of an experiment in which a vibrating Pt–Rh cantilever is made to approach the SAM, with a bias voltage of 0.1 V applied between the tip and the gold substrate. The results indicate that the current increases very rapidly when the tip touches the methyl end groups of the thiolate molecules. To understand this mechanism, the authors present the results of theoretical calculations concerning electron transport between a gold tip and a gold surface with a straight chain $-S-(CH_2)_n-CH_3$ perpendicular to it, the tip being separated from or touching the methyl end group. Their conclusion is that the presence of the organic molecule greatly increases the tunneling probability by about half an order of magnitude per C atom.

Concerning the nature of the contaminant layer of evaporated gold, the results presented by Tonck *et al.*[4] must also be mentioned. Their study concerns the electrical and mechanical interaction between two very close macroscopic gold plated electrodes, one a sphere and the other a plane, studied in dry air. Their $I(z)$ and $V(I)$ curves attest to a tunneling transfer mechanism. By using expression (1), they find $\bar{\Phi}=1.32$ eV. By analyzing the electrical and mechanical results, they conclude that, although the experiments are carried out in dry air, a meniscus exists between the electrodes, and that this meniscus comprises water and hydrocarbons, but not pure water. This conclusion is drawn, among others, from a measure (not cited in the article) of the pull-off force

$$F_{p-o}=4\pi\gamma R \tag{3}$$

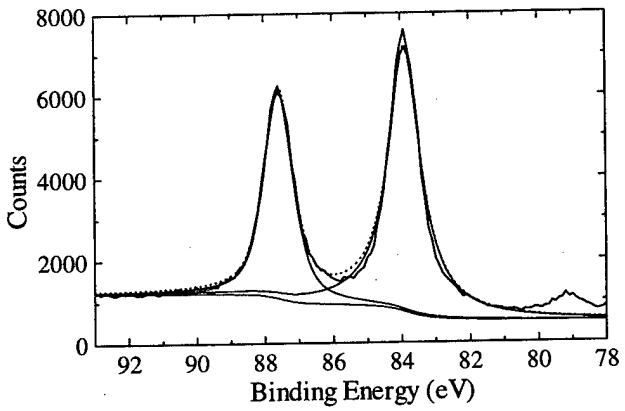


FIG. 8. XPS spectra of the Au 4*f* peaks of the galvanic surface showing the experimental and fitted spectra.

which appears at the mechanical separation of the contact. In Eq. (3), γ is the surface tension of the liquid and R the radius of the sphere.

In their article,[6] Song *et al.* find that the mean barrier height for Pt–Au systems in air is 0.76 eV. Thus, for the same material (vacuum deposited gold) for the plane, and metals with large work functions (Au, Pt, and PtIr) for the other electrode, Tonck *et al.*, Song *et al.*, and the authors of the present article find, in the same conditions (air), rather different values for the tunnel barrier height; namely: 1.32, 0.76, and 0.31 eV, respectively. Our opinion on this point is that the composition and concentration of contaminants lying on the gold surfaces are probably dependent upon the vacuum deposition system as well as upon the composition of the laboratory atmosphere. Hence, the comparison of barrier heights is pertinent only for samples prepared in a given laboratory, with the same vacuum system.

UVO has been shown to be a very effective way of cleaning surfaces[23,26] and is now largely used before thin film deposition. The mechanism of contaminant removal is attributed to the constant formation and destruction of ozone during which processes atomic oxygen is formed as an intermediate product. Hydrocarbons absorb the 253.7 nm radiation and are eliminated as gaseous products. Organic molecules can also react with atomic oxygen to give volatile species. Atomic oxygen is a strong oxidizing agent and several investigations have reported gold oxidation when the exposure times were long enough.[21,22]

In all the studies, gold surfaces are reported to be hydrophilic after exposure to UVO.[21,22,26] Comparing all the studies on surface contaminants, contact angles measurements and UVO cleaning can be confusing because of the multitude of experimental conditions involving exposure to the atmosphere contamination at various stages. It must be remembered that UVO cleaning of a surface makes it very active and that recontamination will occur when it is exposed to the atmosphere. The hydrophilic character of the UVO treated samples is usually correlated with the polar character of the surface oxygen (surface oxide or surface chemisorbed oxygen).

Our experiments are carried out in the laboratory air. Thus the surfaces of the nonlubricated samples are probably covered by two types of foreign compounds: interface (solid?) ones and liquid ones. Since surface analyses are performed under UHV, it is probable that only the elements constituting the interface (solid) foreign compounds are analyzed (Table II elements and hydrogen), most of the water or liquid hydrocarbons being pumped away and remaining just as monolayers. Considering at first the interface (solid) compounds, and in particular their oxygen concentrations, we observe, as previously mentioned, that the values for the nonlubricated samples appear in the same order as the mean barrier heights (see Tables I and II). In other words, the higher the oxygen content of the interface compound, the higher the value for the mean barrier height.

The nature of the liquid foreign compounds on the nonlubricated as received samples, can be assigned from the results presented in Refs. 4 and 22 to a mixture of water and hydrocarbons, the latter being in a larger proportion than the former. During the UVO treatment, the 184.9 nm wavelength transforms the ambient oxygen into both ozone (O_3) and atomic oxygen. Finally the largest wavelength (253.7 nm) interacts with ozone to give molecular and atomic oxygen. Most hydrocarbons will also absorb this UV radiation and volatile products are formed. The surface is thus cleaned from organic contaminants and the metal can be exposed to atomic oxygen which can chemisorb on the surface. According to many authors, the resulting surface is hydrophilic which implies that the liquid overlayer, after UVO, probably contains much more water than hydrocarbons.

The presence of the lubricant on the surface is seen to cause a significant increase of the barrier height values, particularly for the galvanic gold samples. In the latter case, some values for $\bar{\Phi}$ approach or exceed 2 eV, a very large value for measurements carried out in air. At this point, it is worth noting that the lubricant molecules contain fluorine and oxygen which are the most electronegative atoms of the Mendeleïev's table, whatever the chosen scale of electronegativity.

Considering now the six types of surfaces, one observes that the values for $\bar{\Phi}$ are less scattered for the as-received samples than for the UVO treated or lubricated ones. This means that the treated or lubricated surfaces are more heterogeneous. In particular, there are overlaps, for the values of $\bar{\Phi}$, for the unlubricated and lubricated surfaces. This may indicate that the lubricant does not cover all the surface. Hence, we can suppose that for surfaces well covered by the lubricant, the values for $\bar{\Phi}_{\text{mean}}$ would be higher than those presented in Table I. Thus, the main result of our experimental study is that the higher the content of the overlayer of the gold surface in electronegative elements is—here oxygen and fluorine—the higher the apparent barrier height is.

V. CONCLUSION

Though we have gained a good qualitative understanding of what occurs on a physico-chemical level, further work would be required for a quantitative explanation for the barrier height lowering. For this, we need to know (1) the topograpy of the active parts of the electrodes, (2) the exact nature and position of the chemical species (oxides, atoms, and molecules) between the electrodes, and (3) the manner in which the active interface deforms when the tip approaches the sample. Only in such conditions would it be possible to consider modeling the problem and deriving the theoretical $I(z)$ law. In our case, for the nonlubricated samples, one can only speculate about the nature of the interfacial fluid in which the tip approaches the sample. For non-UVO treated samples it may comprise hydrocarbon and water. For UVO treated samples, it may comprise mainly water, because the surface has become hydrophilic, and because water is produced very near the surface by the UVO treatment. For the lubricated samples, the results presented in Fig. 2 attest that the nature of the interface fluid has changed, a fact which was not evident *a priori*. Indeed, in tunneling current mea-

surements, the tip sample distance is very low, 5–10 Å, and the observed increase of the barrier height is due to a replacement, in a large proportion, of the fluid present on the surface by the lubricant. This means that the lubrication process used here is efficient.

In addition, we have pointed out that, in the same conditions, different values of the barrier height are observed by different authors. Our opinion is that, whatever the gold deposition process, the comparison of barrier heights is pertinent only for samples prepared in a given laboratory, with the same process.

From a semiquantitative point of view, two kinds of effects may be involved in the barrier height lowering process. First, as mentioned by Joachim and Sautet,[27] through-bond tunneling can occur along the hydrocarbon chains which are mainly present on the surfaces of our nontreated samples. This may result in most of the barrier height lowering observed with these samples. Second, electrostatic charge effects may also be involved. Indeed the fact that we find low binding energy oxygen atoms indicates that some oxygen atoms are negatively charged. This could explain why for UVO treated samples the barrier height lowering is less pronounced, since the presence of negative charges, within the barrier, reduces the tunneling transfer probability.

[1]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Appl. Phys. Lett. **40**, 178 (1982).
[2]H. J. Mamin, E. Ganz, D. W. Abraham, R. E. Thomson, and J. Clarke, Phys. Rev. B **34**, 9015 (1986).
[3]J. K. Gimzewski and R. Möller, Phys. Rev. B **36**, 1284 (1987).
[4]A. Tonck, F. Houzé, L. Boyer, J. L. Loubet, and J. M. Georges, J. Phys.: Condens. Matter **3**, 5195 (1991).
[5]S. C. Meepagala and F. Real, Phys. Rev. B **49**, 10 761 (1994).
[6]J. P. Song, K. A. Mørch, K. Carneiro, and A. R. Thölén, J. Vac. Sci. Technol. B **12**, 2237 (1994).
[7]L. Olesen, M. Brandbyge, M. R. Sørensen, K. W. Jacobsen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, Phys. Rev. Lett. **76**, 1485 (1996).
[8]M. Antler, Wear **6**, 44 (1963).
[9]M. Antler and M. H. Drozdowicz, Wear **74**, 27 (1982).
[10]C. A. Haque and T. A. Uhrig, IEEE Trans. Compon., Hybrids, Manuf. Technol. **7**, 76 (1984).
[11]M. Antler, IEEE Trans. Compon., Hybrids, Manuf. Technol. **10**, 24 (1987).
[12]J. Périé, M. Périé, L. Boyer, and S. Noël, Wear **128**, 153 (1988).
[13]S. Noël, L. Boyer, F. Houzé, J. Boissel, J. M. Bourin, and E. M. Zindine, Proceedings of the 16th International Conference on Electrical Contacts, 73, Loughborough, England, 1992 (unpublished).
[14]W. H. Abott, Proceedings of the 42nd IEEE Holm Conference on Electrical Contacts, 414, Chicago, IL, 1996 (unpublished).
[15]M. Antler, Proceedings of the 42nd IEEE Holm Conference on Electrical Contacts, 363, Chicago, IL, 1996 (unpublished).
[16]S. Noël, F. Houzé, L. Boyer, Z. Mekhalif, R. Caudano, and J. Delhalle, Proceedings of the 43rd IEEE Holm Conference on Electrical Contacts, Philadelphia, PA, 1997 (unpublished).
[17]S. Noël, L. Boyer, and C. Bodin, J. Vac. Sci. Technol. A **9**, 32 (1991).
[18]C. R. Aita and N. C. Tran, J. Vac. Sci. Technol. A **9**, 1498 (1991).
[19]J. J. Pireaux, M. Liehr, P. A. Thiry, J. P. Delrue, and R. Caudano, Surf. Sci. **141**, 221 (1984).
[20]T. Dickinson, A. F. Povey, and M. A. Sherwood, J. Chem. Soc., Faraday Trans. 1 **71**, 298 (1975).
[21]A. Krozer and M. Rodahl, J. Vac. Sci. Technol. A **15**, 1704 (1997).
[22]D. E. King, J. Vac. Sci. Technol. A **13**, 1247 (1995).
[23]J. R. Vig, J. Vac. Sci. Technol. A **3**, 1027 (1985).
[24]C. G. Worley and R. W. Linton, J. Vac. Sci. Technol. A **13**, 2281 (1995).
[25]M. Salmeron, G. Neubauer, A. Folch, M. Tomitori, D. F. Ogletree, and P. Sautet, Langmuir **9**, 3600 (1993).
[26]N. S. McIntyre, R. D. Davidson, T. L. Walzak, R. Williston, M. Westcott, and A. Pekarsky, J. Vac. Sci. Technol. A **9**, 1355 (1991).
[27]C. Joachim and P. Sautet, Scanning Tunneling Microscopy and Related Methods, 1990 (unpublished), p. 377.

Structural and electrical properties of chemical vapor deposition tungsten overgrowth on physical vapor deposited and metalorganic chemical vapor deposited TiN adhesion layers

Y. C. Peng and L. J. Chen[a)]
Department of Materials Science and Engineering, National Tsing Hua University, Hsinchu, Taiwan, Republic of China

W. Y. Hsieh, Y. R. Yang, and Y. F. Hsieh
United Microelectronics Corporation, Hsinchu, Taiwan, Republic of China

(Received 19 February 1998; accepted 22 May 1998)

The structure and electrical properties of chemical vapor deposited W (CVD-W) films on various physical vapor deposited or metalorganic chemical vapor deposited TiN films have been investigated. The growth orientations of the TiN adhesion layers were controlled by deposition method and film thickness. The growth orientations of CVD-W films were found to depend strongly on the microstructures of TiN. The grain sizes and electrical resistivity of CVD-W were found to increase and decrease, respectively, with the grain sizes of underlying TiN layers. © *1998 American Vacuum Society.* [S0734-211X(98)09704-2]

I. INTRODUCTION

As device dimensions scale down into the submicron regime, the development of new materials and processing technologies such as contact diffusion barrier, adhesion layer, and high aspect ratio via hole filling in multilevel metallizations is highly desired. Chemical vapor deposited tungsten (CVD-W) has been widely used in the very large scale integrated circuits (VLSI) multilevel metallization process because of its excellent electromigration resistance, good step coverage on high aspect ratio contact holes and vias as compared with other refractory metals. However, in order to improve the adhesion properties between W and SiO_2, a metal glue layer is required.[1–3]

Titanium nitride (TiN) has been known to be a very useful diffusion barrier layer and antireflecting layer for aluminum interconnect metallization in the VLSI manufacturing process.[2–8] TiN also meets the requirement for improving the adhesion properties between CVD-W and the SiO_2 insulating layer because it shows good adhesion to SiO_2 and the thermal stability temperature of the TiN/SiO_2 system is relatively high (>1000 °C). The W/TiN structure has been shown to be a stable low-resistance interconnection.[3] In order to improve the step coverage of conventional physical vapor deposited TiN (PVD-TiN) for submicron contacts filling, a variety of deposition methods, such as collimated physical vapor deposited and chemical vapor deposited TiN have been developed.[4,9–12] It is known that the crystal orientation of the TiN films is determined by the deposition method and film thickness. The microstructures and electrical properties of the subsequently deposited CVD-W could be influenced by the initial growth of the TiN adhesion layer. In this article, we report the results of an investigation on the microstructures and electrical properties of CVD-W films on TiN adhesion layers with different preferred growth orientations.

II. EXPERIMENTAL PROCEDURES

Single crystal, 3–5 Ω cm, 8 in. in diameter, phosphorus doped (001) oriented silicon wafers were used in this work. The wafers were first chemically cleaned by a conventional wet cleaning process (Pre-Metal Etching) before loading into a deposition chamber. Four sets of TiN samples deposited by conventional physical vapor deposition, collimated PVD, metalorganic chemical vapor deposition (MOCVD), and MOCVD with *in situ* plasma treatment were prepared.

A reactive dc magnetron sputtering system (AMT plasma system, model PVD HP-5500) was used to prepare the TiN films. A high-purity Ti target (99.995%) with a high basal plane ratio of texture was employed during sputtering. The base pressure of the chamber before deposition was kept below 1×10^{-8} Torr. During the sputtering process, a mixture of nitrogen (purity 99.99%) and argon (purity 99.999%) was fed into the vacuum chamber controlled by separate mass flow controllers. The sputtering power was kept at 8 kW and the total sputtering pressure for a mixture of N_2 and Ar gases was less than 1×10^{-3} Torr during sputtering. The thicknesses of the TiN films were selected to be 10, 50, and 100 nm. For the preparation of collimated PVD TiN samples, the PVD system as mentioned earlier with a honeycomb collimator was used. The aspect ratio of the honeycomb collimator is 1.5:1. The cell size is 5/8 in. In order to increase the deposition rate in the collimated sputtering process, the sputtering power was increased to about three times of that in the conventional PVD process. The substrate temperature of samples deposited by PVD and collimated PVD during film growth was maintained at room temperature. The x values of TiN_x, determined by x-ray fluorescence analysis, of various PVD and collimated PVD samples were kept between 1.08–

a)Electronic mail: ljchen@mse.nthu.edu.tw

TABLE I. Deposition conditions of TiN adhesion layers.

	Base pressure (Torr)	Pressure during deposition (Torr)	Gases used	Power (kW)	Deposition temperature (°C)
Conventional PVD	1×10^{-8}	1×10^{-3}	N_2+Ar	8	room temperature
Collimated PVD	1×10^{-8}	1×10^{-3}	N_2+Ar	22	room temperature
MOCVD TiN	1×10^{-6}	0.5–1	TDMAT+He		400–450
MOCVD TiN with plasma treatment	1×10^{-6}	0.5–1	TDMAT+He (NH_3+H_2 plasma treatment)		400–450

1.10. The resistivity of the PVD and collimated PVD deposited TiN was about 90 $\mu\Omega$ cm.

For the as-deposited MOCVD TiN films, the metalorganic precursor, tetrakis(dimethylamido)-titanium (TDMAT), was introduced into the AMT Centura system with He carrier gas. The deposition pressure was in the range of 0.5–1 Torr, and the deposition temperature was at 400–450 °C. An *in situ* plasma treatment with a mixture of NH_3 and H_2 gases was carried out for every 25-nm-thick TiN deposition for the plasma treated MOCVD TiN samples. Table I lists the deposition conditions for various samples. Auger electron spectroscopy (AES) concentration-depth profile analysis was used to analyze the atomic concentrations in the films. The resistivities of as-deposited MOCVD TiN and plasma treated MOCVD TiN films were about 200 and 120 $\mu\Omega$ cm, respectively.

After the TiN films were prepared, 200-nm-thick CVD-W films were deposited. The CVD-W films were grown by the reduction of WF_6 and SiH_4+H_2 at pressures of 1–1.5 Torr. The deposition temperature of CVD-W was in the range of 400–450 °C.

A high-power thin-film x-ray diffractometer with a wide-angle goniometer utilizing the parallel beam method was used to characterize the diffraction peaks of TiN thin films. The microstructures of TiN thin films were examined by both plan-view and cross-sectional transmission electron microscopy (TEM). The orientation relationships between TiN and (001)Si substrates were determined by analysis of the transmission electron diffraction patterns. The cross-sectional TEM images were taken along the [110] zone axis of the single-crystal Si substrate. A JEOL JEM-2010 TEM operating at 200 kV and a JSM-3620F scanning electron microscope (SEM) were used for TEM and SEM examinations, respectively. Four point probe was used to determine the sheet resistance of the thin films.

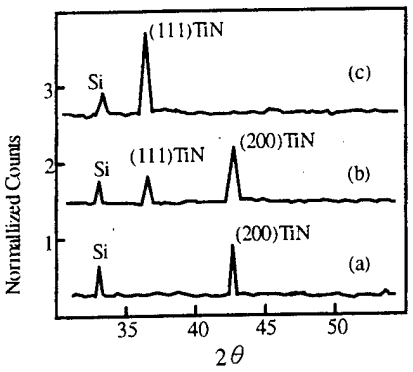


FIG. 1. XRD spectra of PVD TiN films with thicknesses of (a) 10 nm, (b) 50 nm, and (c) 100 nm.

III. RESULTS AND DISCUSSION

A. Growth orientation of TiN adhesion layer

The evolution of preferred crystal orientation from (100) to (111) was observed for PVD TiN with increasing film thickness. Figure 1 shows the x-ray diffraction (XRD) spectra of PVD TiN films with different thicknesses. For thin TiN films (<30 nm), the surface energy was known to be a dominant factor in determining the crystal orientation. Since the (100) plane has the lowest surface energy,[13] preferred (100) growth in thinner films could be achieved. However, as the film thickness was increased, the preferential growth of (111) TiN was found to occur with film thickness exceeding 30 nm. A (100)/(111) TiN two-layer structure was observed to grow sequentially in a 100-nm-thick PVD sample as determined by TEM diffraction analysis. An example is shown in Fig. 2. Since (111)TiN possesses the lowest strain energy, the transition from (100) to (111) growth orientation in thicker films could be explained as the relaxation or reduction of strain energy with increasing film thickness.[13] The conversion from (100) to (111)TiN preferred growth orientation was, therefore, attributed to the increase in strain energy with TiN film thickness.

Film deposition methods were also shown to influence significantly the microstructures of TiN films. Microstructural analysis was conducted for various 100-nm-thick TiN films deposited by conventional PVD, collimated PVD and MOCVD with the *in situ* plasma treatment. XRD spectra

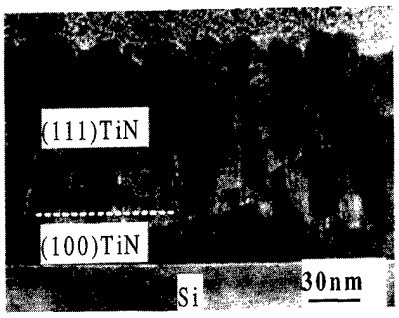


FIG. 2. Cross-sectional TEM micrograph of a 100-nm-thick PVD TiN film.

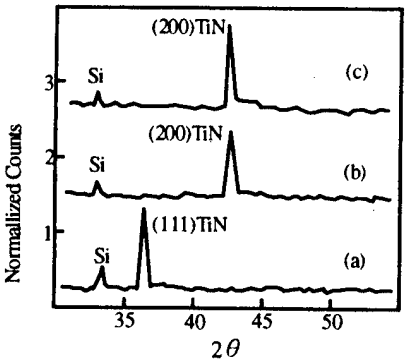


FIG. 3. XRD spectra of TiN films deposited by (a) conventional PVD, (b) collimated PVD, and (c) MOCVD with plasma treatment.

confirmed the preferential growth of (111)TiN in PVD samples as shown in Fig. 3. For the samples deposited by collimated PVD, the preferred growth orientation was (100) and did not change with the film thickness. Comparing the collimated PVD deposited TiN films with the conventional PVD deposited samples, the TiN grain sizes were relatively large in the collimated PVD deposited samples as seen in Table II. It is thought that the high-power process of collimated PVD provides TiN molecules with higher kinetic energy during film growth and leads to a structure with large grains which, in turn, impedes the change of preferred growth orientation with film thickness. Table III lists the stress data measured for conventional and collimated PVD deposited TiN. In Table II, it is seen that the stress value of collimated PVD deposited TiN is higher than the conventional PVD deposited samples as the film thickness reaches 60 nm. The result suggested that the impedance to the change of preferred orientation from (100) to (111)TiN by the formation of large grains in collimated PVD deposited samples led to significantly higher stress than that in conventional PVD deposited samples.

In the MOCVD process, high carbon contamination was found in the as-deposited MOCVD TiN films. The crystal structure of as-deposited MOCVD TiN films was amorphous. Carbon contamination in the MOCVD deposited TiN films was found to increase the resistivity of TiN films. The *in situ* plasma treatment was previously used to eliminate carbon contamination.[4] In the present study, carbon contamination was indeed reduced significantly by the *in situ* plasma treatment. Figures 4(a) and 4(b) show AES depth profiles of as-deposited and plasma treated MOCVD TiN samples, respectively. Preferred growth of (100)TiN film was found in the plasma treated MOCVD TiN film as shown in Fig. 3. The result indicated that the surface energy is the dominant factor in the growth of TiN films. It is thought that the intermittent plasma treatment either relieves the accumulation of strain energy or creates a favorable surface condition to facilitate the preferred (100)TiN film growth. Figures 5(a)–5(d) show the cross-sectional TEM micrographs of 100-nm-thick TiN samples deposited by different methods.

TABLE III. Stess measured for conventional and collimated PVD deposited TiN films.

Film thickness (nm)	60	40
Stress of conventional PVD deposited TiN (dyn/cm^2)	-6.49×10^9	-6.71×10^9
Stress of collimated PVD deposited TiN (dyn/cm^2)	-1.09×10^{10}	-6.72×10^{10}

B. Properties of CVD-W on TiN films deposited by different methods

The growth of CVD-W was found to depend strongly on the microstructures of the TiN underlayer. Figure 6 shows the XRD spectra of CVD-W on various TiN films. (110)W textured films were observed to grow on either the (111) or amorphous TiN underlayer, which were prepared by conventional PVD and MOCVD processes, respectively. In contrast, (100)W textured films were formed on (100) oriented TiN layers, which were subjected to either collimated high-power PVD or CVD with *in situ* plasma treatments. The electrical resistivity of the W overlayer is correlated to the grain size as will be discussed in a later section. Results of structural and electrical characterizations of the CVD-W overgrowth on TiN adhesion layers are summarized in Table II.

In order to understand the nucleation and growth mechanisms of the CVD-W layer, the interface structure and lattice mismatch of these heterostuctural metal systems are taken into consideration. Owing to the small lattice mismatch

TABLE II. Microstructural and electrical properties of CVD-W films on TiN adhesion layers deposited by different methods.

	CVD-W/PVD TiN	CVD-W/PVD TiN (collimated)	CVD-W/MOCVD TiN	CVD-W/MOCVD TiN (plasma treated)
Preferential growth orientation of TiN	(111)TiN	(100)TiN	Amorphous	(100)TiN
Average grain size of TiN (nm)	10–15	30–40	···	15–20
Preferential growth orientation of W	(110)W	(100)W	(110)W	(100)W
Average grain size of W (nm)	75	150	25	75
Resistivity ($\mu\Omega$ cm)	10.8	7.8	14.5	11.2

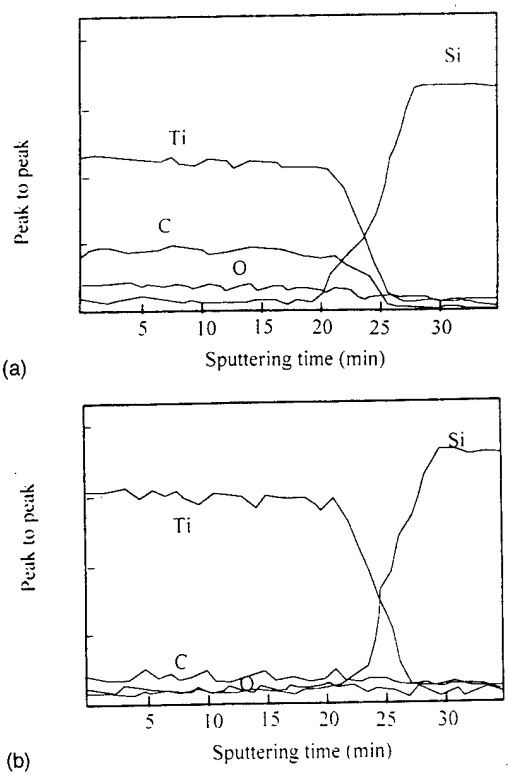


FIG. 4. AES depth profiles of (a) as-deposited MOCVD TiN and (b) MOCVD TiN with *in situ* plasma treatment.

(~0.5%) between (100)W and (100)TiN, the orientation relationships: (100)W||(100)TiN and [011]W||[011]TiN were realized in CVD-W on collimated PVD and plasma treated MOCVD TiN samples. The atomic arrangements of the crystal lattice of bcc (100)W and fcc (100)TiN are shown in Figs. 7(a) and 7(b), respectively. Figures 8(a) and 8(b) show the cross-sectional TEM (XTEM) micrographs of samples with [011]W||[011]TiN and (100)W||(100)TiN. The (110) textured W films are formed on MOCVD deposited amorphous TiN since (110)W is the most densely packed plane of bcc W

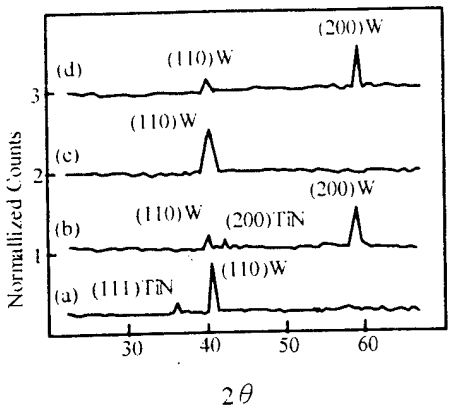


FIG. 6. XRD spectra of the metal layer systems with (a) W/conventional PVD TiN, (b) W/collimated PVD TiN, (c) W/as-deposited MOCVD TiN, and (d) W/MOCVD TiN with *in situ* plasma treatment.

structure and is expected to possess the lowest interface energy. Energetically, (111)W is expected to grow on (111)TiN with a flat surface owing to the small lattice mismatch (~0.5%). However, preferred [111] TiN growth was relatively rough and led to the growth of the most densely packed (110)W films on essentially randomly oriented TiN surfaces although it is textured. The average roughness, taken as the distance between the local maximum and minimum, was measured to be about 8 and 4 nm, respectively, from XTEM images for conventional and collimated PVD deposited 100-nm-thick TiN films.

The grain growth of CVD-W was affected by the interface structures of W and TiN, which would result in different electrical performance. The grain sizes of CVD-W were found to increase with the grain sizes of the underlying TiN layers. Plan-view SEM micrographs revealed that the average grain sizes are the largest and smallest on TiN samples deposited by collimated PVD and MOCVD without plasma treatment, respectively, as seen in Table II. Examples are shown in Figs. 9(a) and 9(b). The amorphous TiN layer pre-

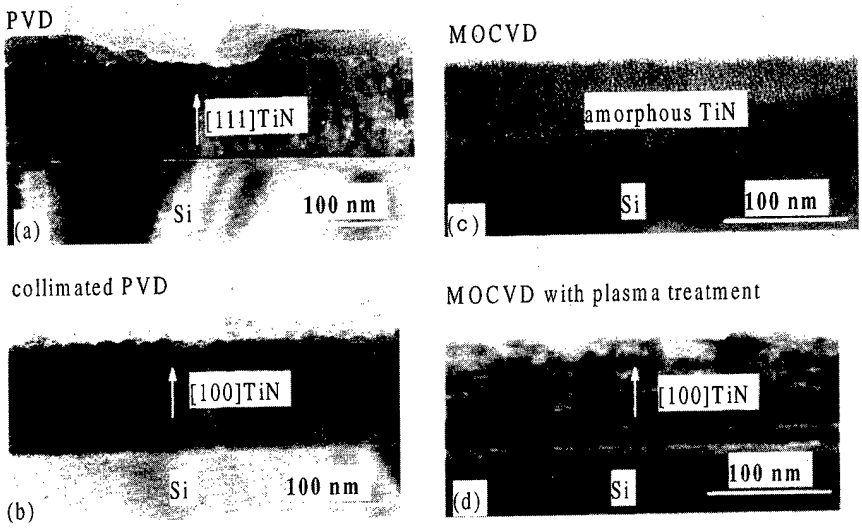


FIG. 5. Cross-sectional TEM images of (a) conventional PVD TiN, (b) collimated PVD TiN, (c) as-deposited MOCVD TiN, and (d) MOCVD TiN with *in situ* plasma treatment.

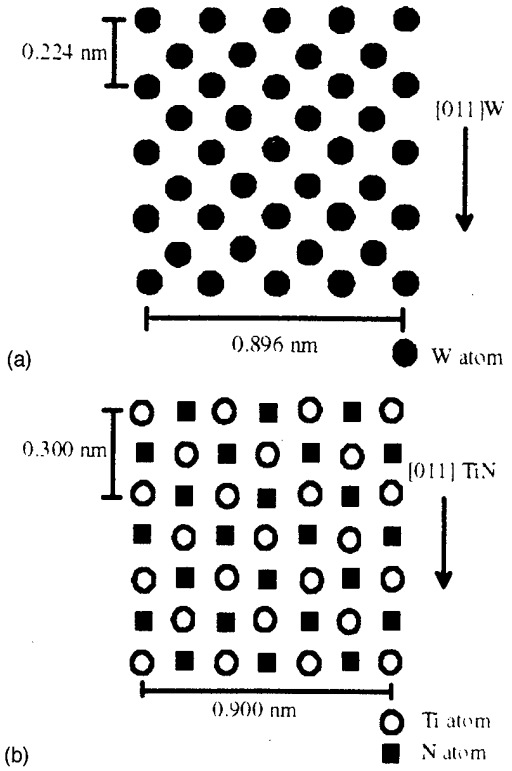


FIG. 7. Schematic diagrams of atomic arrangement in crystal lattices of (a) bcc (100)W and (b) fcc (100)TiN.

pared by MOCVD without plasma treatment apparently provides favorable nucleation sites for W grains. The impingement of more densely nucleated W grains eventually impedes the grain growth. Four point probe measurement

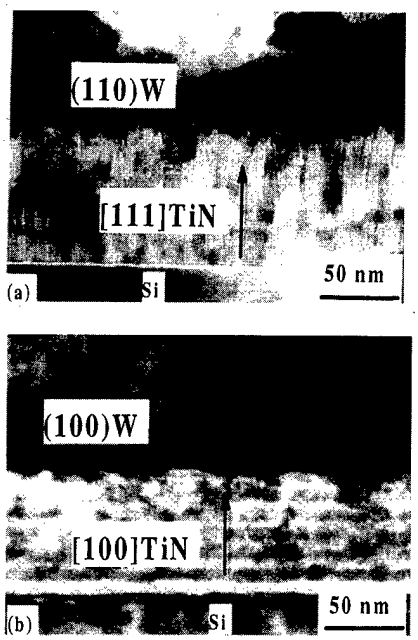


FIG. 8. Cross-sectional TEM micrographs showing the interface structures of (a) CVD-W/(111)TiN deposited by conventional PVD and (b) CVD-W/(100)TiN deposited by MOCVD with *in situ* plasma treatment.

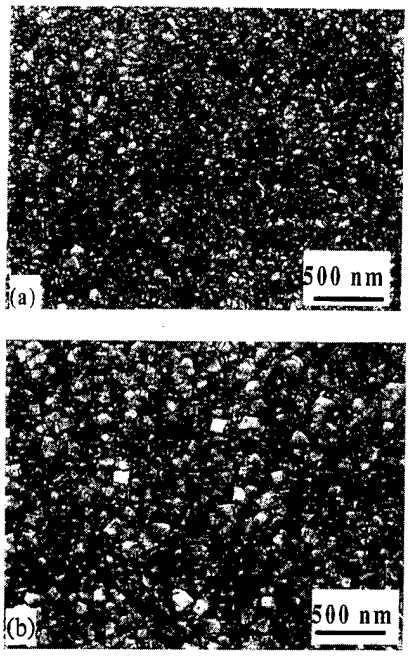


FIG. 9. Plan-view SEM micrographs of (a) (110)W/(111)TiN deposited by conventional PVD and (b) (100)W/(100)TiN deposited by collimated PVD.

showed that the resistivities of the CVD-W grown on different TiN films decrease with the grain sizes of the underlying layers. Since the mean-free path of single-crystal W is about 40 nm, the effect is attributed to the longer mean-free path of electrons and less grain boundary interference in the large grain structures.[14]

IV. SUMMARY AND CONCLUSIONS

The microstructures and electrical properties of CVD-W films grown on TiN adhesion layers deposited by different methods have been investigated. The growth orientations of the TiN adhesion layers were controlled by deposition method and film thickness. Preferred growth orientation of conventional PVD deposited TiN thin films was found to vary with the film thickness. For thin-film samples (10–50 nm), preferred [100] TiN growth is attributed to the possession of the lowest surface energy among all crystal planes. As the film thickness was increased, the strain energy becomes the dominant factor and leads to preferred [111] TiN growth.

For the collimated PVD deposited TiN, only the [100]TiN preferred growth was found. The high-power process during collimated PVD deposition is thought to result in the large grain TiN structure, which in turn, impedes the change of preferred orientation.

For the as-deposited MOCVD TiN film, the crystal structure was found to be amorphous because of the high carbon contamination in the sample. The carbon contamination was reduced by intermittent *in situ* plasma treatment. The preferred growth orientation of the plasma treated MOCVD TiN

film was [100]TiN, which is attributed to the relief of stress during the plasma treatment and the surface energy factor is dominating the film growth.

The growth orientations of CVD-W films were found to depend strongly on the microstructures of TiN. For CVD-W on (111) orientated TiN (deposited by conventional PVD) and amorphous TiN (deposited by MOCVD without *in situ* plasma treatment), the preferred growth orientation of W is (110)W. The most densely packed plane of bcc W structure, (110)W, is of the lowest interface energy on these TiN substrates. In contrast, the preferred orientation (100)W was found in the samples of CVD-W on (100)TiN deposited by collimated PVD or CVD with plasma treatment processes. The small lattice mismatch between [110]W and [110]TiN apparently played a dominant role in the growth of (100)W on the (100)TiN substrate.

The grain sizes and electrical resistivity of CVD-W were found to increase and decrease, respectively, with the grain sizes of the underlying TiN layers. The electrical resistivity of W films was found to be the lowest in films deposited by high-power collimated sputtering. The correlation between electrical resistivity and grain size can be understood since the mean-free path of electrons in W films is about 40 nm.

ACKNOWLEDGMENTS

The research was supported by the National Science Council through Grant No. NSC86-2215-E007-015 and the United Microelectronics Corporation.

[1]E. S. Kim, C. M. Lee, J. G. Lee, and H. B. Im, Mater. Sci. Eng., B **17**, 137 (1993).
[2]S. Kumar, D. R. Chopra, and G. C. Smith, J. Vac. Sci. Technol. B **11**, 1815 (1993).
[3]Y. Nakasaki, K. Suguro, and M. Kashiwagi, J. Appl. Phys. **64**, 3263 (1988).
[4]A. J. Konecni, G. A. Dixit, J. D. Luttmer, and R. H. Havemann, Proceedings of the VLSI Multilevel Interconnection Conference, Santa Clara, CA, 1996, p. 181.
[5]C. Y. Ting, Thin Solid Films **119**, 11 (1984).
[6]C. Y. Ting, J. Vac. Sci. Technol. **21**, 14 (1992).
[7]M. Wittmer and H. Melchoir, Thin Solid Films **93**, 397 (1982).
[8]M. Wittmer, J. Vac. Sci. Technol. A **2**, 273 (1984).
[9]M. Biberger, S. Jackson, G. Tkack, J. Schlueter, B. Janes, C. K. Huang, and L. Ouellet, Thin Solid Films **270**, 522 (1995).
[10]A. Katz, A. Feingold, S. J. Pearton, S. Nakahara, M. Ellington, U. K. Chakrabarti, M. Gera, and E. Lane, J. Appl. Phys. **70**, 3666 (1991).
[11]S. R. Kurtz and R. G. Gordon, Thin Solid Films **140**, 277 (1986).
[12]I. J. Raaijmakers, Thin Solid Films **247**, 85 (1994).
[13]J. Pelleg, L. Z. Zevin, S. Lungo, and N. Croitoru, Thin Solid Films **197**, 129 (1991).
[14]J. E. J. Schmitz, *Chemical Vapor Deposition of Tungsten and Tungsten Silicides* (Noyes, Park Ridge, NJ, 1992), p. 107.

Characterization of TiN barriers against Cu diffusion by capacitance–voltage measurement

Sa-Kyun Rha, Seung-Yun Lee,[a] Won-Jun Lee, Yong-Sup Hwang, and Chong-Ook Park
Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea

Dong-Won Kim
Department of Materials and Science and Engineering, Kyonggi University, Suwon 440-760, Korea

Youn-Seoung Lee and Chung-Nam Whang
Department of Physics, Yonsei University, Seoul 120-749, Korea

(Received 20 January 1998; accepted 29 May 1998)

Sputtered TiN was studied as a diffusion barrier in Cu/TiN/Ti/Si and Cu/TiN/Ti/SiO_2/Si multilayer structures using various characterization methods, and their sensitivities for detecting breakdown of the barrier were compared. It was confirmed by scanning electron microscopy and Auger electron spectroscopy that breakdown of the TiN barrier occurred through out-diffusion of Si in addition to in-diffusion of Cu. Breakdown temperatures varied by more than 100 °C depending on characterization methods, and capacitance–voltage ($C-V$) measurement was most sensitive for detecting the failure of the TiN barrier. The effects of rapid thermal annealing (RTA) on barrier properties of TiN were investigated, and it was found by $C-V$ measurement that the TiN(400 nm) RTA treated at 700 °C in a NH_3 ambient was stable up to 590 °C for 2 h, while the reference TiN (400 nm) was stable up to 450 °C for 2 h. © *1998 American Vacuum Society.*
[S0734-211X(98)13204-3]

I. INTRODUCTION

Cu is an attractive material for interconnection metallization in ultralarge scale integrated (ULSI) devices due to its low resistivity (1.67 $\mu\Omega$ cm for bulk), high reliability against electromigration (EM), and to the feasibility of chemical vapor deposition using organometallic precursors.[1,2] However, there are still many critical issues which remain to be resolved for its use in semiconductor metallization. For example, a thin diffusion barrier preventing Cu diffusion is needed because Cu is known as a fatal impurity in the semiconductor fabrication process. Cu is known to have high diffusivity in silicon and silicon oxide, and when it is dissolved into silicon crystals at interstitial sites, it becomes a deep level dopant[3,4] or forms neutral B–Cu complexes in the case of boron doped silicon.[5,6] These deep levels can act as generation–recombination centers, which decrease carrier lifetime. The formation of B–Cu complexes decreases active concentration of boron in silicon.

TiN is presently one of the most widely used barrier materials in Cu metallization, as well as in aluminum-based metallization.[7–11] Reported failure temperatures of TiN against Cu diffusion differ among researchers; for example, 400–800 °C for furnace annealing,[10–14] and 900 °C for rapid thermal annealing (RTA).[7] The difference in failure temperatures is caused by various film microstructures and characterization methods.

In the present study, the diffusion barrier properties of TiN in multilayer structures encountered in real devices—Cu/TiN/Ti/Si in the case of Cu contact to the Si substrate, and Cu/TiN/Ti/SiO_2/Si in the case of the field outside of the contact hole[15] —were comparatively analyzed using various characterization methods, and the sensitivities of these methods were compared. On the basis of these results, we investigated the effect of RTA temperature and ambient on the barrier properties of TiN.

II. EXPERIMENT

Phosphorus-doped n-type silicon (100) wafers and boron-doped p-type silicon (100) wafers were used as substrates. For samples with a Cu/TiN/Ti/SiO_2/Si structure, 100 nm thick SiO_2 layers were formed on Si wafers using the thermal oxidation method at 900 °C before TiN/Ti deposition. Ti films 20 nm thick were deposited in Ar with total pressure of 3 mTorr and rf power of 3 kW. TiN films were made by reactively sputtering Ti in an ambient of an Ar and N_2 mixture with total pressure of 3.3 mTorr and rf power of 6.5 kW. The temperature of the substrate was maintained at 400 °C for both Ti and TiN deposition. The variation in TiN density was achieved by adjusting the gas ratio of Ar and N_2. Gas ratios of Ar:N_2 were 25:60 for the reference TiN, 25:45 for the Ti-rich TiN, and 25:90 for the N-rich TiN. The thickness of TiN was controlled within 40±3 nm. The density of TiN was calculated from the film thickness and planar density of TiN obtained by Rutherford backscattering spectrometry (RBS). For the investigation of the effect of RTA on the barrier properties of the reference TiN, samples with TiN/Ti/Si and TiN/Ti/SiO_2/Si structures were annealed for 1 min in a tungsten–halogen lamp heated RTA system. The RTA ambients were NH_3, N_2, and H_2, respectively, and the RTA temperature varied from 500 to 700 °C. Finally, Cu films 300 nm thick were deposited by chemical vapor deposition (CVD) using copper(I) hexafluoroacetylacetonate trimeth-

[a]Electronic mail: seungyun@cais.kaist.ac.kr

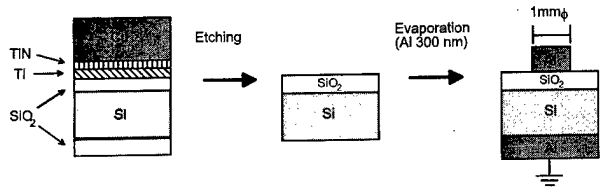


FIG. 1. Schematics of MOS structure for $C-V$ measurements.

ylvinylsilane, Cu(hfac)(tmvs). The substrate temperature and total pressure for the Cu deposition were 180 °C and 0.5 Torr, respectively. Details of the CVD system and the deposition procedure have been reported elsewhere.[16] After Cu deposition, annealing was performed at 1 atm in a resistance heated tube furnace with a flow of H_2(10%)/Ar for 2 h. The annealing temperature varied from 425 to 700 °C, with an interval of 25 °C.

Various characterization methods were employed to investigate the barrier characteristics of TiN: scanning electron microscopy (SEM) observation, sheet resistance measurement, depth profiling by Auger electron spectroscopy (AES), and capacitance–voltage ($C-V$) measurement. For $C-V$ measurement, an Al/SiO_2/Si capacitor structure was created by depositing Al after etching Cu and TiN/Ti layers as shown in Fig. 1. $C-V$ measurement was performed by an HP4192A LF impedance analyzer, where the input voltage frequency was fixed at 10 kHz with a sweeping rate of 0.2 V/s.

III. RESULTS AND DISCUSSION

RBS spectra of TiN films deposited by sputtering with various flow rates of N_2 gas were obtained as shown in Fig. 2. The Ti:N ratios were measured 1:1.02 for the reference TiN (N_2:60 sccm); 1:0.98 for the Ti-rich TiN (N_2:45 sccm); and 1:1.04 for the N-rich TiN (N_2:90 sccm). Formation of the TiN phase was confirmed by x-ray photoelectron spectroscopy (XPS) as shown in Fig. 3, where the binding energy shift of the Ti($2p$) core levels of the three TiN films could be seen. Considering that there is a difference in TiN composition of 6% between Ti-rich TiN and N-rich TiN, and a difference in density of 14%, it seems that the effects of density outweigh the effects of composition. Park *et al.*[17] reported that the high density of TiN is of primary importance in achieving good diffusion barrier performance between Cu and Si. Other properties of various TiN films were also measured as listed in Table I.

Figure 4 shows SEM photographs of the Cu surfaces of the Cu/TiN/Ti/Si multilayer structures after annealing at various temperatures. No shape changes were observed when Ti-rich TiN was annealed up to 525 °C, as shown in Figs. 4(a), 4(b), and 4(c). At annealing temperatures higher than 550 °C, pinholes began to appear on the Cu surface, and its surface morphology became rough, as shown in Fig. 4(d). As shown in Figs. 4(f) and 4(j), respectively, at 500 °C pinholes began to appear in the reference and N-rich TiN. Further temperature increases caused an increase in the number and

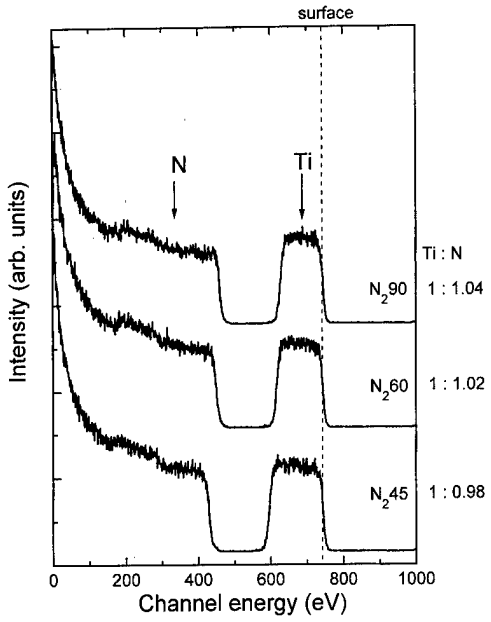


FIG. 2. RBS spectra of TiN films deposited by sputtering with varying N_2 gas flow rate (sccm) at 25 sccm Ar flow.

the size of pinholes, and the surface became much rougher as shown in Figs. 4(g)–4(h) and Figs. 4(k)–4(l).

The failure spot on the Cu surface after annealing at 500 °C was examined as shown in Fig. 5(a), and AES spectra were obtained from positions A and B [Figs. 5(b) and 5(c)]. While the spectrum obtained from point A indicated the existence of Cu, that from B showed Si and O, indicating that Si diffused out to the Cu surface at point B. AES depth profilings were performed at points A and B as shown in Fig. 6. The diffusion of Cu at the Ti/Si interface was not observed in the AES depth profile of point A. However, a large

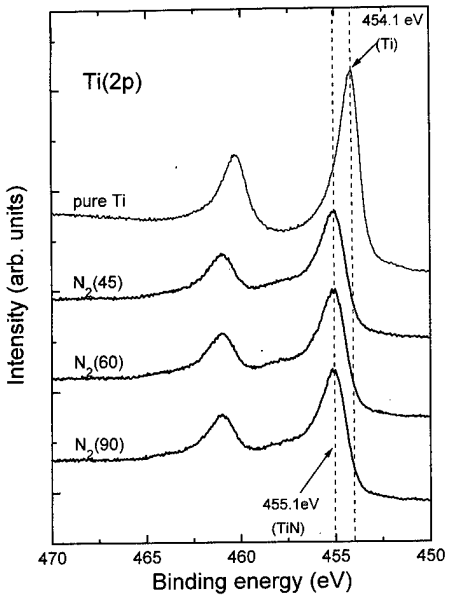


FIG. 3. XPS spectra of TiN films deposited by sputtering with varying N_2 gas flow rate (sccm) at 25 sccm Ar flow.

TABLE I. Properties of TiN film with various Ti:N ratios.

	Ti-rich	Reference	N-rich
$Ar:N_2$ ratio during sputtering	25:45	25:60	25:90
Thickness	41.8±1.0 nm	40.4±0.8 nm	38.6±0.8 nm
Density	4.94 g/cm^3	4.47 g/cm^3	4.24 g/cm^3
Sheet resistance	16.61 Ω/□	17.12 Ω/□	18.49 Ω/□
Stoichiometry of TiN	Ti:N=1:0.98	Ti:N′=1:1.02	Ti:N=1:1.04

amount of Cu trace was found at position B. From these results, it can be concluded that the breakdown of the TiN barrier occurs through out-diffusion of Si, as well as in-diffusion of Cu. According to Chamberlain,[18] Cu diffuses primarily through the grain boundary in TiN. Upon annealing, pinholes began to appear on a small portion of the Cu surface, their size ranging from 2 to 50 μm as shown in Figs. 4(f) and 5(a). Considering that 40 nm thick TiN had a grain size of less than 40 nm, it is thought that the diffusion of Cu proceeds mainly along weak regions in the grain boundary.

AES analyses of the Cu surfaces were performed to detect the breakdown of the TiN barrier. The Si peak was observed on the Cu surface of the Ti-rich TiN when it was annealed at 525 °C. However, reference and N-rich TiN showed the Si peak when annealed at 500 and 475 °C, respectively.

Sheet resistance was also measured to detect TiN barrier breakdown. As annealing temperature increased sheet resistance gradually decreased due to the reduction of crystal defects and grain growth in the Cu film, but resistance abruptly increased at some annealing temperatures because of the silicidation of Cu with the out-diffused Si and the decrease in Cu volume. While Cu diffusion through the Ti-rich TiN layer was observed at an annealing temperature of 550 °C through SEM characterization, and at 525 °C through AES

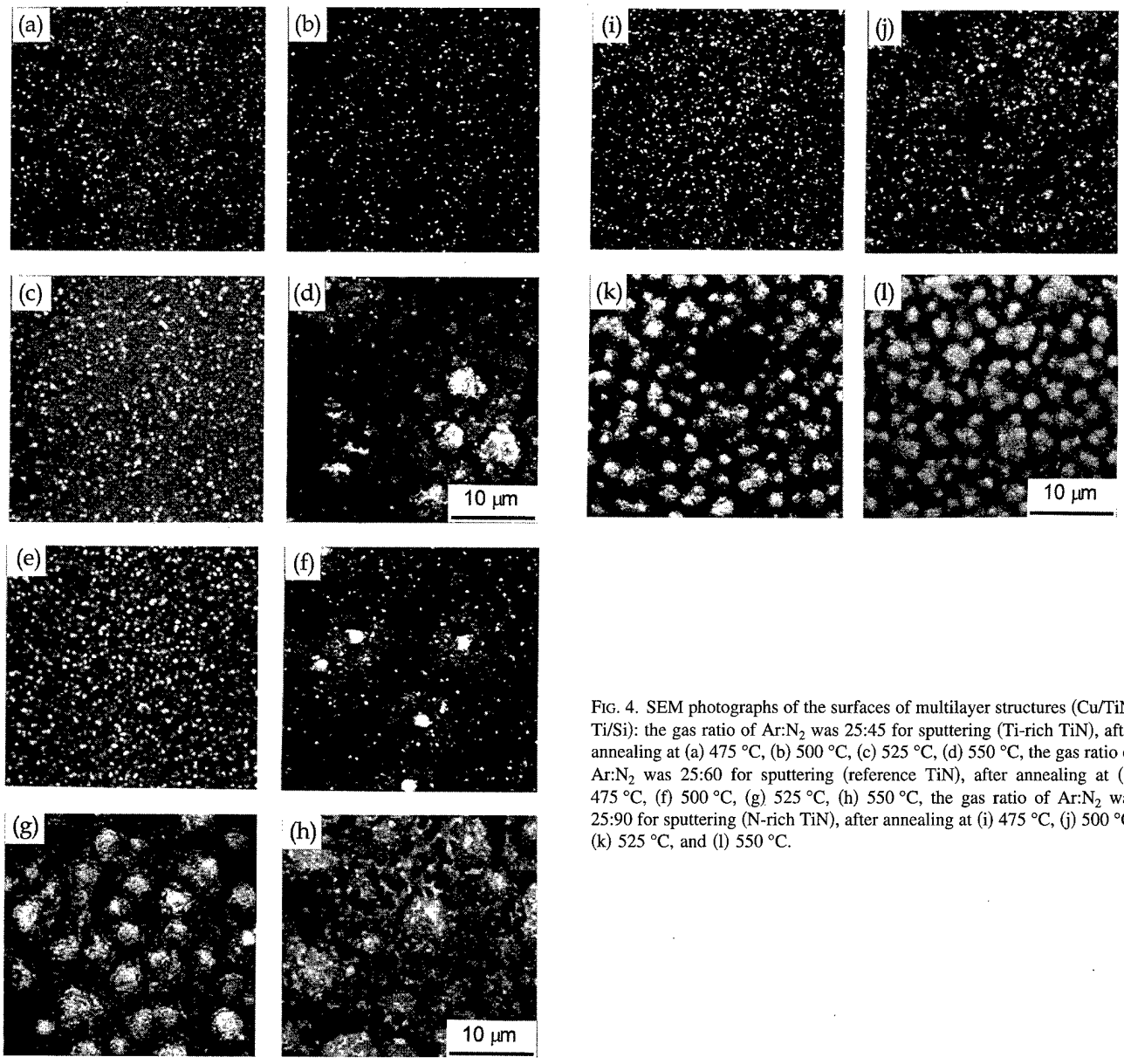


FIG. 4. SEM photographs of the surfaces of multilayer structures (Cu/TiN/Ti/Si): the gas ratio of $Ar:N_2$ was 25:45 for sputtering (Ti-rich TiN), after annealing at (a) 475 °C, (b) 500 °C, (c) 525 °C, (d) 550 °C, the gas ratio of $Ar:N_2$ was 25:60 for sputtering (reference TiN), after annealing at (e) 475 °C, (f) 500 °C, (g) 525 °C, (h) 550 °C, the gas ratio of $Ar:N_2$ was 25:90 for sputtering (N-rich TiN), after annealing at (i) 475 °C, (j) 500 °C, (k) 525 °C, and (l) 550 °C.

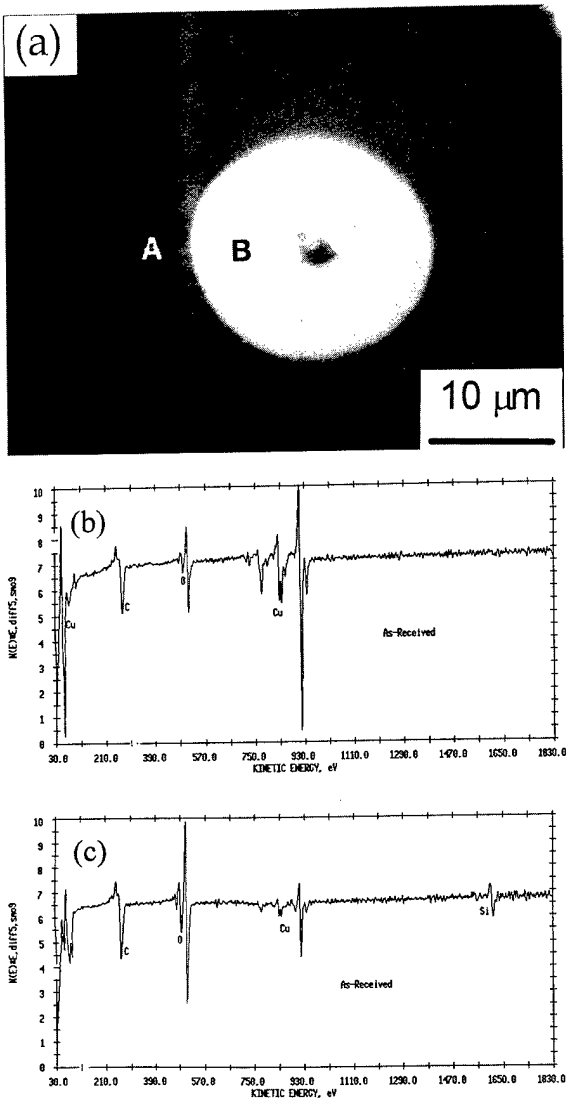


FIG. 5. Cu/reference TiN/Ti/Si structure was annealed at 500 °C. (a) SEM photograph of the "white spot" on the surface of the Cu film; AES spectra obtained from (b) point A and from (c) point B in the SEM photograph.

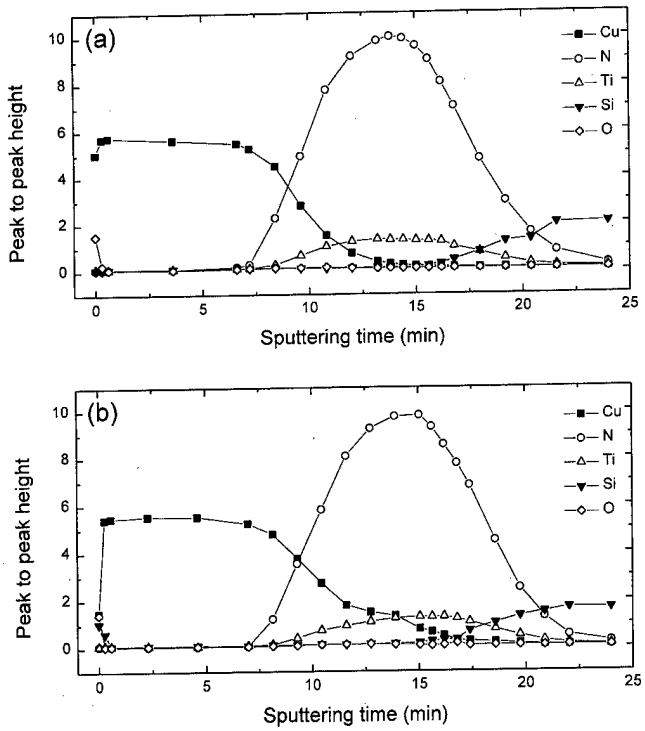


FIG. 6. AES depth profiles obtained from (a) point A and from (b) point B in Fig. 5(a).

characterization, sheet resistance increased after annealing at temperatures between 575 and 600 °C. For the reference TiN and the N-rich TiN, sheet resistance increases were observed at 525 and 500 °C, respectively.

Figure 7 shows the $C-V$ plot obtained from the metal–oxide-semiconductor MOS capacitor of Al/SiO_2/p-type Si/Al, which was made from a structure of Cu/TiN/Ti/SiO_2/p-type Si annealed at various temperatures. Using p-type Si wafers, we observed a decrease in inversion capacitance with annealing temperature and obtained a straight line dissimilar to a typical $C-V$ plot at 800 °C. As previously mentioned, the formation of B–Cu complexes caused by diffused Cu results in a decrease in active boron concentration.[5,6] Since capacitance across the surface depletion region is proportional to $\sqrt{}$ active concentration,[19] inversion capacitance decreases with decreasing active concentration (Fig. 7). Leakage current measured at 800 °C was more than two orders larger than those measured at other annealing temperatures, so it is thought that the oxide does not work as an insulator at 800 °C. While the $C-V$ plots are identical regardless of measuring frequency in the case of p-type Si wafers, $C-V$ characteristics varied with measuring frequency in the case of n-type Si wafers (Fig. 8). The difference in $C-V$ characteristics may be caused by the difference in doping concentrations and minority-carrier lifetime between boron-doped and phosphorus-doped Si wafers. It is well known that Cu in Si produces a deep donor level of 0.49 eV above the valence band. According to the Shockley–

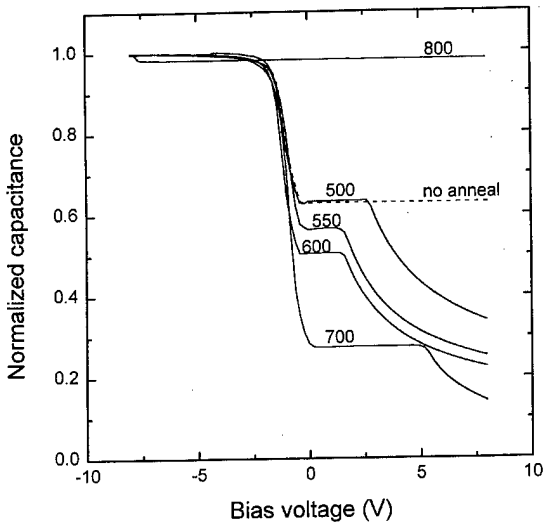


FIG. 7. $C-V$ plots obtained from MOS capacitor of Cu/reference TiN/Ti/SiO_2/p-type Si structure.

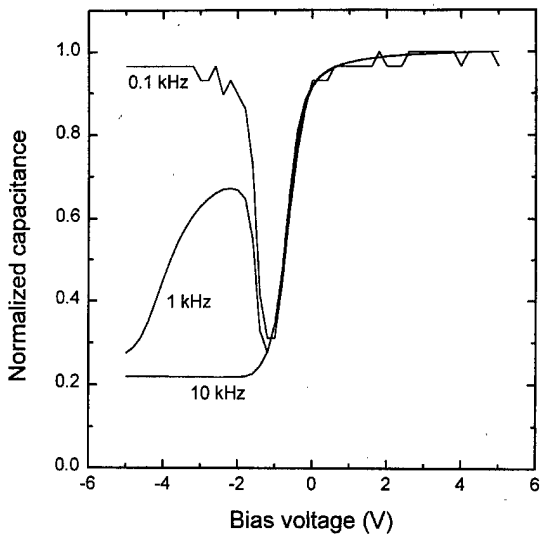


FIG. 8. $C-V$ plots obtained from MOS capacitor of Cu/reference TiN/Ti/SiO_2/n-type Si structure.

Hall–Read model,[20,21] these generation–recombination centers near intrinsic levels can lower carrier lifetime in such a way that inversion capacitance increases up to oxide capacitance, similar to the low frequency sweeping characteristics of the $C-V$ plot.[4] Therefore $C-V$ measurement can be used as a sensitive and preferable testing method to detect the breakdown of the Cu diffusion barrier, even though formation of the MOS structure is required for a test. Figure 9 shows the $C-V$ plot obtained from the MOS capacitor of Al/SiO_2/n-type Si/Al. As annealing temperature increases, inversion capacitance also increases, due to the diffused Cu which serves as a generation–recombination center. When Ti-rich TiN was annealed, Cu diffusion into the Si substrate was detected at 500 °C as shown in Fig. 9(a). In the case of reference TiN and N-rich TiN, inversion capacitance began to increase at 475 and 450 °C, respectively.

Breakdown temperatures of the three different TiNs were measured by various methods, as shown in Fig. 10. The criteria for each measurement were determined as follows: in SEM observation, the temperature at which the pinholes appeared; in AES analysis, the temperature at which Si peaks were observed on the sample surface or where Cu peaks were observed in the Si substrate; in sheet resistance measurement, the temperature at which sheet resistance began to increase; in $C-V$ measurement, the temperature at which inversion capacitance began to increase. It is consistent that breakdown temperature increases with the increasing density of TiN, irrespective of detection method. It can also be seen that $C-V$ measurement detects the lowest breakdown temperature among the tests. AES analysis shows lower breakdown temperatures than SEM observations, which is more sensitive than sheet resistance measurement. Previously, many authors reported different breakdown temperatures for TiN barriers;[7,10–14] however, as shown in Fig. 10, breakdown temperatures can vary by more than 100 °C, depending on the testing methods used.

The breakdown temperature of the RTA-treated TiN was

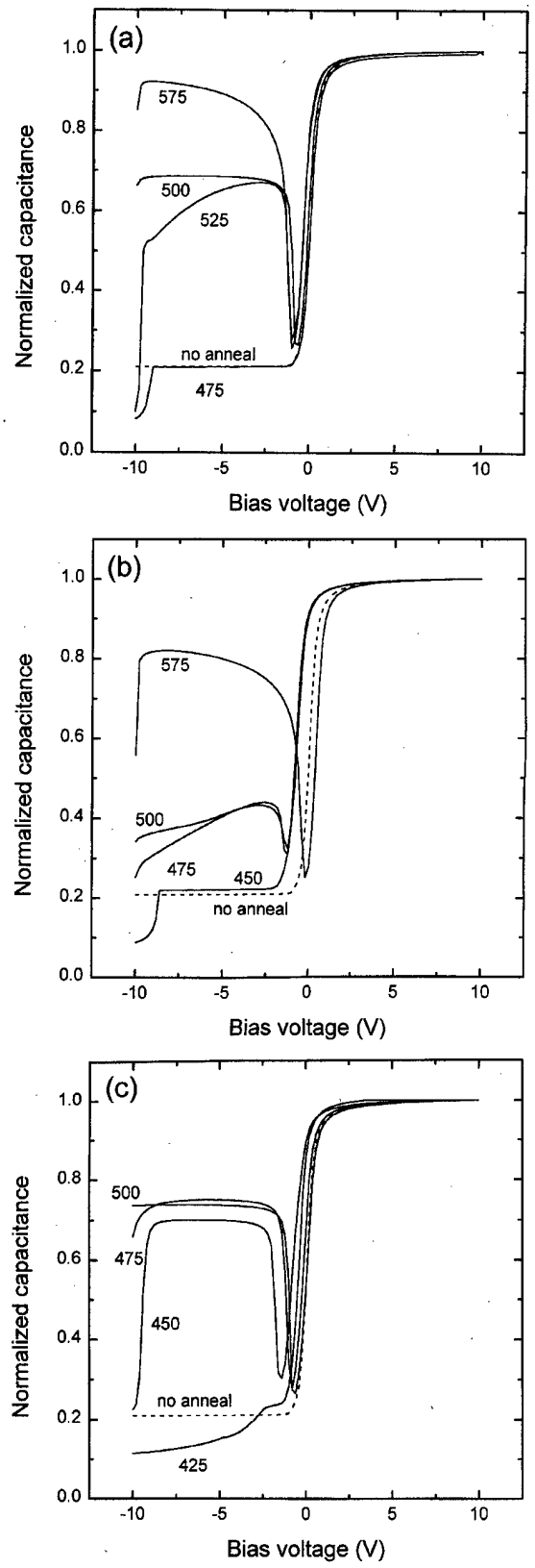


FIG. 9. $C-V$ plots obtained from MOS capacitor of Cu/TiN/Ti/SiO_2/n-type Si structure: (a) Ti-rich TiN; (b) reference TiN; and (c) N-rich TiN.

measured using the previously mentioned characterization methods, and it was confirmed again that $C-V$ measurement was the most sensitive method. As shown in Fig. 11, the breakdown of TiN RTA treated at 500 °C in H_2/Ar ambient

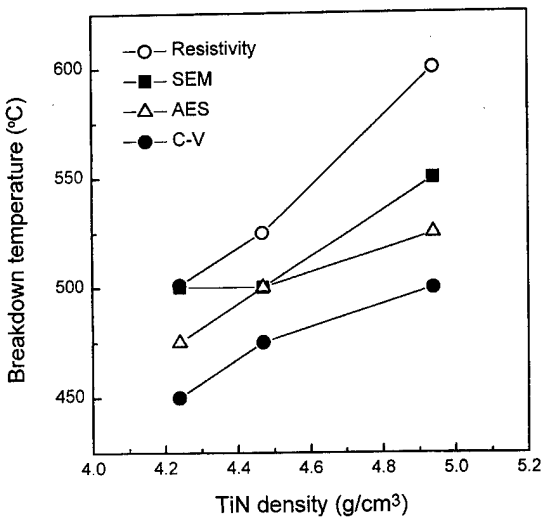


FIG. 10. Breakdown temperatures of the three TiNs measured by various characterization methods.

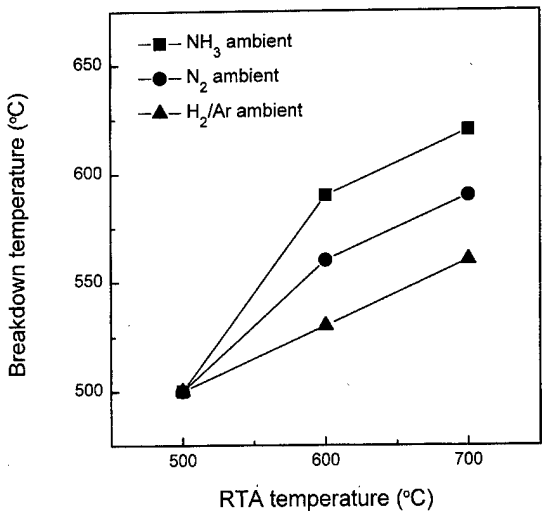


FIG. 12. Breakdown temperatures of TiN RTA treated with various ambients and temperatures measured by $C-V$ measurement.

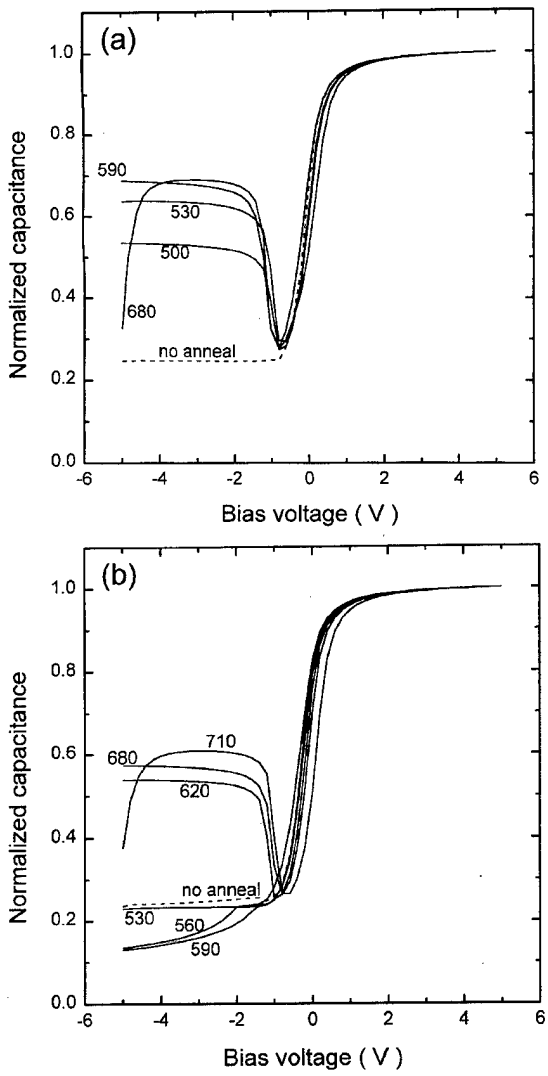


FIG. 11. $C-V$ plots obtained from MOS capacitor of Cu/TiN/Ti/SiO_2/n-type Si structure: (a) TiN RTA treated at 500 °C in H_2/Ar ambient and (b) TiN RTA treated at 700 °C in NH_3 ambient.

occurred at 500 °C, while the TiN RTA treated at 700 °C in NH_3 ambient failed at 620 °C. NH_3 RTA improved the barrier property of TiN most, followed by N_2 RTA. Moreover, the barrier property improved with increasing RTA temperature, as shown in Fig. 12. It seems that nitrogen densifies TiN films by migrating to porous regions and that hydrogen reduces TiO_2 which is involved in TiN films, and so deteriorates the barrier property. Since NH_3 readily decomposes into N_2 and H_2 at temperatures over 500 °C, the TiN RTA treated in NH_3 ambient shows the highest failure temperature. We observed the formation of Ti silicide—a barrier against Cu and Si diffusion—at an RTA temperature higher than 600 °C by x-ray diffraction (XRD); thus it appears that improvement of the barrier property of TiN with RTA temperature was caused by Ti silicide formation.

IV. CONCLUSION

TiN films with various densities were studied as Cu diffusion barriers using various characterization methods, which showed that the denser TiN film was stable up to the higher failure temperature. Among the characterization methods, $C-V$ measurement best detected failure of the TiN barrier. $C-V$ characteristics after failure of the TiN barrier varied according to whether p-type or n-type Si wafers were used as a substrate. A decrease in inversion capacitance was observed in the case of p-type Si wafers; on the other hand, the opposite result was obtained with n-type Si wafers. The failure temperature of the TiN RTA treated at 700 °C in NH_3 ambient increased by more than 140 °C according to $C-V$ measurement.

ACKNOWLEDGMENT

The authors acknowledge the support of LG Semicon Co., Ltd. in this study.

[1]S. P. Muraka, *Tungsten and Other Advanced Metals for VLSI Applications 1990*, edited by G. C. Smith and R. Blumenthal (Materials Research Society, Pittsburgh, PA, 1991), p. 179.
[2]P. Singer, Semicond. Int. **Sept.**, 34 (1993).
[3]N. Toyama, Solid-State Electron. **26**, 37 (1983).
[4]Y. Shacham-Diamond, A. Dedhia, D. Hoffstetterand, and W. G. Oldham, J. Electrochem. Soc. **140**, 2427 (1993).
[5]S. K. Estreicher, Phys. Rev. B **41**, 5447 (1990).
[6]M. O. Aboelfotoh and B. G. Svensson, Phys. Rev. B **44**, 742 (1991).
[7]I. Sumi, M. Blomberg, and J. Saarilahti, J. Vac. Sci. Technol. A **3**, 2233 (1985).
[8]N. Fumimura, N. Nishida, L. Ito, and Y. Nakayama, Mater. Sci. Eng., A **108**, 153 (1989).
[9]A. Armigliato and G. Valdre, J. Appl. Phys. **61**, 390 (1987).
[10]S. Q. Wang, I. J. Raaijmakers, B. J. Burrow, S. Suthar, S. Redkar, and K. B. Kim, J. Appl. Phys. **68**, 5176 (1990).
[11]J. O. Olowolafe, J. Li, J. W. Mayer, and E. G. Colgan, Appl. Phys. Lett. **58**, 469 (1991).
[12]C. A. Chang and C. K. Hu, Appl. Phys. Lett. **57**, 617 (1990).
[13]J. O. Olowolafe, C. J. Mogab, R. B. Ciregory, and M. Kottke, J. Appl. Phys. **72**, 4099 (1992).
[14]I. Sumi, M. Maenpaa, M. A. Nicolet, and M. Luomajavi, J. Electrochem. Soc. **130**, 1215 (1983).
[15]S.-K. Rha, W.-J. Lee, S.-Y. Lee, D.-W. Kim, and C.-O. Park, J. Mater. Res. **12**, 3367 (1997).
[16]W.-J. Lee, J.-S. Min, S.-K. Rha, S.-S. Chun, C.-O. Park, and D.-W. Kim, J. Mater. Sci.: Mater. Electron. **7**, 111 (1996).
[17]K.-C. Park, K.-B. Kim, I. J. M. M. Raaijmakers, and K. Ngan, J. Appl. Phys. **80**, 5674 (1996).
[18]M. B. Chamberlain, Thin Solid Films **91**, 155 (1982).
[19]R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed. (Wiley, New York, 1986), p. 378.
[20]W. Shockley and W. T. Read, Phys. Rev. **87**, 835 (1952).
[21]R. N. Hall, Phys. Rev. **87**, 387 (1952).

Thermal stability of selective chemical vapor deposited tungsten contact and effects of *in situ* N_2 plasma treatment

M. T. Wang,[a] P. C. Wang, M. C. Chuang,[b] L. J. Chen,[c] and M. C. Chen
Department of Electronics Engineering, National Chiao-Tung University, Hsinchu, Taiwan

(Received 8 August 1997; accepted 3 April 1998)

This work investigates the thermal stability of Al/W/p^+-n junction diodes, in which the W contact was filled using selective chemical vapor deposition to a thickness of about 450 nm and served as diffusion barrier between the Al and the Si substrate. The effects of *in situ* N_2 plasma treatment on the barrier effectiveness were also investigated. The Al/W(450 nm)/p^+-n junction diodes can sustain a 30 min furnace annealing up to 575 °C. With an *in situ* N_2 plasma treatment on the W surface caused a thin layer of WN_x to form on the W surface, and the nitrified layer of WN_x/W acting as barrier between the Al and the Si substrate effectively suppressed WAl_{12} formation at elevated temperatures, resulting in a significant barrier improvement. N_2 plasma treatment at 100 W for 300 s enabled the Al/WN_x/W(450 nm)/p^+-n junction diodes to sustain thermal annealing at temperatures up to 625 °C without degradation of electrical characteristics. © *1998 American Vacuum Society.* [S0734-211X(98)00304-7]

I. INTRODUCTION

Because Si is highly soluble in Al alloy and exhibits high diffusivity along Al grain boundaries, Al/Si contacts tend to fail due to Si dissolving in Al and Al spiking into Si substrates.[1,2] The Al–Si alloy also becomes supersaturated during cooling cycles and excess Si precipitates at the Al/Si interface, resulting in degradation of contact characteristics. Therefore, the introduction of a barrier layer between the Al metallization and the Si substrate is of fundamental importance in preventing junction spiking and avoiding contact degradation.[2]

Tungsten has been considered as a good contact barrier for protecting shallow junctions from aluminum spiking and achieving low contact resistance.[3–8] Tungsten nitride is also an effective diffusion barrier between Si and contact metals because it is chemically more stable than pure tungsten, and has a greater capacity for suppressing interdiffusion between Al and Si.[9–15] Generally, tungsten nitride is deposited by reactively sputtering a pure W target in Ar–N_2 mixed ambient.[9,10] It has been reported that reactively sputtered tungsten nitride far surpassed pure W metal as a diffusion barrier between Al and Si substrates.[9] However, it is difficult to deposit tungsten nitride with excellent barrier properties into contact holes of submicron dimension using sputtering techniques because of potential step coverage problems. In this respect, selective chemical vapor deposition of tungsten (selective CVD-W) is a most attractive technique for filling deep submicron contact holes in ultralarge scale integrated (ULSI) interconnect applications.[8]

It has been reported that a selective 73-nm-thick CVD-W film acted as an effective diffusion barrier between Al and Si substrate at 450 °C for 30 min; however, it failed at 520 °C.[6] It has also been reported that no discernible reaction took place at an Al/W interface at temperatures up to 500 °C with a W film deposited using CVD; however, reaction occurred at an Al/W interface with sputtered W even at 450 °C and a WAl_{12} compound formed.[4] Moreover, it was found that N_2 plasma treatment of a CVD-W surface before Al deposition prevents Al–W alloy formation and suppresses the increase of sheet resistance for the Al/W structure.[15] Although these studies have provided much valuable information for the application of thin CVD-W films (less than 100 nm) as diffusion barriers, nevertheless, little study has been made of the thermal stability of thicker CVD-W film.[8]

Generally, W plugs used in industry are 450 nm or more in thickness.[16] In this study, 100 and 450-nm-thick CVD-W films were selectively deposited to fill contact holes and acted as diffusion barriers between Al and Si substrates. The film resistivities of the selective CVD-W were found to be about 10 $\mu\Omega$ cm. This technique offers the advantage of producing a fully self-aligning contacts and barrier-formation processes. In addition, the resultant structures also possess more planar surfaces for subsequent metallization processing. To study the effects of N_2 plasma treatment, an *in situ* N_2 plasma treatment was performed on tungsten surfaces prior to Al metallization. After the treatment, thin WN_x layers formed on the surfaces of the W layers, which efficiently suppressed diffusion between W and Al during post-metallization thermal annealing. The results of this study might be useful in producing selective CVD-W-filled plugs for ULSI interconnection applications.

II. EXPERIMENT

To evaluate the barrier effectiveness of a selective CVD-W layer, various samples of Al/barrier/p^+-n junction diodes were prepared. The starting material was n type, ⟨100⟩-oriented Si wafers with 4–7 Ω cm nominal resistivity. After RCA standard cleaning, the wafers were thermally oxi-

[a]Electronic mail: mcchen@cc.nctu.edu.tw
[b]Present address: National Nano Device Laboratory, 1001 Ta Hsueh Rd., Hsinchu, Taiwan.
[c]Present address: Department of Submicron Technology Development, ERSO/ITRI, Hsinchu, Taiwan.

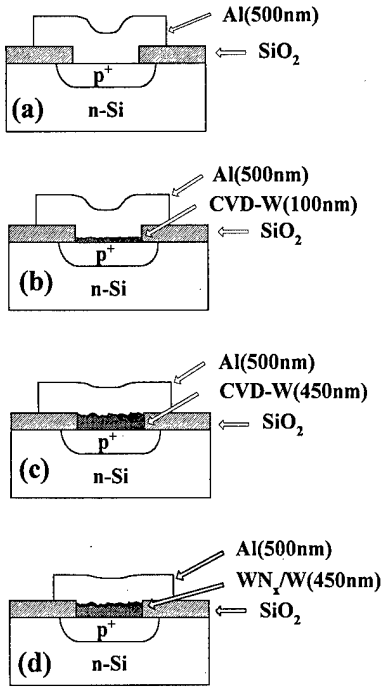


FIG. 1. Schematic cross sections of four differently metallized p^+-n junction diodes (a) Al/p^+-n, (b) Al/W(100 nm)/p^+-n, (c) Al/W(450 nm)/p^+-n, and (d) Al/WN_x/W(450 nm)/p^+-n.

dized to grow a 500 nm oxide layer. Diffusion areas with sizes of 500×500 and 1000×1000 μm^2 were defined on the oxide-covered wafers using conventional photolithographic techniques. The p^+-n junctions 300 nm deep were formed using BF_2^+ implantation at 30 keV to 3×10^{15} cm^{-2} followed by furnace annealing at 900 °C for 30 min in N_2 ambient.

After junction formation, the wafers were split into four groups for the preparation using various metallization structures: Al/p^+-n, Al/W(100 nm)/p^+-n, Al/W(450 nm)/p^+-n, and Al/WN_x/W(450 nm)/p^+-n. For the Al/W(100 nm)/p^+-n and Al/W(450 nm)/p^+-n diodes, the contact holes were selectively deposited with CVD-W layers of 100 and 450 nm, respectively, followed by Al metallization. For the Al/WN_x/W(450 nm)/p^+-n diodes, an *in situ* N_2 plasma treatment was performed on the selectively deposited CVD-W layers of 450 nm thickness before the Al metallization; thus, thin WN_x layers formed on the W surfaces. The schematic cross sections of these differently metallized p^+-n junction diodes are shown in Fig. 1. The contact holes in the Al/W(450 nm)/p^+-n and Al/WN_x/W(450 nm)/p^+-n diodes, were almost fully filled by selective CVD-W; this provided nearly planar surfaces for subsequent metallization processing.

Prior to selective CVD-W deposition, the wafers were dipped in dilute HF (50:1) for 30 s, followed by a rinse in deionized water for 3 min and a spin dry. The wafers were then loaded into a load-locked cold wall W-CVD system (ULVAC ERA-1000S) within 5 min and transferred to the deposition chamber without exposure to the atmosphere. The ERA-1000S is a fully automatic single wafer CVD system equipped with a cluster of multichambers, including a load/unload, buffer and two deposition chambers. This system employs a robot unit in the buffer chamber for wafer transfer in vacuum. The aluminum-alloy reactor was watercooled and was kept at a high vacuum base pressure by a turbopump. The base pressure of the CVD chamber was 1×10^{-6} Torr. In this study, W films were chemically vapor deposited using the SiH_4 reduction of WF_6 process under the following conditions: substrate temperature 300 °C, total gas pressure 100 mTorr, WF_6 flow rate 40 sccm, SiH_4 flow rate 10 sccm, and H_2 carrier gas flow rate 1000 sccm. After selective CVD-W deposition, one group of wafers was further treated with *in situ* N_2 plasma without breaking the vacuum under the following conditions: N_2 flow rate 80 sccm, total gas pressure 30 mTorr, treatment time 300 s, and plasma generation power ranging from 50 to 200 W. Finally, Al metallization was applied to all samples.

FIG. 2. SEM micrographs showing an overfilled 0.5 μm contact hole using selective CVD-W techniques.

To investigate thermal stability of the differently metallized junction diodes, samples were thermally annealed in an N_2 flow furnace for 30 min at temperatures ranging from 350 to 650 °C. Electrical characteristics were measured using an HP-4145B semiconductor parameter analyzer. Unpatterned samples with various W/Si, Al/W/Si, and Al/WN_x/W/Si structures were also prepared for material analysis. Sheet resistance of the multilayer structures was measured using a four-point probe. X-ray diffraction (XRD) analysis was used for phase identification. Auger electron spectroscopy (AES) was used to observe WN_x formation following the N_2 plasma treatment. Scanning electron microscopy (SEM) was used to observe surface morphology as well as cross-sectional microstructure, and secondary ion mass spectroscopy (SIMS) was used for elemental depth profile measurement.

III. RESULTS AND DISCUSSION

A. Selective CVD-W and N_2 plasma treatment

Figure 2 shows overfilled 0.5 μm contact holes resulting from using a selective CVD-W technique. Excellent selectivity, good uniformity, and low resistivity W (about 10 $\mu\Omega$ cm) were obtained with a WF_6/SiH_4 flow rate of 40/10 sccm. Thus, this WF_6/SiH_4 flow rate was used to de-

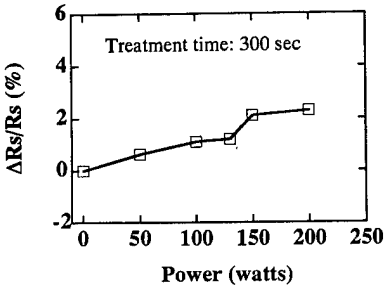


FIG. 3. Percentage change of sheet resistance vs N_2 plasma power in W(450 nm)/Si samples.

posit tungsten during barrier study in this work. Details of the CVD-W selectivity study were reported elsewhere[17] and are not discussed further in this article.

Post-deposition N_2 plasma treatment resulted in a sheet resistance change in the W/Si samples. Figure 3 shows the percentage change of sheet resistance versus N_2 plasma power for a W(450 nm)/Si sample. The sheet resistance increased slightly with increasing plasma power, presumably due to formation of thin WN_x layers on W surfaces, and the thicknesses of WN_x layers increased with N_2 plasma power. The measured AES depth profiles, as shown in Fig. 4, provide evidence for these presumptions. With N_2 plasma treatment at 100 W, a thin WN_x layer about 2 nm thick was formed on the W surfaces [Fig. 4(b)], while WN_x layer thicknesses were about 5 nm in the samples treated with N_2 plasma at 200 W [Fig. 4(c)]. Since metal nitride generally has a higher resistivity than corresponding pure metal, the presence of a thin WN_x layer on the W surface obviously resulted in a slight increase in sheet resistance in the W/Si samples.

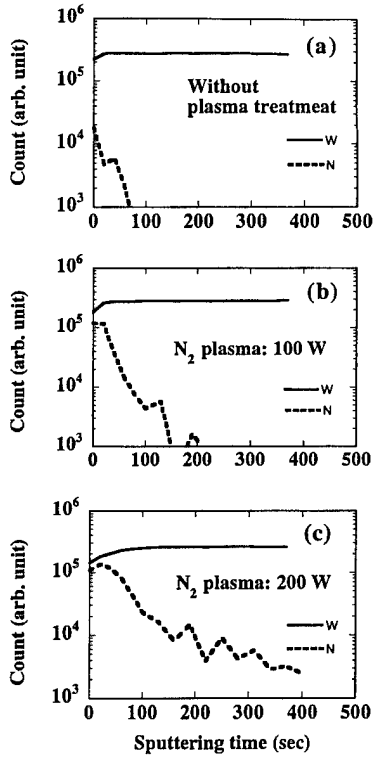


FIG. 4. AES depth profiles of W/Si samples (a) as-deposited, (b) after N_2 plasma treatment at 100 W, and (c) after N_2 plasma treatment at 200 W. The sputtering rate was about 0.1 nm/s.

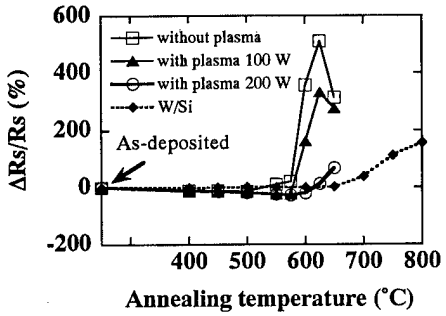


FIG. 5. Percentage change in sheet resistance plotted against annealing temperature for the W/Si sample as well as the Al/W/Si samples with and without N_2 plasma treatment on the W surface.

B. Sheet resistance measurement and XRD analysis

Sheet resistance measurement and XRD analysis were used to evaluate the barrier capability of the W films. Figure 5 shows the percentage change in sheet resistance plotted against annealing temperature for the W/Si, Al/W/Si, and Al/WN_x/W/Si samples. The Al/W/Si samples whose W barriers did not receive N_2 plasma treatment showed increases in sheet resistance after annealing at 550 °C. With N_2 plasma treatment on the W surface, sheet resistance for the Al/WN_x/W/Si samples remained constant up to at least 575 °C. Although the WN_x layer was very thin, it showed an excellent barrier ability to suppress the reaction between Al and W, resulting in improved thermal stability for the Al/WN_x/W/Si structure.

The increase in sheet resistance for the Al/W/Si structure reflects the consumption of conductive aluminum due to WAl_{12} formation, as confirmed by the x-ray diffraction pattern shown in Fig. 6. The WAl_{12} compound appeared for the Al/W/Si sample annealed at 550 °C [Fig. 6(b)]. After annealing at 600 °C, phase signal intensity of the WAl_{12} increased and a part of it was converted into WAl_5 phase,[18] W_5Si_3 and WSi_2 signals both appeared as well [Fig. 6(c)]. Moreover, the Al signal disappeared while the α-W signal remained, implying that the 500 nm Al layer had been completely converted into WAl_{12} and WAl_5 but the 450 nm W layer was only partially consumed. The presence of W_5Si_3 and WSi_2 signals indicates that reactions took place at the W/Si interface during the 600 °C annealing.

The results of XRD analysis of Al/WN_x/W/Si samples treated with N_2 plasma at 200 W are illustrated in Fig. 7. No compound phase of W–Al or W–Si was observed after samples were annealed at temperatures up to 600 °C [Fig. 7(b)], which confirms the integrity of the multilayer structure. Apparently, N_2 plasma treatment has the effect of increasing the thermal stability of the Al/barrier/Si structure. After annealing at 650 °C, $W(Si, Al)_2$ and WSi_2 signals ap-

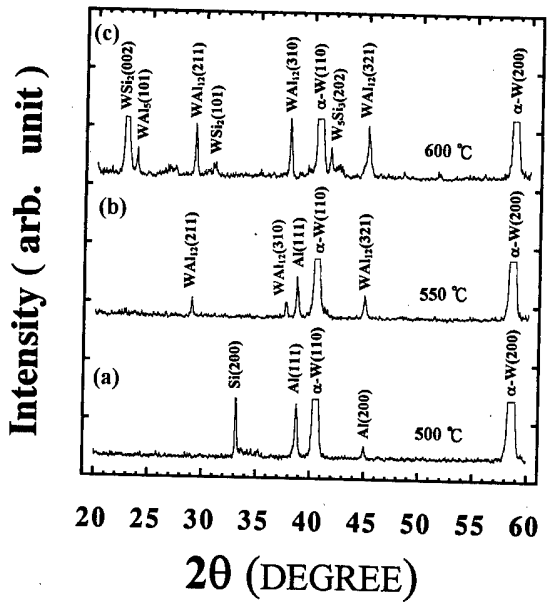


FIG. 6. XRD spectra for the Al/W(450 nm)/Si sample after annealing at (a) 500, (b) 550, and (c) 600 °C.

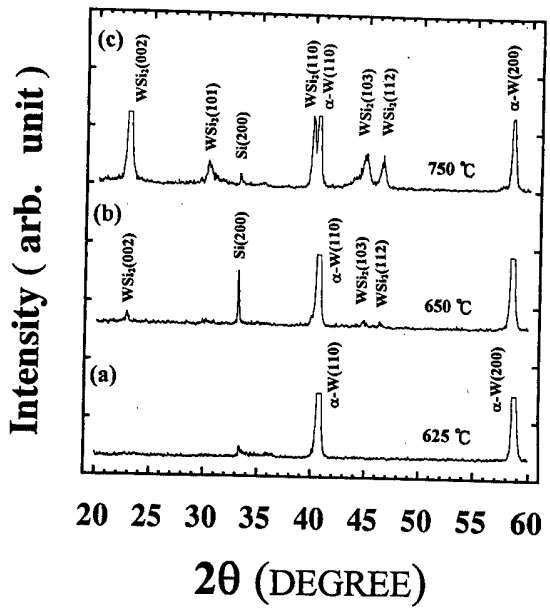


FIG. 8. XRD spectra for the W/Si sample after annealing at (a) 625, (b) 650, and (c) 750 °C.

peared and the Al signal remained, while no WAl_{12}-related signals were observed [Fig. 7(c)]. The W/Si samples without Al overlayers were also investigated for comparison. The increase in sheet resistance for the W/Si structure after annealing at temperatures above 700 °C (Fig. 5) presumably resulted from the consumption of conductive tungsten due to the formation of WSi_2, as confirmed by the XRD analysis results shown in Fig. 8. In fact, the XRD spectra show that weak WSi_2-phase signals began appearing at 650 °C, indicating reaction between W and the Si substrate. The WSi_2-signal intensity increased with increasing annealing temperature, but the α-W signal remained even after annealing at 750 °C. This indicates that the 450 nm W layer was only partially consumed after annealing at 750 °C. By comparing the XRD spectra of W/Si samples (Fig. 8) with those of Al/W/Si and Al/WN_x/W/Si samples (Figs. 6 and 7), we found that the formation of WAl_{12} or $W(Si, Al)_2$ promoted the formation of WSi_2 for the Al/W/Si structure. Similar results have been reported in which the presence of an Al overlayer in the Al/W/Si structure led to a reduction in thermal stability of the underlying W/Si interface.[5] Comparative thermal stability results for the various layered structures are given in Table I.

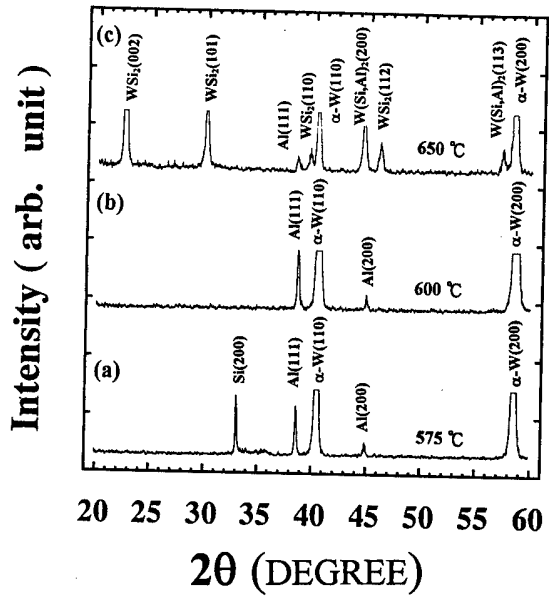


FIG. 7. XRD spectra for the Al/WN_x/W(450 nm)/Si sample after annealing at (a) 575, (b) 600, and (c) 650 °C. The W surface was treated with N_2 plasma at 200 W.

C. Electrical measurements

Barrier capability of the self-aligned selective CVD-W layers was investigated by evaluating the thermal stability of the Al/W/p^+-n junction diodes using electrical measurements. Figure 9 illustrates the distributions of reverse bias leakage current density measured at -5 V for the Al/p^+-n, Al/W(100 nm)/p^+-n, and Al/W(450 nm)/p^+-n junction diodes annealed at various temperatures; the diodes had a diffusion area of 500×500 or 1000×1000 μm^2, and at least 30 diodes were measured in each case. Electrical characteristics of the Al/p^+-n diodes deteriorated after annealing at 400 °C [Fig. 9(a)], while the Al/W(100 nm)/p^+-n diodes remained stable up to 500 °C [Fig. 9(b)]. Moreover, the Al/W(450 nm)/p^+-n diodes were able to retain their integrity up to 575 °C [Fig. 9(c)]. It is clear that the thermal stability of the Al/W/p^+-n diodes can be improved further by using a thicker W film. Reaction of Al and W at elevated temperatures leads to the formation of WAl_{12}, and the barrier capability of a thin W film will be determined by the W consumption. For the Al/W(450 nm)/p^+-n diodes, partial consumption of the thicker W layer due to the WAl_{12} forma-

TABLE I. Comparative results of thermal stability for various layered structures.

Annealing temperature	W/Si	Al/W/Si	Al/WN_x/W/Si (N_2 plasma 100 W)	Al/WN_x/W/Si (N_2 plasma 200 W)
500 °C	X[a]	X	X	X
550 °C	X	WAl_{12}	X	X
575 °C	X	WAl_{12}, WAl_5	X	X
600 °C	X	WAl_{12}, WAl_5, $W(Al,Si)_2$, WSi_2, W_5Si_3	WSi_2, $W(Al,Si)_2$	X
650 °C	WSi_2	WAl_5, $W(Al,Si)_2$, WSi_2	WSi_2, $W(Al,Si)_2$	WSi_2, $W(Al,Si)_2$

[a]"X" indicates no observation of compound phase.

tion did not deteriorate the integrity of W/Si interface; thus the diodes were stable even after annealing at 575 °C.

For comparison, junction diodes with W(450 nm)/p^+-n structures, but without Al overlayers were also fabricated for thermal stability study. Figure 10 shows that the W(450 nm)/p^+-n diodes retained their integrity up to 700 °C and showed only slight degradation after annealing at 750 °C. This implies that degradation of the Al/W(450 nm)/p^+-n diodes can be attributed to the presence of the Al overlayer.

The barrier capability of CVD-W layer can be effectively improved by *in situ* N_2 plasma treatment prior to Al metallization. Figure 11 shows the statistical distributions of reverse bias leakage current density for the variously annealed Al/WN_x/W(450 nm)/p^+-n junction diodes, in which the WN_x layer on the W surface was formed by N_2 plasma treatment. About half of the diodes whose W barriers were treated with N_2 plasma at 50 W survived thermal annealing at 600 °C, retaining a leakage current density of less than 100 nA/cm^2, as shown in Fig. 11(a); by contrast, none of the Al/W(450 nm)/p^+-n diodes survived the same thermal annealing at 600 °C [Fig. 9(c)]. Because the thickness of the WN_x layer increased with increasing N_2 plasma power (Fig. 4), further improvement of barrier capability can be obtained by increasing the N_2 plasma power. With N_2 plasma treatment at 100 W, the Al/WN_x/W(450 nm)/p^+-n diodes were able to retain their integrity up to 625 °C, as shown in Fig. 11(b). Increasing the N_2 plasma power to 200 W did not improve the thermal stability of the Al/WN_x/W(450 nm)/Al/WN_x/W(450 nm)/p^+-n diodes, as shown in Fig. 11(c). Thermal stability temperatures for Al/WN_x/W(450 nm)/p^+-n diodes as high as 650 °C were not achieved in this study, presumably because of closeness to the Al melting point of 660 °C, and the formation of WSi_2 at the W/Si interface (Fig. 7). Based on these experimental results, we conclude that selective CVD-W layers with post-deposition *in situ* N_2 plasma treatment form effective diffusion barrier between Al and Si substrate.

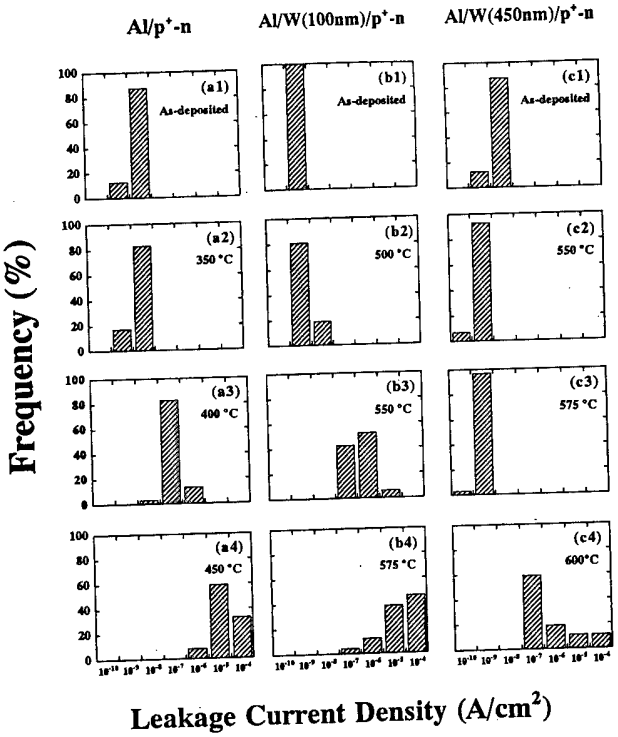


FIG. 9. Histograms showing the distributions of reverse bias leakage current density measured at -5 V for (a) Alp^+-n, (b) Al/W(100 nm)/p^+-n, and (c) Al/W(450 nm)/p^+-n junction diodes annealed at various temperatures.

D. SEM observation

Scanning electron microscopy (SEM) was used to investigate the surface and cross-sectional morphologies of the Al/barrier/p^+-n junction diodes. Figure 12 shows SEM micrographs of the Al/W(450 nm)/p^+-n junction diodes before and after thermal annealing. The as-deposited samples have rough Al surfaces because of the rough surfaces of the CVD-W underlayers [Fig. 12(a)]. Both the surface morphology and the W/Si interface remained unchanged after annealing at 500 °C [Fig. 12(b)]. After annealing at 600 °C, precipitates were found at the W/Si interface and protrusions

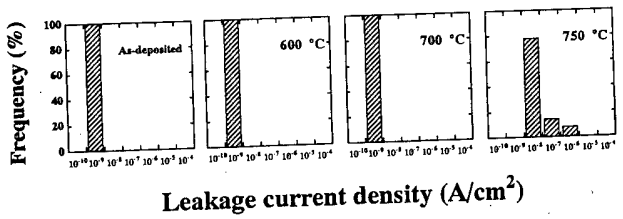


FIG. 10. Histograms showing the distributions of reverse bias leakage current density measured at −5 V for the W(450 nm)/p^+-n junction diodes annealed at various temperatures.

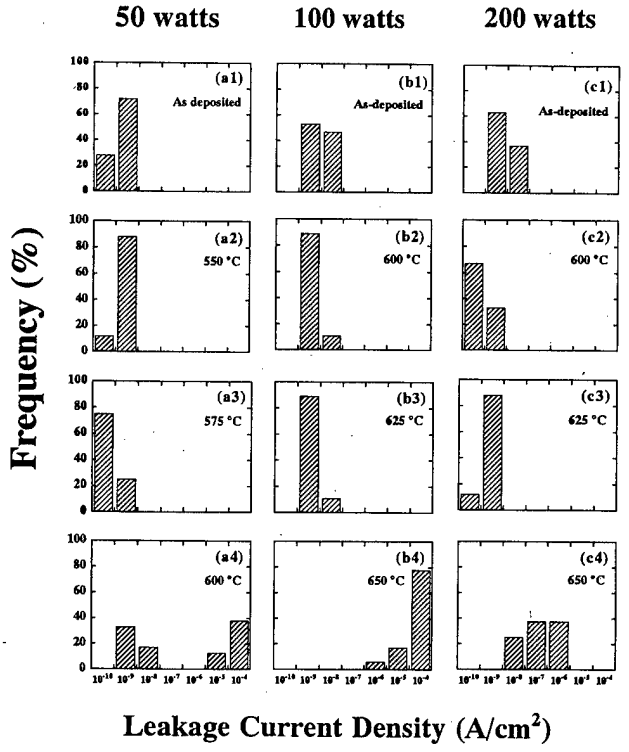


FIG. 11. Histograms showing the distributions of reverse bias leakage current density measured at −5 V for the Al/WN_x/W(450 nm)/p^+-n junction diodes annealed at various temperatures; the W layers of the junction diodes were treated with N_2 plasma at (a) 50, (b) 100, and (c) 200 W.

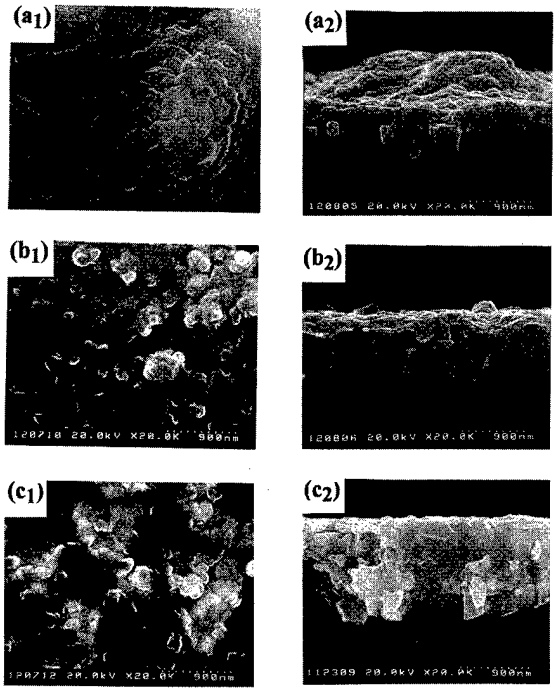


FIG. 13. Top-view (left) and cross-sectional view (right) SEM micrographs for (a) 600 °C annealed Al/WN_x/W(450 nm)/p^+-n diodes whose W barrier was treated with N_2 plasma at 50 W, (b) 600 °C annealed Al/WN_x/W(450 nm)/p^+-n diodes whose W barrier was treated with N_2 plasma at 100 W, and (c) 650 °C annealed Al/WN_x/W(450 nm)/p^+-n diodes whose W barrier was treated with N_2 plasma at 100 W.

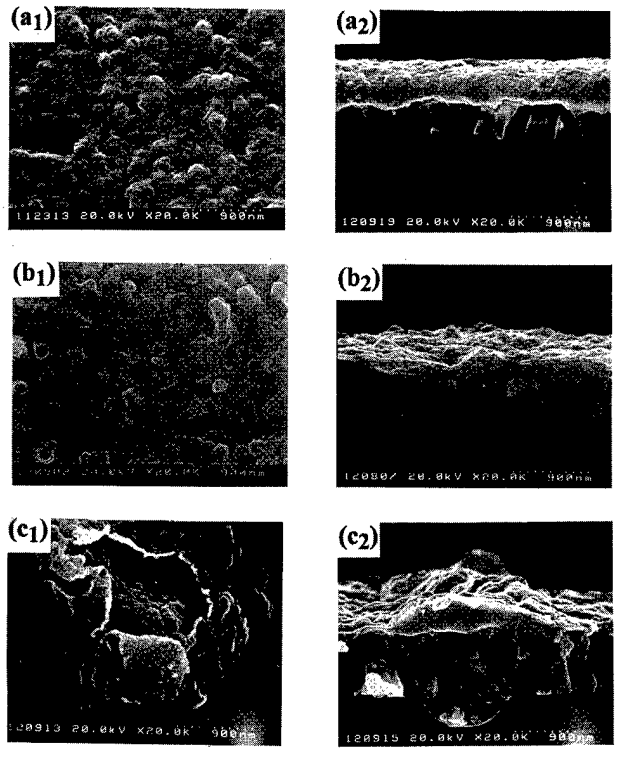


FIG. 12. Top-view (left) and cross-sectional view (right) SEM micrographs for the Al/W(450 nm)/p^+-n junction diodes (a) as-deposited, (b) annealed at 500 °C, and (c) annealed at 600 °C.

were observed on the surface, presumably due to WAl_{12} formation [Fig. 12(c)]. Assuming total conversion of the 500-nm-thick Al film into the WAl_{12}, only about 40 nm of W-layer thickness was consumed and the remaining 410 nm should be able to function as a barrier layer. However, nonuniform formation of WAl_{12} might lead to a highly localized thinning of the W layer; thus, either Al atoms or WAl_{12} compound can be very close to the W/Si interface. As a result, close proximity of Al atoms to the Si substrate can easily induce reaction between Al and Si during thermal annealing at 600 °C. Moreover, it is possible that the formation of WSi_2 and WAl_{12} introduces microcracks that cannot be observed by SEM but might provide fast diffusion paths for Al atoms to react with the Si substrate.

Figure 13 shows top-view and cross-sectional SEM micrographs of the thermally annealed Al/WN_x/W(450 nm)/p^+-n junction diodes. Instead of protrusions on the surface and large precipitates at the W/Si interfaces for the 600 °C annealed Al/W(450 nm)/p^+-n samples [Fig. 12(c)], surface cracks and small precipitates at the W/Si interface were found on the 50 W plasma-treated Al/WN_x/W(450 nm)/p^+-n diodes after annealing at 600 °C, as shown in Fig. 13(a). Figure 13(b) shows the 600 °C annealed Al/WN_x/W(450 nm)/p^+-n samples whose barriers were treated with N_2 plasma at 100 W; compared with the as-deposited Al/W(450 nm)/p^+-n shown in Fig. 12(a), no obvious difference is observable with respect to surface morphology and W/Si interface smoothness. The thin WN_x layer that formed

TABLE II. Comparative thermal stability results for p^+-n junction diodes with different W-barrier layers as determined using various characterization techniques.

Barrier/metallization structures of p^+-n junction diodes	Leakage current	SEM	SIMS
Al/W(100 nm)/p^+-n	500 °C	500 °C	
Al/W(450 nm)/p^+-n	575 °C	575 °C	<600 °C
Al/WN_x/W(450 nm)/p^+-n (N_2 plasma at 100 W)	625 °C	625 °C	>600 °C
Al/WN_x/W(450 nm)/p^+-n (N_2 plasma at 200 W)	625 °C	625 °C	>600 °C
W(450 nm)/p^+-n	700 °C	WSi_2 formed at 700 °C	

on the W surface due to the N_2 plasma treatment, was able to suppress the WAl_{12} compound formation and kept the W/Si interface unchanged up to 600 °C. After annealing at 650 °C, all of the diodes failed, presumably due to the low melting point of Al (660 °C). Compound [WAl_{12}, and WAl_5, and $W(Si, Al)_2$] formation along with tungsten silicidation deteriorated the Al/WN_x/W(450 nm)/p^+-n diodes completely, as shown in Fig. 13(c). Comparative thermal stability results for the p^+-n junction diodes with different W-barrier layers determined using various characterization techniques are given in Table II.

E. SIMS analysis

SIMS depth profiles were used to examine elemental distributions in the multilayer structures. Figure 14 shows the SIMS depth profiles for the Al/W(450 nm)/Si and Al/WN_x/W(450 nm)/Si samples before and after thermal annealing at 600 °C. Without N_2 plasma treatment, the interdiffusion of Al, W, and Si led to severe deterioration of the Al/W (450 nm)/Si structure after 600 °C annealing, as shown in Fig. 14(b). The Al/WN_x/W(450 nm)/Si samples whose W barriers were treated with N_2 plasma at 200 W, had W and Si profiles that remained almost unchanged, while Al diffused only slightly into W layers, as shown in Fig. 14(d). This indicates that the thermal stability of the Al/W barrier/p^+-n diodes can be significantly improved using N_2 plasma treatment on the W surfaces.

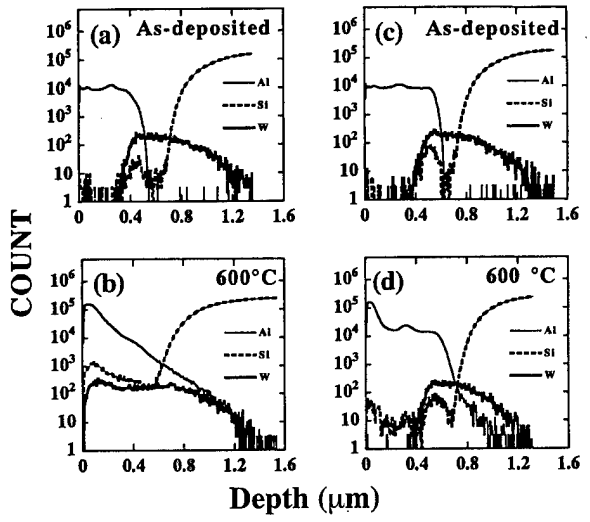


FIG. 14. SIMS depth profiles for Al/W(450 nm)/Si (a) as-deposited and (b) 600 °C annealed, and for Al/WN_x/W(450 nm)/Si (c) as-deposited and (d) 600 °C annealed.

IV. SUMMARY

The effect of *in situ* N_2 plasma treatment on the thermal stability of selective CVD-W contacted p^+-n junction diodes with Al metallization was systematically investigated. Without N_2 plasma treatment, the Al/W(100 nm)/p^+-n and Al/W(450 nm)/p^+-n junction diodes were able to sustain a 30 min furnace annealing up to 500 and 575 °C, respectively. With *in situ* N_2 plasma treatment on W surfaces at a power of 100 W, the Al/WN_x/W(450 nm)/p^+-n junction diodes were able to sustain thermal annealing up to 625 °C without degradation of their electrical characteristics. Reaction between Al and CVD-W at elevated temperatures resulted in formation of WAl_{12}, WAl_5, and $W(Si, Al)_2$, leading to acceleration of the reaction between W and Si substrates. The N_2 plasma treatment resulted in formation of a thin WN_x layer on the W surface, and the Al/WN_x/W/Si structure effectively suppressed W–Al compound formation, leading to improvement of thermal stability.

ACKNOWLEDGMENT

This work was supported by the National Science Council (ROC) under Contract No. NSC86-2215-E-009-040.

[1]S. Wolf, *Silicon Processing for the VLSI Era* (Lattice, Sunset Beach, CA, 1990), Vol. 2, p. 111.
[2]S. R. Wilson, C. J. Tracy, and J. L. Freeman, Jr., *Handbook of Multilevel Metallization for Integrated Circuits: Materials, Technology, and Applications* (Noyes, Park Ridge, NJ, 1993), pp. 35 and 38.
[3]L. Gutai, M. Delfino, and J. M. De Blasi, Mater. Res. Soc. Symp. Proc. **178**, 265 (1987).
[4]T. Hara, S. Enomoto, N. Ohtsuka, and S. Shima, Jpn. J. Appl. Phys., Part 1 **24**, 828 (1985).
[5]Y. Pauleau, F. C. Dassapa, P. Lami, J. C. Oberlin, and F. Romagna, J. Vac. Sci. Technol. B **6**, 817 (1988).
[6]B. W. Shen, G. C. Smith, J. M. Anthony, and R. J. Matyi, J. Vac. Sci. Technol. B **4**, 1369 (1986).
[7]Y. Shioya, M. Maeda, and K. Yanagida, J. Vac. Sci. Technol. B **4**, 1175 (1986).
[8]W. K. Yeh, K. Y. Chan, T. C. Chang, M. S. Lin, and M. C. Chen, J. Electrochem. Soc. **143**, 2053 (1996).

[9]H. P. Kattelus, E. Kolawa, K. Affolter, and M.-A. Nicolet, J. Vac. Sci. Technol. A **3**, 2246 (1985).
[10]F. C. So, E. Kolawa, X. A. Zhao, E. T-S. Pan, and M.-A. Nicolet, J. Appl. Phys. **64**, 2787 (1988).
[11]E. K. Broadbent, A. E. Morgan, J. M. Flanner, B. Coulman, D. K. Sadana, B. J. Burrow, and R. C. Ellwanger, J. Appl. Phys. **64**, 6721 (1988).
[12]M. Takeyama, K. Sasaki, and A. Noya, Jpn. J. Appl. Phys., Part 1 **31**, 3424 (1992).
[13]Y. T. Kim, C. W. Lee, and S. K. Min, Jpn. J. Appl. Phys., Part 1 **32**, 6126 (1993).
[14]A. Noya, M. Takeyama, K. Sasaki, E. Aoyagi, and K. Hiraga, Jpn. J. Appl. Phys., Part 1 **33**, 1528 (1994).
[15]M. Sekiguchi, T. Fujii, and M. Yamanaka, Jpn. J. Appl. Phys., Part 1 **35**, 1111 (1996).
[16]F. C. So, E. Kolawa, X. A. Zhao, and M.-A. Nicolet, Thin Solid Films **153**, 507 (1987).
[17]M. T. Wang, M. S. Tsai, C. Liu, W. T. Tseng, T. C. Chang, L. J. Chen, and M. C. Chen, Mater. Chem. Phys. **51**, 75 (1997).
[18]H. H. Hoang, Proceedings of the IEEE International Reliability Phys. Symposium (IRPS), 1988, p. 273.

Effects of sintering on Au/Ti/GaAs Schottky barrier submicron metal-semiconductor field-effect transistors characteristics

M. M. Ahmed[a)]
Faculty of Electronic Engineering, GIK Institute of Engineering Sciences and Technology, Topi, Swabi, Pakistan

(Received 19 June 1997; accepted 17 April 1998)

This investigation reports on the effects of sintering (25 °C up to 450 °C) on Au/Ti/GaAs Schottky barrier submicron metal-semiconductor field-effect transistors with an interfacial layer. In short channel devices the presence of an interfacial layer at the metal-semiconductor junction is inevitable. This reduces the transconductance and gives rise to the nonideal $I-V$ characteristics. The adverse effects of the interfacial layer can be reduced significantly by sintering the finished devices at 350 °C in H_2/N_2 ambient for 30 s. This enhances the Schottky response of the device and an improvement up to 40% in the device extrinsic transconductance has been achieved. © *1998 American Vacuum Society.* [S0734-211X(98)02304-X]

I. INTRODUCTION

In GaAs field-effect transistors (FETs) the Schottky junction plays a key role in the performance of the device. The height of the Schottky barrier is very sensitive to the surface condition of the semiconductor. An ideal Schottky barrier on GaAs or any other compound semiconductor is an epitaxially deposited film.[1] This type of Schottky contact generally eliminates the presence of interface states at the junction. In a practical FET fabrication process however, especially when a gate recess step is involved, interfacial layer growth is inevitable.

The presence of an interfacial layer at a Schottky barrier contributes to nonideality in the $I-V$ characteristics, because the interface states give rise to parallel processes of current flow via capture and emission of electrons passing through the barrier.[2] The trapping of charge in the interface states reduces the barrier height, and is a common cause of the soft reverse characteristics of a Schottky contact.[3,4] The interfacial layer has been assumed to be of atomic dimensions, such that it is transparent to electrons but can support a potential across it.[5] This gives rise to a voltage-dependent resistance at the interface. A finite portion of the gate potential is masked at the interface due to this resistance and the net potential that will alter the depletion underneath the gate is smaller than the applied voltage.[5–7] This reduces the gain and increases the threshold voltage of the devices.[3,6]

The importance of charge exchange with interface states on the microwave behavior of the device can be understood since it is a process which operates on a scale of time which is long compared with a microwave signal period. For instance, the characteristic time for emptying of an interface state may be around 1 ms,[2] whereas at 10 GHz the signal period is 100 ps. In contrast, the mechanism of electron transport over the barrier is extremely fast, involving time scales of the order of the dielectric relaxation time in an n-type GaAs semiconductor (typically <1 ps).[2]

This article describes an annealing process by which the adverse effects of interface states can be offset significantly and, hence, improved device performance can be achieved.

II. DEVICE DESCRIPTION

The complete device fabrication sequence is given elsewhere.[8] In brief, GaAs epilayers were grown by molecular beam epitaxy on Cr doped semi-insulating substrates. The layer structure of the wafer, as shown in Fig. 1, from bottom to top, consisted of a 50 nm thick undoped GaAs layer, a 50×20 nm thick GaAs and AlGaAs superlattice, an AlGaAs buffer layer, 1.5 μm GaAs smoothing layer, and finally a 0.2 μm Si doped GaAs channel layer (5×10^{17} cm^{-3}). A mesa was defined by an optical lithographic process. Ohmic contact delineation was accomplished by electron beam lithography. For Ohmic metallization AuGeNi 200 nm was applied in a proportion that represents an alloy Au:Ge:Ni 88:12:5 by weight.[9] A 100 nm thick layer of Au was then evaporated on the Ohmic metal to improve the contact resistance and surface morphology.[10] The Ohmic metallization layers (AuGeNi/Au) were evaporated at a base pressure of 1×10^{-6} Torr in a single sequence without breaking the vacuum. The contacts were alloyed by rapid electron beam annealing at the beam power density of 3 W cm^{-2} for 3 s.[11] Gate lithography was carried out with a high resolution electron beam lithography system operated at 60 kV, with probe size less than 300 Å. The nominal gate length, L_g of the FETs was 0.23 μm.

The recess was etched in a H_2SO_4:H_2O:H_2O_2 5 : 200 : 1 solution. The sample was then put into the oxide strip solution, i.e., diluted HCl, to dissolve the oxide layer before the Schottky metal evaporation. Immediately afterwards it was transferred into the evaporator for gate metal evaporation. The gate metallization, Ti/Au 30/200 nm, was then evaporated at a base pressure of 1×10^{-6} Torr, and the extra metal was lifted off in acetone.

III. RESULTS AND DISCUSSION

Uncertainty in the removal of native oxide and regrowth of an oxide layer before deposition of the Schottky metal

a)Electronic mail: mansoor@giki.sdnpk.undp.org

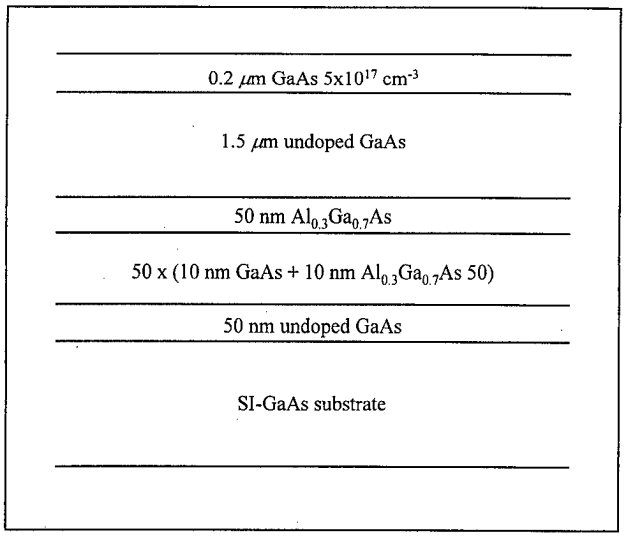


FIG. 1. Layer structure of the wafer used for GaAs MESFETs fabrication.

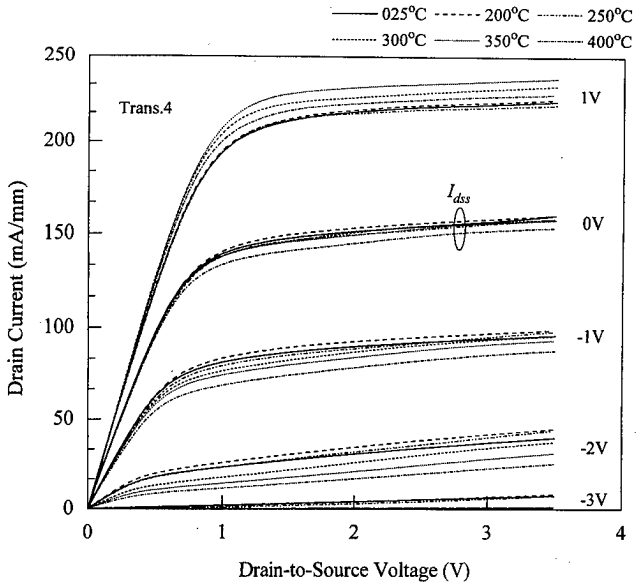


FIG. 2. Output characteristics of a 0.23 μm gate length Au/Ti/GaAs Schottky barrier MESFET after 30 s heat treatment at various temperatures.

could generate nonuniformity at the metal-semiconductor interface. To study this nonuniformity, four devices were selected from each of four different chips. These devices have the same L_g value, but with different surface states as assessed by their gate leakage current.[1,3,12] The difference in their surface state densities could be associated with the nonuniform cleaning of submicron gate trenches before the Schottky metal evaporation.[13–16] Chips were sintered in stages, up to 450 °C in H_2/N_2 ambient for 30 s. The devices were first electrically characterized on a dc probe station and then sintered at 200 °C, and measured again. The temperature was increased in steps of 50 °C, and measurements were performed after each heat treatment.

Sintering the Au/Ti/GaAs Schottky barrier at some elevated temperature can change the interface quality and, hence, the device performance. Wu and Yang[17] have studied the effective metal screening and Schottky barrier formation in metal GaAs structure by preparing Ti/Pt/GaAs and Pt/Ti/GaAs Schottky diodes in ultrahigh vacuum. After the first layer of evaporation Ti or Pt, they heated the samples *in situ* to 300 °C for 1 min; this increased the barrier height for Ti/GaAs Schottky barrier by 60 mV, whereas no change has been reported for Pt/GaAs Schottky barriers, which suggests that in the case of Ti/GaAs Schottky junctions there could be a finite density of interface states, whereas for Pt/GaAs the density of these states may be so small, or missing all together, that their impact on the electrical characterization of the device is virtually undetectable.

The change in metal-semiconductor FET (MESFET) Schottky barrier interface could be realized by comparing the $I-V$ characteristics measured before and after the sintering. Figure 2 shows the output characteristics of a MESFET measured at room temperature, after sintering the chip at various temperatures. Examination of the figure shows that at 0 V gate voltage, the values of drain current (I_{dss}) are almost the same up to 350 °C, and start decreasing slightly at 400 °C. A 1.0 V positive bias at the gate allows more current to flow through the active channel, especially between 300 and 400 °C sintering temperatures, compared with the unsintered characteristic. This indicates that for sintered devices a greater magnitude of applied potential is used to alter the depletion underneath the Schottky barrier gate and the masking of gate potential by the oxide layer is reduced.[3,6,7] This implies that the interface quality has improved after sintering.

A convenient way of assessing the change in device behavior is by plotting the extrinsic transconductance (g'_m) of the device as shown in Fig. 3. The solid curve, which represents the device performance prior to any heat treatment, is the reference curve. Sintering at 200 °C causes almost no change in the device performance but it starts improving with higher temperature heat treatment. The figure shows

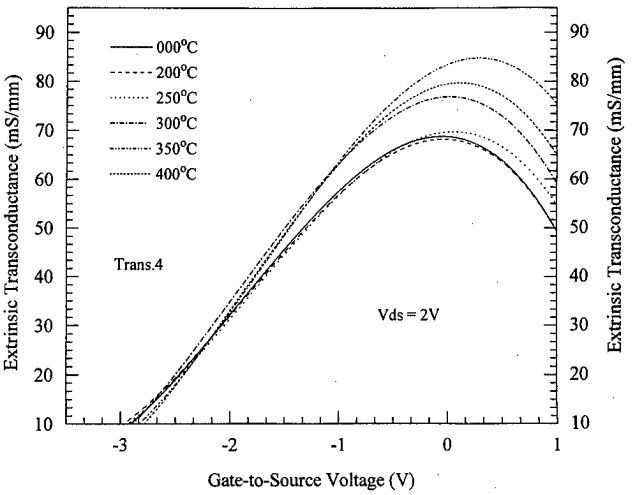


FIG. 3. Extrinsic transconductance vs sintering temperature of a 0.23 μm gate length Au/Ti/GaAs Schottky barrier MESFET.

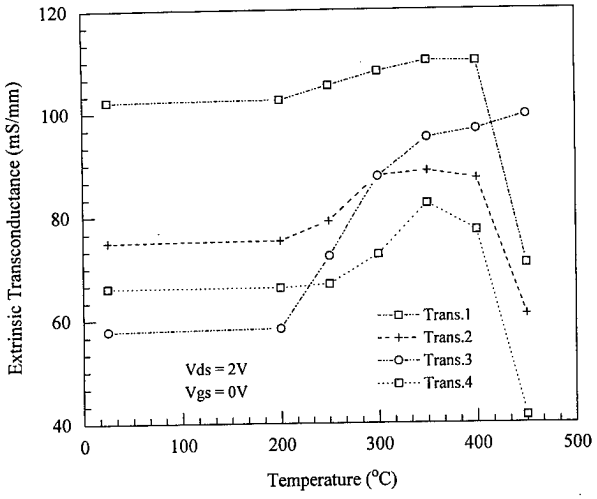


FIG. 4. Improvement in the extrinsic transconductance of four Au/Ti/GaAs Schottky barrier MESFETs by sintering at various temperatures.

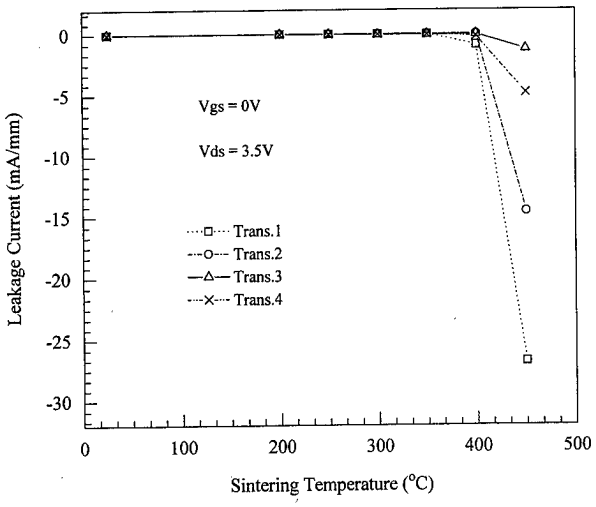


FIG. 5. Gate leakage current vs sintering temperature of four Au/Ti/GaAs Schottky barrier submicron GaAs MESFETs.

that the g'_m of the device is improved by increasing the sintering temperature and it has highest g'_m relative to untreated characteristic after 350 °C sintering. Sintering at 400 °C reduces the device g'_m.

The variations in g'_m as a function of sintering temperatures of four different devices are plotted in Fig. 4. The figure shows a consistent improvement in all the four devices under consideration. These four devices have different surface-state densities, which decrease with increasing values of g'_m (Fig. 4). Examination of the figure shows that the improvement in the device g'_m depends on the quality of the device. A low value of g'_m of an unsintered device could be associated with a high density of interface states.[3,6,13]

The applied gate voltage of a Schottky barrier with an interfacial layer can be expressed as $V_g = V_i + V_D$,[5,7] where V_i is the voltage drop across the interfacial layer and V_D represents the voltage drop across the depletion layer. The net potential that will change the depletion layer is $|V_g| - V_i$. This indicates that the transconductance of the device will compress. A high compression in the g'_m value may then be associated with a high density of interface states. In Fig. 4 transistor 1 which has high gain and the lowest density of surface states shows only 7% improvement in its g'_m at 350 °C, whereas transistor 3 which has the highest density of surface states shows 40 improvement at the same sintering temperature.

The decrease in g'_m at high sintering temperatures (i.e., 450 °C) may be due to the high value of the gate leakage current, I_g. Figure 5 shows the increase in I_g as a function of sintering temperature. The increased value of I_g may indicate that the barrier height has been reduced due to interdiffusion. Interdiffusion is a function of the nature of interfacial oxide layer. As we can realize the barrier lowering by the change in I_g, therefore the degradation in the device performance will be severe if the magnitude of I_g is high after the sintering. By examining Figs. 4 and 5 one can see that after sintering the devices which have lower I_g values at 450 °C show high improvement in their g'_m. These devices, presumably, have thicker interfacial layers, which indicates that the presence of oxides at the interface reduces the interdiffusion of Ti/GaAs contact. Although the presence of interfacial layer causes a reduction in the device g'_m, this experiment shows that we can offset the detrimental effects of this layer considerably by sintering the chip at a suitable temperature.

We can, therefore, conclude that sintering the GaAs MESFETs having Ti/GaAs Schottky barriers at 350 °C improves the device performance significantly. The level of improvement depends on the density of interface states. For high interface-state densities devices the observed improvement is also high.

IV. SUMMARY

In submicron GaAs MESFETs the presence of an interfacial layer at the Schottky barrier junction causes the performance of the device to deteriorate. The degradation in the device performance will be severe if the density of interface states is high. The growth of an interfacial layer in FET fabrication process is inevitable, especially when the value of L_g is very small (submicron device). The Schottky $I-V$ measurements showed that there are serious adverse effects on the Schottky junction due to the presence of an interfacial layer, which can be offset significantly by sintering the chip at 350 °C for 30 s.

ACKNOWLEDGMENTS

The author would like to thank H. Ahmed at Cavendish Laboratory for many helpful discussions and for his technical support in the fabrication and characterization of GaAs MESFETs. Financial support from Ministry of Education, Islamabad is also acknowledged.

[1]E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts* (Clarendon, London, 1988).
[2]P. H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs* (Artech House, London, 1989).
[3]M. M. Ahmed, H. Ahmed, and P. H. Ladbrooke, J. Vac. Sci. Technol. B **13**, 1519 (1995).

[4]H. C. Card and E. H. Rhoderick, J. Phys. D **4**, 1602 (1971).
[5]C. Y. Wu, IEEE Trans. Electron Devices **41**, 5947 (1982).
[6]C. Chen, A. R. Calawa, W. E. Courtney, L. J. Mahoney, S. C. Palmateer, M. J. Manfra, and M. A. Hollis, IEEE Trans. Electron Devices **39**, 1797 (1992).
[7]Y. S. Lou and C. Y. Wu, IEEE Trans. Electron Devices **41**, 558 (1994).
[8]M. M. Ahmed and H. Ahmed, J. Vac. Sci. Technol. B **15**, 306 (1997).
[9]N. Braslau, J. B. Gunn, and J. L. Staples, Solid State Electron. **10**, 381 (1967).
[10]S. J. Chau and S. H. Lee, Thin Solid Films **200**, 211 (1991).
[11]R. A. McMahon, D. G. Hasko, and H. Ahmed, Rev. Sci. Instrum. **56**, 1257 (1985).
[12]P. H. Ladbrooke, Solid State Electron. **16**, 743 (1973).
[13]C. L. Chen and K. D. Wise, IEEE Electron Device Lett. **4**, 341 (1983).
[14]M. J. Howes and D. V. Morgan, *Gallium Arsenide, Material, Devices, and Circuits* (Wiley, New York, 1985).
[15]M. M. Ahmed, Ph.D. dissertation, University of Cambridge, Cambridge, 1996.
[16]M. M. Ahmed, J. Vac. Sci. Technol. B **15**, 1621 (1997).
[17]X. Wu and E. S. Yang, IEEE Electron Device Lett. **11**, 315 (1990).

Advanced emitters for next generation rf amplifiers

K. L. Jensen,[a] J. E. Yater, E. G. Zaidman, M. A. Kodis, and A. Shih
Code 6840, Electronic Science and Technology Division, Naval Research Laboratory, Washington, DC 20357-5347

(Received 10 September 1997; accepted 18 January 1998)

Next generation rf amplifiers, in particular the inductive output amplifiers (IOAs), will require electron sources characterized by high current density, high brightness, low emittance, and the ability to be emission gated. The strong interaction between the beam and the resonant or synchronous electromagnetic field may enable compact, highly efficient, and moderate gain *X*-band power booster amplifiers. An analysis of amplifier demands on generic emitter candidates is provided. Of the emitter candidates available, two (namely, field emitter arrays and wide-band-gap semiconductors) are amenable to an analysis predicated on a simple model of field emission from hyperbolas and ellipsoids. The simple model is used to investigate two problems of critical concern: for field emitter arrays (FEAs), we will investigate the conditions under which important space charge effects exist, and from the model predict optimum FEA performance characteristics for rf IOAs; for wide-band-gap materials, the simple model identifies parameters which limit the maximum electron transport across a metal-semiconductor interface, thereby limiting current flow. [S0734-211X(98)06904-2]

I. INTRODUCTION

The need for higher frequency and power, greater efficiency, and smaller size of next generation rf amplifiers place stringent demands on the performance and characteristics of electron sources. Present rf amplifiers require more than a 10:1 beam area convergence ratio. Given the present capabilities of thermionic emitters and the needs of future rf amplifiers incorporating traveling wave tubes (TWTs), large diameter electron beams and sophisticated electron gun designs are required. Advanced amplifier designs require smaller electron beam diameters in order to achieve higher power and frequency. Reducing the diameter of the (low current density) beam at the emitter to that required for the TWT beam tunnel requires a highly convergent magnetic field; relaxing the beam convergence required would reduce the complexity of the magnetic field profile and relax the machining tolerances required to avoid beam scalloping. Shortening the circuit would reduce the size of the magnet needed with consequent reduction in the weight and volume of the amplifier, and, in some cases, the power consumption. Uniformity of emission over the surface area of the emitter is therefore critical: a laminar and uniform beam results in the best beam transport for the least magnetic field and reduces losses to the tube walls. Likewise, reduced emittance electron beams are also critical. Emittance is caused by nonuniform emission and thermal velocities at the source, as well as nonaxisymmetric current density. In addition to the losses in amplifier performance incurred by poor emittance beams, further demerits include discrepancies in tube performance as compared to its modeled performance, as (due to the complexity of the modeling effort) uniform beams are typically presumed in the simulations.[1]

The relationship between several of the amplifier requirements and the demands thereby entailed of the emitters are listed in Table I. Our attention is restricted to TWTs, a broadband device whose circuit structure is a helix rather than a resonant cavity (Klystrode),[2] and which is utilized in the vacuum power booster component of the microwave power module (MPM).[3] In addition to the requirements of future amplifiers, the next generation MPM may incorporate emission gating of the electron beam to boost efficiency and further reduce the size of the helix circuit, which will place further austere demands on advanced emitters.[4–8]

Several emitter candidates vie for consideration, but two in particular are under consideration: Field emitter arrays (FEAs) promise exceptionally high current densities (>2000 A/cm^2), but improvements are needed in emission uniformity and robustness in the tube environment; wide-band-gap (WBG) materials show potential for high current density at low applied field, but the emission mechanisms and ultimate capabilities are poorly understood and not sufficiently quantified. A characterization of the theoretical performance of these candidates is desirable. In both cases, the predicted performance (power, efficiency, gain, length) of an amplifier is fundamentally tied to the performance characteristics of the emitter, in particular to the Fowler–Nordheim slope (B_{FN}) of the current–voltage relationship. For FEAs, the experimental B_{FN}, as well as A_{FN}, can be tainted by effects related to the high current density (space charge). For WBGs, the current through the metal-WBG tunneling barrier (the "Schottky barrier") limits the total current through the structure. A simple analytical model of emission current from hyperbolic and elliptical cones provides a theoretical tool capable of qualitatively analyzing both issues, albeit that the simplicity of the models precludes a quantitative treatment—it is widely appreciated that the complexity of emission from FEAs and into WBG materials obviates a literal interpretation of the models we provide. As justification for proceeding with geometrical models, we note that such

[a]Electronic mail: kevin.jensen@nrl.navy.mil

TABLE I. Amplifier demands on next generation emitters.

Requirement on source	Benefit
Smaller electron beam/ higher current density	Lowers beam convergence ratio, relaxes machining tolerances, gives better beam focusing
High transconductance electron source	Increases gain, reduces drive power
Beam prebunching	Increases efficiency, reduces TWT helix length
Reduced beam interception by gate and grid	Increased robustness of emitter and longer lifetime
Reduced emittance	Laminar and uniform beam give best beam transport for least magnetic field, losses to the walls reduced, better focusing, power output increases
Resistance to sputtering	Increased robustness of emitter and longer lifetime
Microcathode	Reduced cathode-gate transit time, high-frequency operation

models are the basis for much of the simulation work for using finite boundary and finite element approaches. For FEAs, the existence of migrating nanoprotrusions, and for WBG materials, the nonlinear Schottky barrier (which mandates a full quantum mechanical description using an effective potential including electron correlation and exchange effects), makes definitive statements based on abstract models rather perilous. Nevertheless, we argue that the reduction of a complex problem to simple-model components is warranted and has adequate pretext. Our goal here is to validate qualitative and simple models, and indicate where their weaker components can be augmented with more comprehensive treatments.

Attention shall be restricted to emitter candidates which show potential for providing levels of performance demanded by particular rf amplifier devices. We shall briefly discuss the requirements of an emission gated TWT (Twystrode) on the emitter candidate in terms of compactness, efficiency, and gain, and provide a simple and conservative analytical model for predicting required emitter characteristics. Analytical formulas for calculating total current from rotationally symmetric geometrical surfaces shall be given. From the model we will estimate whether an experimental arrangement for characterizing FEAs may be affected by space charge effects for ring cathodes appropriate for a next generation vacuum power booster. Further, we will provide a simple estimate of the limiting current through a metal-semiconductor interface based on barrier height, doping density, and interface roughness, and ascertain what combination of parameters will theoretically result in emitter characteristics suitable for a next generation VPB.

II. AMPLIFIER PERFORMANCE AND CATHODE REQUIREMENTS

The principles of Twystrode operation have been extensively discussed and shall only be hi-lighted here. In a TWT, a continuous electron beam is bunched by its interaction with the slow wave supported by the helix in a process referred to as "velocity modulation" of the electron beam. For the last portion of its flight, the kinetic energy of a bunch is then extracted from the beam and converted to rf power. In contrast, "density modulation" occurs by modulating the extraction voltage so that the beam is bunched at the emitter itself, eliminating that portion of the helix which is employed to prebunch the beam. In addition to a reduction in size, density modulation also largely eliminates the undesirable velocity spread, enabling a more radical taper on the helix and therefore a greater conversion of beam power to rf power.

The efficiency of the Twystrode amplifier has been found to depend strongly on the ratio between the peak and average current. If the gate voltage is sinusoidally varied, i.e., if $V_{\text{gate}} \equiv V_{\text{pk}} - V_{\text{rf}}[1-\cos(\omega t)]$ and the relationship between the gate voltage and the emitted current follows a Fowler–Nordheim relationship, i.e., $I(V) = A_{\text{FN}} V^2 \exp(-B_{\text{FN}}/V)$, then it may be shown that the ratio may be well approximated by

$$\frac{I_{\text{ave}}}{I_{\text{pk}}} = e^{-z} I_0(z) + [(2z+1)I_1(z) - 2zI_0(z)]ze^{-z} \frac{(B_{\text{FN}} + V_{\text{pk}})V_{\text{pk}}}{(B_{\text{FN}} + 2V_{\text{pk}})^2},$$
$$z = \frac{V_{\text{rf}}}{V_{\text{pk}}}\left(2 + \frac{B_{\text{FN}}}{V_{\text{pk}}}\right), \tag{1}$$

where $I_n(z)$ is a Bessel Function of order n. Equation (1) may be used to deduce the voltage amplitude V_{rf} from the current ratio, from which the drive power may be estimated.

The evaluation of the output power is a nontrivial affair and is best accomplished using sophisticated computer models of the beam-wave interaction present in the TWT.[9,10] Nevertheless, an analytical estimate may be obtained by integrating the electron equations of motion, which govern the evolution of the electron velocity v, the electric field E, and the electron phase ϕ as a function of the axial coordinate z [not to be confused with the z of Eq. (1)]; they are given by[11]

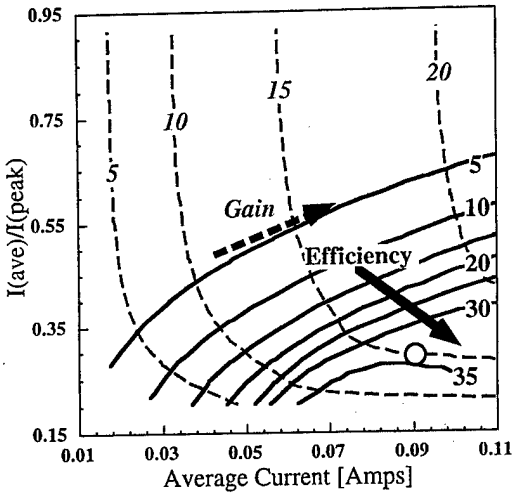


FIG. 1. Simulation of gain and efficiency as a function of average current and average to peak current ratio for an FEA-based TWT (Twystrode). Circuit and array parameters correspond to Table II. The small circle in the lower right-hand corner represents the amplifier target of 35% efficiency and 15 dB gain.

$$\frac{\partial}{\partial z} v(z) = -\frac{e}{m v_{\mathrm{ph}}(z)} E(z)\cos(\varphi(z)+\theta),$$

$$\frac{\partial}{\partial z} \varphi(z) = \left(\frac{\omega}{v_{\mathrm{ph}}}\right)\left(\frac{v(z)}{v_{\mathrm{ph}}(z)} - 1\right), \tag{2}$$

$$\frac{\partial}{\partial z} E(z) = \frac{Z(z) c \omega^2}{v_0 v_{\mathrm{ph}}(z)} \frac{1}{2\pi} \int_{-\pi}^{\pi} d\omega t\, I(\omega t) v(z) \cos(\varphi(z)+\theta),$$

where v_0 is the initial velocity of the electron, I is the current, ω is the frequency; the phase velocity v_{ph} and the impedance Z are dependent on the helix geometry and quadratic taper. A Runge–Kutta solution of Eq. (2) is obtained by using a 24-point Gaussian Quadrature scheme to evaluate the integral. The resulting simulation, CASSANDRA, though pessimistic by comparison to more thorough numerical simulations, is fast and provides estimates of the drive power and optimum length. By "pessimistic," it is meant that the beam current is typically more taxing on an emitter candidate for a given gain and efficiency than would be obtained from advanced codes because the output power is underestimated by comparison to more comprehensive 1D codes;[12] nevertheless, CASSANDRA gives qualitatively correct behavior, and is therefore utilized to infer required emitter characteristics given particular amplifier specifications here. A more complete description of the code and its methodology shall be given elsewhere.[13] A typical calculation is shown in Fig. 1, in which a desired gain and efficiency for a Twystrode (Table II) is used to extrapolate the required average current and current ratio demanded of the emitter.

Analytical estimates of current from rotationally symmetric hyperbolic and elliptical geometries, and issues associated with area and field enhancement factors, have been investigated by several authors.[14–18] We shall sketch the methodology employed previously.[19,20] If the conducting surfaces are parallel to an axis of an orthogonal coordinate system in which Laplace's equation is separable, then the solution of the potential boundary value problem is straightforward, as is finding the field along the conducting surfaces.[21] For a differential surface element $d\Omega$, the differential current element dI is related to the current density $J_{\mathrm{FN}}(F) \equiv a_{\mathrm{fn}} F^2 \exp(-b_{\mathrm{fn}}/F)$[22] by $dI = J_{\mathrm{FN}}[F(\Omega)]d\Omega$. In the coordinate system defined by $z = a\cosh(\alpha)\cos(\beta)$ and

TABLE II. Twystrode input parameters for "CASSANDRA."

Symbol	Definition	Value
Amplifier specifications		
η	Electronic efficiency	35%
Gain	Gain	15 dB
P_{rf}	Output power	82 W
Emitter characteristics		
A_{FN}	Fowler–Nordheim intercept	3.4 A/V^2
B_{FN}	Fowler–Nordheim slope	321 V
I_{pk}	Peak current	338 mA
$I_{\mathrm{ave}}/I_{\mathrm{pk}}$	Current ratio	0.28
Circuit parameter		
ω	Frequency	10 GHz
ω_0	Resonant frequency	10 GHz
ω_{bw}	Band width	30% of ω_0
V_{anode}	Anode voltage	2.5 kV
Z_0	Initial circuit impedance	2.5 Ω
R_0	Circuit resistance	1.4 Ω
Ring parameters		
d_{tt}	Tip-to-tip distance	0.50 μm
d_{bg}	Oxide thickness	0.30 μm
L	Ring outer diameter	610 μm
W	Ring width	30 μm

$\rho = a \sinh(\alpha) \sin(\beta)$, surfaces of constant α define rotationally symmetric hyperbolic structures (to be used to model FEAs) and surfaces of constant β likewise define rotationally symmetric ellipsoidal structures (to be used for the WBG simulations). If the field at the apex of each structure is defined as F_{tip}, then for each geometry, the field along the surface of the emitter is

$$F(\alpha,\beta_0)|_{\text{hyperbola}} = F_{\text{tip}} \frac{\sin \beta_0}{\sqrt{\sin^2 \beta_0 + \sinh^2 \alpha}},$$
$$F(\alpha_0,\beta)|_{\text{ellipse}} = F_{\text{tip}} \frac{\sinh \alpha_0 \sin \beta}{\sqrt{\sin^2 \beta + \sinh^2 \alpha_0}}, \tag{3}$$

where β_0 is the cone angle of the hyperbola, and $\coth \alpha_0 = R$ is the ratio of the major to the minor axis of the ellipsoid. Expressions for F_{tip} depend on the emitter under consideration. The current from a single emitter is then given by

$$I_{\text{tip}} = J_{\text{FN}}(F_{\text{tip}}) \int_0^{x_{\max}} \left(\frac{F(x)}{F_{\text{tip}}}\right)^2 \times \exp\left(\frac{b_{\text{fn}}(F(x)-F_{\text{tip}})}{F_{\text{tip}}F(x)}\right) 2\pi\rho \, dl(x), \tag{4}$$

where $x = \alpha$ or β depending on whether a hyperbolic or elliptical emitter is being considered. We have used $d\Omega = 2\pi\rho \, dl$, where the differential length is given by $[1 + (\partial_\rho z)^2]^{1/2} \partial_x \rho \, dx$. The integral is typically referred to as the "area factor" and is tip-field dependent. Because of the rapidity with which the exponential vanishes in Eq. (4), the log of the integrand may be Taylor expanded about $x = 0$ to first order, the upper limit may be approximated by (∞) and Eq. (4) integrated, thereby providing an analytical estimate to tip current.

For simulations reported herein, Eq. (4) invokes the Fowler–Nordheim representation of the current density J as a function of applied field. In standard boundary element and finite element simulations of field emitter arrays, this represents common practice. For the WBG studies investigated here, the applicability of such an approach is far less clear-cut, and requires some justification. By virtue of its inherent complexity, the description of electron transport from metal into semiconductor under reverse bias (emitter) conditions is impeded by the difficulty in correctly accounting for a myriad of effects, from interface roughness, tunneling, and statistical effects due to the fermionic nature of the electrons which affects the effective potential from which the tunneling current is calculated. Further, because of the reliance on reverse biasing across the Schottky barrier, an optimal WBG emitter seeks to magnify the "roughness" of the metal-semiconductor interface, making the generalization of any coherent one-dimensional theory to the actual three-dimensional interface non-trivial at best; for forward biasing (electron transport from WBG to the metal), where the interface is desired to be much smoother, one-dimensional theories have greater reign. Finally, in the presence of current flow, the barrier at the metal-semiconductor contact changes by comparison to the static case.[23] Recourse to a Fowler–Nordheim equation (even with the modifications we propose below), in light of these difficulties, represents a remarkably simplistic approach to an inordinately complex problem; the motivation for proceeding therefore requires discussion.

Qualitatively, tunneling through the metal-semiconductor barrier governs the behavior of the current–voltage relationships into a WBG material.[24] A fundamentally superior estimation of current density is necessitated by explicitly including electron exchange and correlation potentials[25] into the effective potential used in the one-dimensional Schrödinger's equation, from which the tunneling current from the metal into the conduction band may be calculated.[26–28] Embarking on such a full-blown quantum mechanical treatment is not our aim, however. Our intention is to investigate the utility of an interface roughness model, for which a qualitative description of the tunneling current is adequate: the reduction of a complex problem to simple-model components to be validated, and subsequently augmented with successively better approximations, is warranted. A qualitative model which is amenable to future refinement ("bottom-up" approach) rather than a complete theory for which the ellipsoidal model is an asymptotic particular case ("top-down" approach), is indicative of our desire to identify those factors which most affect total current and their relative impact, and represents a pragmatic approach to the problem of finding the distribution of electrons injected into the conduction band of the WBG, as needed for electron transport simulations (e.g., Monte Carlo) to ascertain emitted current.

III. FEA CHARACTERIZATION AND SPACE CHARGE

A. Analytical model of a field emitter

To a good approximation, the field at the apex for a gated structure may be shown to be

$$F_{\text{tip}} = \left(\frac{\pi}{\ln(ka_g/a_s)} - \tan^2 \beta_c\right) \frac{V_{\text{gate}}}{a_s}, \tag{5}$$

where a_g is the gate radius, a_s is the tip radius, k is the field factor determined from boundary element simulations[29] (although an analytical form has been developed; see Ref. 13) and β_c is the cone angle.

A distribution in the B_{FN} and A_{FN} parameters may be accounted for. Departing from previous work,[20] which considered only a linear distribution in B_{FN}, the present treatment considers the sources of that distribution in tip radius (a_s) and incorporates its effect on B_{FN} and A_{FN} explicitly, which is a fundamentally superior approach. For simplicity, assume that no variations in the work function occur. Assume a linear distribution in the value of tip radii exits such that $a(s) \equiv a_s(1+s)$, where $0 \leq s \leq \Delta s$. Then we have

$$I_{\text{array}} = \frac{N_{\text{tips}}}{\Delta s} \int_0^{\Delta s} I_{\text{tip}}[F(s, V_{\text{gate}})] ds, \tag{6}$$

where I_{tip} is as given in Eq. (4). A linear distribution in tip radii is not equivalent to a linear distribution in B_{FN} values. Further, such a linear distribution may not be the most plau-

sible one, but, as observed in the investigation of various distributions (e.g., Gaussian, Top Hat, double delta) in B_{FN}, equivalent behaviors result. It is therefore expedient to choose that distribution which transparently reveals the length scale of the sharpest of the nanoprotrusions. The radius of curvature should not be too much smaller than the scale of the nanoprotrusions presumed to give rise to the current; consequently, a lower bound of approximately 30 Å occurs. A linear distribution transparently incorporates such a lower bound whereas, say, a Gaussian distribution would not.

By noting that I_{tip} rapidly decreases with increasing tip radius, and performing a Legendre least-squares fit, we may cast I_{array} in the Fowler–Nordheim form $A_{FN}V^2 \times \exp(-B_{FN}/V)$, where (in atomic units[30])

$$A_{FN}=2\pi(a_s\cos\beta_c)^2 N_{tips} a_{fn}\left(\frac{\beta_g^3}{x_0 b_{fn}}\right)\exp\left[2+\frac{4\delta^2}{3x_0^2}\right]\times\left(\frac{e^{b_0\Delta s}-1}{b_0\Delta s}\right),$$

$$B_{FN}=\frac{b_{fn}}{\beta_g}+\frac{\delta^2+2x_0^2}{x_0^3}, \tag{7a}$$

where $\delta=V_{rf}/V_{pk}(V_{pk}-2V_{rf})$ and $x_0=(V_{pk}-V_{rf})/V_{pk}(V_{pk}-2V_{rf})$, and a_{fn} and b_{fn} are Fowler–Nordheim terms. While Eq. (7a) superficially resembles its analog in Ref. 20, there are pronounced differences which are implicit in the last term of A_{FN}. The distribution factor b_0 is a complicated function of tip field, field enhancement factor, cone angle, and b_{fn}, rather than an adjustable parameter. Under the assumption that $\beta_g \sin(\beta_c)\ll 1$, it is approximately equal to

$$b_0=2+\left(3+\frac{x_0 b_{fn}}{\beta_g}\right)\left[\frac{(\beta_g a_s+\tan^2\beta_c)^2}{\pi\beta_g a_s}-1\right]. \tag{7b}$$

B. Tip radius

The value of the tip radius may be inferred by comparing Eq. (7a) to the slope of experimental data cast in a Fowler–Nordheim-like form. Typical values range between 40 and 100 Å, which appears to be in stark contrast to scanning electron microscope photographs and reported values (but not always[31]). However, emitter tips have undulations, grain boundaries, and nanoprotrusions which are expected to be on the order of 40–100 Å; in fact, Purcell *et al.*,[32] and separately, Charbonnier[33] have argued that field emission from nanoprotrusions, whose radii of curvature are comparable to values claimed above, exist on a field emitter, and further, that these nanoprotrusions migrate along the emitter shank towards the emitter tip, that several nanoprotrusions contribute to I_{tip}, and finally, that emission directed toward the gate may initiate circumstances which lead to the formation of an arc. A smooth emitter without surface protrusions with a work function of clean molybdenum is presupposed by the hyperbolic tip model; high resolution electron micrographs of molybdenum tips belie this, as typical emitters have larger tip radii but show surface undulations, and emitted current shows noisy behavior compatible with the migrating nanoprotrusion model. Consequently, a hyperbolic emitter with tip radius a_s should be taken as a heuristic, albeit useful, model, and "tip radius" should be considered shorthand for the effective, not actual, tip radius.

C. Space charge

Two sources of space charge effects exist which may affect the characterization of the emitter[20]; we shall simply summarize them here. In the first, space charge effects will spoil a Fowler–Nordheim characterization of an array if the anode field is insufficient to draw the emitted current away, causing current to be reflected back to the gate: the anode current is then limited by Child's law.[34] A Fowler–Nordheim plot of the anode current versus the gate voltage will then exhibit negative curvature at large gate voltages. This regime occurs when the total emitted current from a tip exceeds I_{child} (where I_{child} is appropriately modified from its one-dimensional form to accommodate beam spreading) and a virtual cathode forms above the gate which reflects charge back to the gate. The second effect is due to the depression of the tip field due to the presence of emitted charge. The gate current is calculated from $I_{gate}=I_{tip}-I_{child}$ (I_{gate} is zero if the right-hand side is negative). Near the space charge limited region, the sheets of charge present between cathode and anode will cause a greater number electrons emitted with an axial velocity to be deflected back to the gate, and so the gate current should rise though this effect is absent in the simple model.

That these effects will negatively impact the Fowler–Nordheim parameterization of the FEA can be inferred, as a simple linear fit would underestimate B_{FN} and A_{FN}, even providing different estimates of these parameters for different anode potentials, though it is known that anode field does not significantly contribute in the extraction of current from an FEA. In other words, a simple Fowler–Nordheim parameterization of an emitter will depend on the anode voltage used, even though the anode voltage does not significantly contribute to the magnitude of the field at the emitter tip. Rather, the anode voltage is used to draw away the emitted current; if it is insufficient, then the onset of space charge effects will affect a least-squares fit used to give the FN parameters, and the gate current will rise. A rise in gate current to high values is known to be correlated with the onset of processes such as the release of adsorbates which can contribute to arc formation, thereby destroying the array. Identifying the space charge regime is therefore critical.

D. Characterization of an IOA-compatible FEA

A study of the experimental data examined in Ref. 19 indicates that, for example, estimates of B_{FN} and A_{FN} can be off by as much as 16% in the former and a factor of 9 in the latter when space charge effects are present. The significant differences in fitting parameters produce large differences in estimates of transconductance. The turnover at high voltages, typical of space charge effects, affect the Fowler–Nordheim

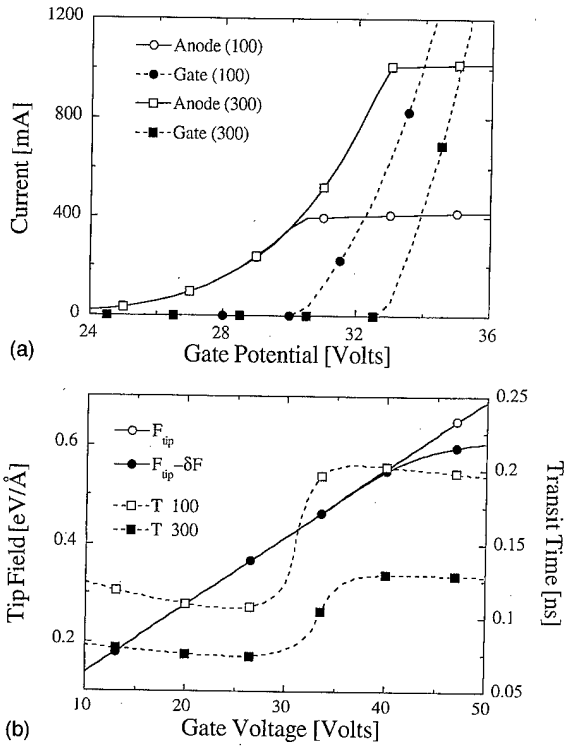


FIG. 2. (a) Current vs gate potential for anode voltages of 100 and 300 V for an FEA adequate for the rf amplifier characteristics corresponding to the circle in Fig. 1, and parameters given in Tables II and III. In the one-dimensional model used, gate current is nonzero only when the emitted current exceeds the space charge limit as given in Ref. 20. (b) Transit time from gate to anode and tip field at emitter apex for the array of (a). Transit time shifts from a space-charge free behavior to a space charge limited behavior at a gate voltage of approximately 30 V, corresponding to the onset of a plateau for the anode current in (a). Tip field suppression, due to the presence of charge between the anode and the gate, becomes significant at high currents.

parameters, thereby affecting estimates of the array's transconductance, a crucial parameter in judging the array's rf compatibility.

In Fig. 2(a), the effects of space charge on the characterization of an FEA-VPB compatible array for the parameters listed in Tables II and III are shown, in which the total (tip) current is equal to the sum of the gate current and the anode current. For IOA operation, the FEA must produce I_{pk} at a gate voltage of $V_{pk}=V_g$. For the values of anode voltage and anode to gate separation used, the anode field will be insufficient to draw the emitted current away, i.e., the gate current will rise. Consequently, the characterization will be negatively affected. The second source of space charge effects, namely, tip field suppression (denoted by δF), is seen in Fig. 2(b) to not contribute significantly for the anode potentials chosen, and for currents needed for TWT operation, is not expected to.

For the parameters chosen, an anode field on the order of 1.4 kV/cm near the array is required to prevent the return of electrons to the gate at an emission current of 400 mA. Lower fields will not prevent the formation of a virtual cathode above the array which would serve to deflect electrons to the gate. Such an analysis is in qualitative agreement with the findings of Imura *et al.*, who investigated the relation of anode potential to space charge effects in the operation of an FEA-based TWT.[35] Even if the value of I_{child} is overestimated due to an overestimation of the "beam spread factor,"[20] the design value for the anode field of 14 kV/cm, as used in the proposed emission gated Twystrode,[6] is comfortably beyond the minimum field identified above.

TABLE III. FEA input parameters for space charge simulation.

Symbol	Definition	Value/range
ϕ	Work function	3.2 eV
a_s	Tip radius	36 Å
a_g	Gate radius	0.15 μm
β_c	Cone angle	15°
k	Field factor	6
V_{rf}	Modulation amplitude	6 V
V_{pk}	Peak voltage	34 V
C_{unit}	Unit cell capacitance	$\approx 4d_{tt}^2/\epsilon_0 d_{bg}$
D	Anode-gate separation	500 μm
N_{tips}	Number of emitters	$\pi[L^2-(L-W)^2]/d_{tt}^2$
Δs	Distribution of radii	0.2

IV. CURRENT THROUGH A SEMICONDUCTOR-METAL INTERFACE

A. Background

WBG materials are being explored as possible candidates for advanced emitter structures for vacuum electronic devices, but the emission mechanisms and ultimate capabilities are poorly understood. Once electrons have been transported across the metal/semiconductor barrier, the bulk transport properties and surface properties of the wide-band-gap material determine the quality of the electron emission that can be achieved. Several wide-band-gap materials (diamond, AlN, and BN) have been shown to have a low or negative electron affinity that results in a small or negligible energy barrier at the surface/vacuum interface; electron emission into vacuum is anticipated to be more easily accomplished at such surfaces. While wide-band-gap materials possess many desirable properties, their capabilities as electron emitter devices must be well characterized to understand the emission process and thereby optimize the emitted beam. The appropriate surface conditions must be determined for producing strong, stable emission into vacuum, and a low or negative electron affinity surface must be identified that is stable and robust in the operating environment. The surface morphology and compositional uniformity must be characterized in order to examine the factors that determine the emission uniformity, and the scattering and recombination processes must be understood that affect the energy distribution of the transported electrons and the overall transport efficiency to the surface. Finally, the conduction path through the emitter structure must be identified since a material with nonuniform composition or various defects may result in electron transport paths that are not present in uniform, homogeneous material. The methodology for theoretically understanding elec-

tron transport through the WBG material, e.g., by Monte Carlo simulations[36,37] is well developed. In contrast, the understanding of electron injection mechanisms within the cathode structure to provide sufficient electron current to the emitter surface is less understood, and requires a theoretical foundation beyond the one-dimensional models typically invoked.

We have examined the secondary emission properties of single-crystal and chemical vapor deposition (CVD) diamond samples. Some surfaces of diamond are found to produce extremely high secondary yields, where yield is defined as the ratio of the total emitted electron current to the incident electron current. However, high secondary emission yields are only observed from diamond surfaces observed to have a low or negative electron affinity. It is clear that a low or negative electron affinity is critical to the emission of these low-energy electrons, and in fact, the electron yield decreases rapidly as the electron affinity increases from negative to positive. It also appears that the transport of the low-energy secondary electrons is very efficient in the diamond samples since the measured emission yields are not significantly affected as the electron transport distance is increased. While the energy distribution of the secondary electrons is centered near the bottom of the conduction band, the electron-hole recombination time in the diamond is long enough to permit the transport of these electrons to the surface, at least for the range of transport distances we are concerned with here. Consequently, we shall treat electron motion within the WBG material as essentially ballistic, assume that the conduction band is aligned with E_{vac} (i.e., no surface barrier), and concentrate our efforts on analyzing electron tunneling into the conduction band of the semiconductor from a metal.

B. Potential model

If the interface between a metal and a wide-band-gap (*n*-doped) semiconductor is visualized as an array of protrusions existing in a field generated by the depletion region, as suggested by Geiss *et al.*,[38] and exploited by Lerner *et al.*,[39] then the simple FEA theory mentioned previously may be used to estimate electron transport into the wide-band-gap material. Here, however, instead of modeling the protrusions as hyperbolic emitters, it is more convenient to model them as ellipsoids in the presence of a background electric field. The limiting case of the potential for a sphere in a background field is a standard problem in electrodynamics, solvable using a Legendre polynomial expansion[40]; the use of an ellipsoid is a generalization of that problem using the oblate spherical coordinate system.

We shall assume the simplest model of the interface, shown in Fig. 3, to calculate the transport of electrons through the contact barrier; as argued previously, we may assume that the transport of electrons through the semiconductor is ballistic, and further, that any surface barrier at the vacuum interface has a negligible impact on the current collected by the anode (a low- or negative-electron-affinity approximation). The background electric field in the ellipsoidal

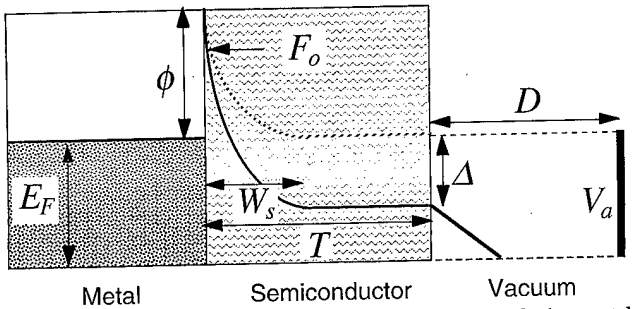


FIG. 3. Schematic one-dimensional potential model of the metal-semiconductor-vacuum system, indicating barrier height ϕ, depletion width W_s, and bias Δ.

model of surface roughness, as shown in Fig. 4, is then given by the fields which exist near the metal-semiconductor interface, and is due to the existence of the depletion region. The problem superficially resembles a field emission problem, except that the potential profile near the contact barrier, rather than being linear as for field emission, is approximately quadratic.

Under the assumption that the electron density in the depletion region vanishes, and is equal to the doping density outside the depletion region, Poisson's equation is readily solved; defining

$$W_s = \sqrt{\frac{K_s(\phi+\Delta)}{2\alpha_{fs}\pi\hbar c N_D}} \tag{8}$$

allows us to conveniently express the potential $V(x)$ as

$$V(x) = \mu - \Delta + (\phi+\Delta)\left(1-\frac{x}{W_s}\right)^2. \tag{9}$$

For $x > W_s$, we assume that $V(x) = \Delta$, which we shall approximate by $\Delta = V_a T/D$, where T is the thickness of the sample, and D is the surface to anode separation, as in Fig. 3. The background field that the ellipsoids experience is then $F_0 \equiv -\partial_x V(x=0) = 2(\phi+\Delta)/W_s$. In this formulation, the

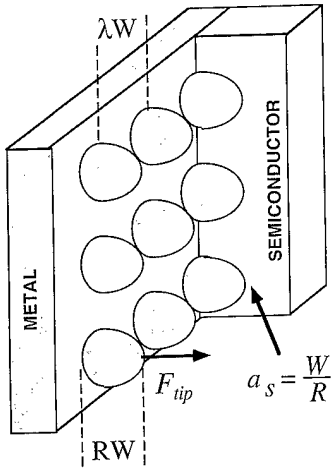


FIG. 4. Idealized model of metal-semiconductor interface roughness. Compared to the width of the ellipsoids, R controls ellipsoid height, and λ controls the spacing of the ellipsoids on a square grid.

problem superficially resembles field emission into the vacuum, and therefore, the analysis behind the Fowler–Nordheim equation may be brought to bear.

C. Field enhancement and area factor

Proceeding as for the hyperbolic emitter, the potential in the vicinity of the boss is

$$V(\alpha,\beta)=-F_0 RW\cos(\beta)\left(\frac{\eta}{\eta_0}-\frac{Q_1[\eta]}{Q_1[\eta_0]}\right), \tag{10}$$

where (α,β) are the spheroidal coordinates, W is the radius of the minor axis of the ellipse, $Q_1(x)=\frac{1}{2}x\ln[(x-1)/(x+1)]-1$ is a Legendre polynomial, $\eta=\cosh(\alpha)$, and $\eta_0=R/\sqrt{(R^2-1)}$. Taking the gradient of Eq. (10) and evaluating it at the apex of the ellipsoidal boss gives

$$F_{\text{tip}}=\frac{F_0}{(\eta_0^2-1)Q_1(\eta_0)}. \tag{11}$$

Unlike the analogous case for FEAs, F_{tip} does not vanish as the anode potential vanishes, due to the behavior of the potential in the depletion region. This will have consequences for the Fowler–Nordheim representation of the current–voltage relationship, in particular, the low voltage region, an issue to which we shall return. Solving for the area factor gives

$$b_{\text{area}}\approx\pi R a_s^2\sqrt{\frac{2\pi F_{\text{tip}}}{F_{\text{tip}}+b_{\text{fn}}}}, \tag{12}$$

where, analogous to the field emitter case,[19] the area factor depends upon the applied field, though for field emitters, the area factor is, to leading order, proportional to F_{tip}, rather than $\sqrt{F_{\text{tip}}}$, as here.

D. Current from an area

A complication which exists for the analysis is that the tunneling barriers may be small (i.e., ϕ may be small). Consequently, we shall use the linearized WKB approximation to the current density[41]

$$J_{\text{LWKB}}(F)=a_{\text{fn}}F^2e^{-b_{\text{fn}}/F}\frac{c_{\text{fn}}\pi/\beta_T}{\sin(c_{\text{fn}}\pi/\beta_T)}\times(1-[c_{\text{fn}}\mu+1]e^{-c_{\text{fn}}\mu}),$$
$$c_{\text{fn}}=\frac{2}{\hbar F}\sqrt{2m\phi}t_0, \tag{13}$$

where $\beta_T=1/k_BT$ is the inverse temperature and $m=$ the effective mass of the electron within the semiconductor; for small barrier heights, the c_{fn} terms cannot be ignored, as is typically done for FEAs. A further complication is introduced by the quadratic nature of the potential. An analysis of the derivation which leads to the Fowler–Nordheim equation for J_{FN} shows that it is sufficient to augment the WKB transmission coefficient via $T_{\text{WKB}}(E)\approx\exp[-c_0-c_1(\mu-E)]\Rightarrow\exp[-c_0-\theta_0-(c_1+\theta_1)(\mu-E)]$, where

$$\theta_0=\frac{\sqrt{2m}}{\hbar}\sqrt{(\phi+\Delta)\left(1-\frac{x_0}{W_s}\right)^2-\Delta}\,(x_{\max}-x_0),$$
$$\theta_1=\frac{m}{\hbar^2\theta_0}(x_{\max}-x_0)^2, \tag{14}$$

where $x_{\max}=W_s\{1-\sqrt{[\Delta/(\phi+\Delta)]}\}$ occurs where $V(x_{\max})-E=0$, and x_0 occurs where the sum of $V(x)$ and the image charge contribution are maximized. In essence, a component to the integral in the WKB approximation has been added to account for the greater width of a quadratic barrier as opposed to a linear one. Consequently, the current from an array of elliptical hemispheres is given by $I_{\text{WBG}}\approx C_W N_{\text{tips}} b_{\text{area}} J_{\text{LWKB}}(F_{\text{tip}})$, where C_W accounts for the modifications to the transmission coefficient and is equal to

$$C_W=e^{-\theta_0}\left(\frac{F_{\text{tip}}c_{\text{fn}}}{3\theta_1F_0+2F_{\text{tip}}c_{\text{fn}}}\right)^2. \tag{15}$$

E. Simulation of current–voltage characteristics

By examining the parametric dependence of I_{WBG}, we may draw several conclusions about how the current through the metal-semiconductor barrier may be increased: (i) Reducing the work function lowers the barrier height; (ii) roughening the surface (increasing R) increases the field enhancement factor, hence F_{tip}; (iii) increasing the doping density shortens the barrier width; (iv) reducing the dielectric constant makes the image charge lowering of the barrier more pronounced and reduces the barrier width. These observations are borne out in simulations and displayed in Fig. 5 for the parameters listed in Table IV.

The choices in parametric values endeavor to replicate the literature: The work function of 4.0 eV is analogous to Geis *et al.* and Pupeter *et al.*[42] Variation in the work function down to 2 eV follows Choi *et al.*[43] and Lerner *et al.* Temperature is room temperature. The ellipsoidal separation to width ratio λ is an adjustable parameter and was inferred from the grit size and surface roughness arguments of Geis *et al.* (Lerner *et al.* assumes a much lower work function of 1.9 eV, and therefore interprets λ to be 800, reflecting the extreme dependence of current on barrier height). The electron effective mass ratio of 0.2 follows Choi *et al.* and Givargizov *et al.*[44] The dielectric constant of 10.5 was chosen high to allow for more general WBG materials, and represents a departure from diamond-like parameters, for which it is 5.5. Choosing the doping density in the 10^{18} cm^{-3} regime is common. The ellipsoidal height to width ratio of 1.25 is an adjustable parameter, though we constrained ourselves to hemisphere-like parameters. The ellipsoidal width of 60 Å is near Geis *et al.* (100 Å) and Lerner *et al.* (82 Å). Sample thickness was chosen at 25 μm, reflective of our experimental work, but not significantly different than the 10 μm maximum value utilized by Cutler *et al.*[45] Anode spacing followed Kang *et al.* Area was chosen as 1 mm^2, a display pixel size, and similar to the value used by Zhu *et al.*[46] In the simulations, "baseline" was taken to mean representative

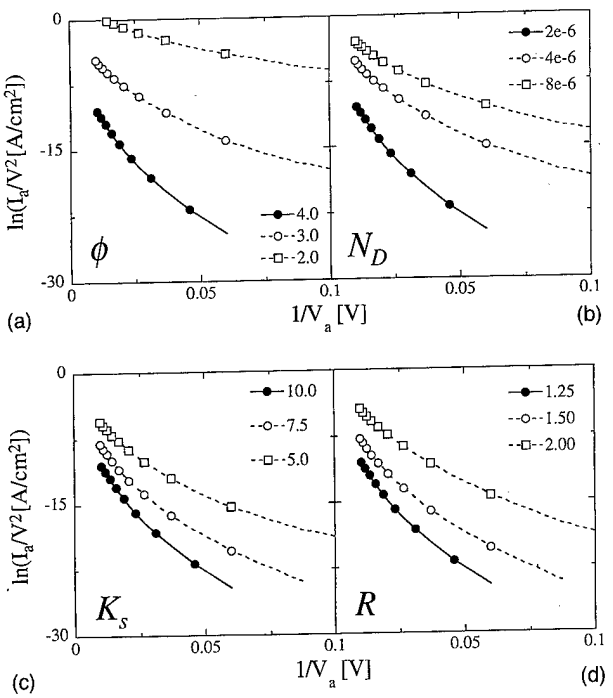


FIG. 5. (a) Total current through a 1 mm^2 sample for baseline parameters with work function (height of Shottky Barrier) variation; current increases as the barrier height drops. (b) Variations in the doping density; current increases as the barrier width decreases, as per Eq. (8). (c) Variations in the dielectric constant; current increases as the barrier width decreases, as per Eq. (8). (d) Variations in the ellipsoidal height; current increases as the tip field increases, as per Eq. (11).

points (similar to the literature) around which we may vary our parameters for particular cases, not a particular case itself.

A current of 1.8 mA is obtained from a 1 mm^2 area, corresponding to a current density of 180 A/cm^2 for an anode at approximately 53 V and 100 μm away from the surface. A Fowler–Nordheim fit gives $A_{FN} \approx 33\ \mu A/V^2$, and $B_{FN} \approx 246$ V. Small changes in ϕ, R, N_D, and K_s can cumulatively have a large effect on increasing the current. Consequently, scaling the parameters to a micron-sized gated structure suggests that WBG emitters may be appropriate candidates for next generation IOAs.

TABLE IV. Baseline WBG parameters.

Symbol	Definition	Value/range
f	Barrier height	4.0
kBT	Temperature	0.0259 eV
l	Ellipsoid separation to width ratio	2
r	Electron effective mass to rest mass ratio (m^*/m_0)	0.2
Ks	Dielectric constant	10.5
ND	Doping density	2×10^{18} #/cm^3
R	Ellipsoid height to width	1.25
W	Ellipsoid width radius	60 Å
T	Sample thickness	25 μm
D	Anode-surface separation	100 μm
Area	Emitter area	1 mm^2

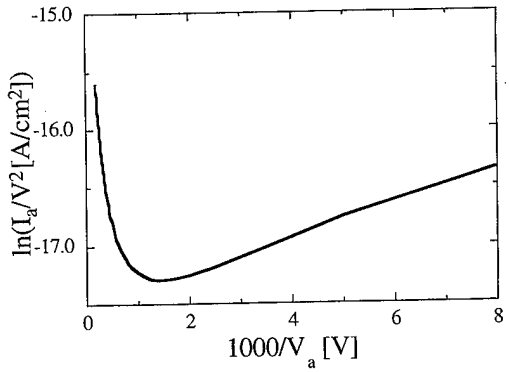


FIG. 6. Behavior of Fowler–Nordheim representation of $I_{anode}(V_{anode})$ due to the nonvanishing of F_0 as V_{anode} vanishes. In the low-V regime, current becomes approximately linear with respect to voltage.

A novel and interesting consequence of the behavior of F_0, namely, that it does not vanish as the anode potential vanishes, couples with the behavior of c_{fn} to produce an interesting artifact in a Fowler–Nordheim plot of current versus voltage, namely, the tail associated with the low voltage region of the Fowler–Nordheim plot. At small anode potentials, the current becomes approximately linear with respect to the anode potential, causing a deviation from the linear relationship of $\ln(I/V^2)$ vs $1/V$. For diamond-like parameters at a high (n-type) doping density, the results are shown in Fig. 6. The parameters for this simulation, where different from Table IV, are $D=2000\ \mu$m, $\lambda=3.4$, $T=0.5\ \mu$m, $W=80$ Å, $K_s=7.0$, $\phi=5.5$ eV, and $N_D=5\times10^{19}$ #/cm^3. This behavior resembles similar behavior found by Kang *et al.*[47] though it is important to bear in mind that Kang's study was for p-type diamond, for which the transport of electrons through and the potential within the diamond is significantly different than the n-type model assumed here. It is suggested that the origin of the tail in the low voltage regime is due to the nature of the potential barrier between the metal and semiconductor (i.e., the field at the barrier does not vanish as the anode voltage vanishes), rather than the behavior of the bulk semiconductor potential (for which the differences between p- and n-type behavior are pronounced), and to the extent that p- and n-type barriers at the semiconductor-metal interface are analogous,[48] their effect on tunneling current share certain analogous features. The existence of a surface barrier at the vacuum interface would cause the upward climb to cease at sufficiently low anode voltages.

V. CONCLUSION

The requirements placed on advanced cathodes by next generation inductive output amplifier designs was explored using a simple model of an emission-gated TWT. The performance levels targeted by that analysis (particularly the Fowler–Nordheim characterizations) were translated into material and geometry requirements of two electron emitter candidates, FEAs and WBGs, using an analytical model of the emission from a field emitter tip and from interface roughness, respectively. The analytical models of the emit-

ters were applied to two problems. First, the analytical theory of a unit cell field emitter was used to calculate tip field and array current, and was applied to show how the onset of curvature at high gate voltages affects $I(V)$ data for FEAs needed in an IOA. Unit cell effects can also contribute to tip field suppression in addition to that contributed by charge between the gate and anode. For IOA requirement levels, it was shown that unless care is taken to insure that the extraction field is adequate, the Fowler–Nordheim characterization of field emitters (in terms of A_{FN} and B_{FN}) will be in error, and further, that knowledge of the anode voltage needed for FEA characterization is important in order to avoid failure events associated with significant gate current. Second, because electron transport through the bulk of a wide-band-gap layer is sufficiently ballistic that current limitation from these structures may be due to transport across the metal wide-band-gap material interface, a simple model, based on generalizations of the FEA model, was used to calculate total current across such a surface. The parametric dependence of the tunneling current on various material and geometric parameters was analyzed to highlight non-Fowler–Nordheim-like behavior in the current–voltage relationships, and to show what typical material and geometric parameters are required to achieve IOA performance levels.

ACKNOWLEDGMENTS

The authors would like to thank: Dr. K. Nguyen (KN Research), Dr. M. Hollis, and Dr. R. A. Murphy (MIT-LL), M. Green (CPI), C. Spindt (SRI), and F. Charbonnier (LRI) for discussions regarding inductive output amplifiers, the IOA tube environment, arcing issues, and space charge. The authors gratefully acknowledge partial funding provided by the Office of Naval Research.

[1]See Vacuum Electronics Annual Review Abstracts (Palisades Institute for Research Services, Inc. 1996), which are the conference proceedings of the joint Monterey Power Tube Conference and the Vacuum Electronics Annual Review, Naval Postgraduate School, Monterey, CA, May 21–23, 1996.
[2]J. A. Arnaud, *Encyclopedia of Physics*, 2nd ed., edited by R. G. Lerner and G. L. Trigg (VCH Publishers, New York, 1991).
[3]R. H. Abrams and R. K. Parker, IEEE MTT-S Intl. Micr. Symp. Dig. **1**, 107 (1993).
[4]E. G. Zaidman and M. A. Kodis, IEEE Trans. Electron Devices **38**, 2221 (1991).
[5]J. P. Calame, H. F. Gray, and J. L. Shaw, J. Appl. Phys. **73**, 1485 (1993).
[6]M. A. Kodis, K. L. Jensen, E. G. Zaidman, B. Goplen, and D. N. Smithe, IEEE Trans. Plasma Sci. **24**, 970 (1996).
[7]K. L. Jensen, E. G. Zaidman, K. Nguyen, M. A. Kodis, and M. Garven, International Conference on Plasma Science (IEEE-ICOPS), San Diego, CA, May 19–22, 1997.
[8]K. L. Jensen, R. H. Abrams, and R. K. Parker, J. Vac. Sci. Technol. B **16**, 749 (1998).
[9]D. N. Smithe, B. Goplen, M. A. Kodis, and N. R. Vanderplaats, IEEE-ICOPS, Madison, WI, 1995.
[10]H. P. Freund and E. G. Zaidman, International Conference on Plasma Science (IEEE-ICOPS), San Diego, CA, May 19–22, 1997; E. G. Zaidman and H. P. Freund, International Conference on Plasma Science (IEEE-ICOPS), Boston, MA, June, 1997.
[11]K. T. Nguyen, H. P. Freund, E. G. Zaidman, N. Vanderplaats, and M. A. Kodis, Proceedings of the 1996 Microwave Power Tube Conference, May 1996.
[12]The CASSANDRA code has been validated by comparison to the 1-D TWT code CHRISTINE developed at NRL: T. Antonsen and B. Levush, "CHRISTINE: A Multifrequency Parametric Simulation Code for Traveling Wave Tube Amplifiers," NRL Technical Report NRL/FR/6840–97-9845, May 5, 1997.
[13]K. L. Jensen, J. Appl. Phys. **83**, 7982 (1998).
[14]E. E. Martin, F. M. Charbonnier, W. W. Dolan, W. P. Dyke, H. W. Pitman, and J. K. Trolan, Wright Air Development Division Technical Report 59–20, January 1960 (see also J. P. Barbour, W. W. Dolan, J. K. Trolan, E. E. Martin, and W. P. Dyke, Phys. Rev. **92**, 45 (1953)); W. P. Dyke, J. K. Trolan, E. E. Martin, and J. P. Barbour, Phys. Rev. **92**, 1043 (1953).
[15]H. G. Kosmahl, IEEE Trans. Electron Devices **38**, 1534 (1991).
[16]P. H. Cutler, J. He, N. M. Miskovsky, T. E. Sullivan, and B. Weiss, J. Vac. Sci. Technol. B **11**, 387 (1993).
[17]D. Nicolaescu, J. Vac. Sci. Technol. B **11**, 392 (1993).
[18]K. L. Jensen, E. G. Zaidman, M. A. Kodis, B. Goplen, and D. N. Smithe, J. Vac. Sci. Technol. B **14**, 1942 (1996); **14**, 1947 (1996).
[19]K. L. Jensen, P. Mukhopadhyay-Phillips, E. G. Zaidman, K. Nguyen, M. A. Kodis, L. Malsawma, and C. Hor, Appl. Surf. Sci. **111**, 204 (1997).
[20]K. L. Jensen, M. A. Kodis, R. A. Murphy, and E. G. Zaidman, J. Appl. Phys. **82**, 845 (1997).
[21]G. Arfken, *Mathematical Methods for Physicists*, 2nd ed. (Academic, New York, 1970). See in particular Chapter 12, Section 11 (Legendre Functions: Application to Spheroidal Coordinate Systems).
[22]A. Modinos, *Field, Thermionic, and Secondary Electron Spectroscopy* (Plenum, New York, 1984).
[23]Heinz K. Henisch, *International Series of Monographs on Physics No. 70* (Clarendon, Oxford, 1984).
[24]See *III-Nitride, SiC and Diamond Materials for Electron Devices*, Symposium Proceedings of the MRS No. 423, edited by D. K. Gaskill, C. D. Brandt, and R. J. Nemanich (1996).
[25]R. P. Feynman, *Frontiers in Physics* (Benjamin Cummings, Reading, MA, 1983), Vol. 36, and references therein.
[26]N. D. Lang and W. Kohn, Phys. Rev. B **1**, 4555 (1970).
[27]F. García-Moliner and F. Flores, *Introduction to the Theory of Solid Surfaces* (Cambridge University Press, Cambridge, 1979).
[28]M. van Schilfgaarde, A. Sher, and A.-B. Chen, J. Cryst. Growth **178**, 8 (1997).
[29]K. L. Jensen, E. G. Zaidman, M. A. Kodis, B. Goplen, and D. N. Smithe, J. Vac. Sci. Technol. B **14**, 1942 (1996); **14**, 1947 (1996); E. G. Zaidman, K. L. Jensen, and M. A. Kodis, *ibid.* **14**, 1994 (1996).
[30]Atomic units will be used throughout, and are given by electron volts (eV), femtoseconds (fs), angstroms (Å); the electron charge (e) is unity. Current density therefore has units of $e/(\text{fs Å}^2)$ and current of (e/fs). The permittivity of free space $\epsilon_0 = e^2/4\pi\alpha_{fs}\hbar c$, where $\alpha_{fs} = 1/137.04$ is the fine structure constant.
[31]W. D. Goodhue, P. M. Nitishin, C. T. Harris, C. O. Bozler, D. D. Rathman, G. D. Johnson, and M. A. Hollis, J. Vac. Sci. Technol. B **12**, 693 (1994).
[32]S. T. Purcell, V. T. Binh, and R. Baptist (preprint obtained from V. T. Binh).
[33]F. Charbonnier, J. Vac. Sci. Technol. B **16**, 880 (1998).
[34]C. K. Birdsall and W. B. Bridges, *Electron Dynamics of Diode Regions* (Academic, New York, 1966).
[35]H. Imura, S. Tsuida, M. Takahasi, A. Okamoto, H. Makishima, and S. Miyano, *Technical Digest of the IEEE-IEDM* (Washington, DC, 1997), p. 721.
[36]R. P. Joshi, K. H. Schoenbach, C. Molina, and W. W. Hofer, J. Appl. Phys. **74**, 1568 (1993).
[37]P. H. Cutler, Z.-H. Huang, N. M. Miskovsky, P. D'Ambrosio, and M. Chung, J. Vac. Sci. Technol. B **14**, 2020 (1996).
[38]M. W. Geis, J. C. Twichell, and T. M. Lysczarz, J. Vac. Sci. Technol. B **14**, 2060 (1996).
[39]P. Lerner, P. H. Cutler, and N. M. Miskovsky, J. Phys. IV **C5**, 39 (1996).
[40]L. Eyges, *The Classical Electromagnetic Field* (Dover, New York, 1972).
[41]K. L. Jensen, J. Vac. Sci. Technol. B **13**, 516 (1995).
[42]N. Pupeter, A. Göhl, T. Habermann, E. Mahner, G. Müller, H. Piel, Ph. Niedermann, and W. Hänni, J. Vac. Sci. Technol. B **14**, 2056 (1996).
[43]W. B. Choi, J. Liu, M. T. McClure, A. F. Myers, V. V. Zhirnov, J. J.

Cuomo, and J. J. Hren, J. Vac. Sci. Technol. B **14**, 2050 (1996).
[44]E. I. Givargizov, V. V. Zhirnov, A. V. Kuznetsov, and P. S. Plekhanov, J. Vac. Sci. Technol. B **14**, 2030 (1996).
[45]P. H. Cutler, Z.-H. Huang, N. M. Miskovsky, P. D'Ambrosio, and M. Chung, J. Vac. Sci. Technol. B **14**, 2020 (1996).
[46]W. Zhu, G. P. Kochanski, S. Jin, and L. Seibles, J. Vac. Sci. Technol. B **14**, 2011 (1996).
[47]W. P. Kang, J. I. Davidson, M. Howell, B. Bhuva, D. L. Kinser, D. V. Kerns, Q. Li, and J. F. Xu, J. Vac. Sci. Technol. B **14**, 2068 (1996).
[48]See, for example, the discussion in Ref. 23 (H. K. Henisch), p. 49.

Field emission from nitrogen doped tetrahedral amorphous carbon prepared by filtered cathodic vacuum arc technique

L. K. Cheah,[a] X. Shi, B. K. Tay, and Z. Sun
Ion Beam Processing Lab, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

(Received 18 August 1997; accepted 6 April 1998)

A study of field emission from nitrogen doped tetrahedral amorphous carbon (ta-C:N) films prepared by the filtered cathodic vacuum arc (FCVA) deposition technique is reported. Field emission from ta-C:N films deposited on Si substrates was investigated in a diode configuration at room temperature and base pressure of 2.0×10^{-6} Torr. The $J-E$ curves shift significantly towards the low potential side with increasing nitrogen concentration. The lowest field at which field emission was obtained was 10 V μm^{-1}. A current density of 0.1 mA mm^{-2} (assuming the entire film surface is emitting) at 50 V μm^{-1} was obtained from these films. Electronic parameters, i.e., the band gap energy and activation energy were measured in order to construct an energy band diagram for the heterojunction structure, and the field emission mechanism is proposed based on this structure. © *1998 American Vacuum Society.* [S0734-211X(98)00904-4]

I. INTRODUCTION

The filtered cathodic vacuum arc FCVA deposition technique was reported to be an efficient method of producing high quality, microparticle free ta-C at room temperature.[1] It has been shown using electron energy-loss spectroscopy (EELS) that as much as 87% of the carbon atoms in the ta-C films produced by the FCVA deposition technique form an amorphous tetrahedral (sp^3) structure.[1] The ta-C is p-type semiconducting material and has a band gap around 2.6 eV. These electronic properties were reported to be adjustable by incorporation of nitrogen during deposition.[2] Thus, ta-C:N films could be useful for electronic applications, in our case, field emission cathodes. The objective of this article is to describe the field emission properties of ta-C:N prepared by the FCVA deposition technique and to point out the effect of nitrogen flow rate on the emission properties by using a parallel plane diode configuration. The electrical properties such as band gap energy and activation energy were measured to construct a band structure diagram for the thin film and the emission mechanism is discussed.

II. EXPERIMENTAL DETAILS

The FCVA deposition technique has been discussed elsewhere.[1] Ionized nitrogen gas was introduced into the deposition chamber through a secondary ion beam source independent of the FCVA source. The base pressure was varied between 1.0×10^{-5} and 2.0×10^{-4} Torr which depended on the nitrogen flow rate which was monitored using a mass flow controller. The impinging C and N ion energies were fixed at 100 and 180 eV, respectively. The ta-C:N surface morphology was inspected using an atomic force microscope (AFM) in the tapping mode using a Dimension 3000 Scanning Probe Microscope by Digital Instruments. The optical band gap (Tauc band gap) is determined by $\varepsilon_2(E) = \text{const}[(E-E_g)^2/E^2]$.[3] ε_2 is the complex part of the dielectric constant and determined by a Jobin Yvon UVISEL phase modulated spectral ellipsometer. E is the photon energy, E_g is the Tauc band gap proposed by Tauc.[3] The film structure model used for the simulation in the spectroscopic ellipsometer is based on a four-layer model developed by Shi *et al.*[4] The electrical measurements were obtained by sputtering 100 nm thick titanium metal contacts to form gap cells (5 mm $\times$0.5 mm interelectrode) to the film deposited on quartz. The conductivity was measured using a Keithley 617 electrometer. To determine the activation energy of the film, the conductivity measurements were taken over the temperature range from 300 to 500 K. Samples of different nitrogen concentrations deposited on $\langle100\rangle$ p-type Si substrates with a resistivity of ~0.01 Ω cm were measured.

III. EXPERIMENTAL RESULTS

A. Some properties on ta-C:N films

The AFM surface morphology of Fig. 1 indicates that the surface of the ta-C:N films deposited on Si is very smooth and exhibits a root mean square (rms) surface roughness of about 0.3 nm, over an area of 1×1 μm^2. Few macroparticles and surface defects were observed under an optical microscope (50×). The results show that the roughness increases from 0.2 to 0.4 nm with increasing nitrogen flow rate from 2 to 10 sccm.

Figure 2(a) shows that the optical band gap decreases from 2.4 to 1.1 eV with increasing nitrogen flow rate from 2 to 10 sccm. The activation energy decreases from 1.08 to 0.07 eV when the nitrogen flow rate increases from 2 to 10 sccm [Fig. 2(b)]. These results show that the incorporation of nitrogen during deposition significantly shifts the Fermi level towards the conduction band.

B. Field emission measurement

Figure 3 shows the $I-V$ curves from the samples of different nitrogen concentrations. For all measurements plotted

[a]Electronic mail: pa1870668@ntu.edu.sg

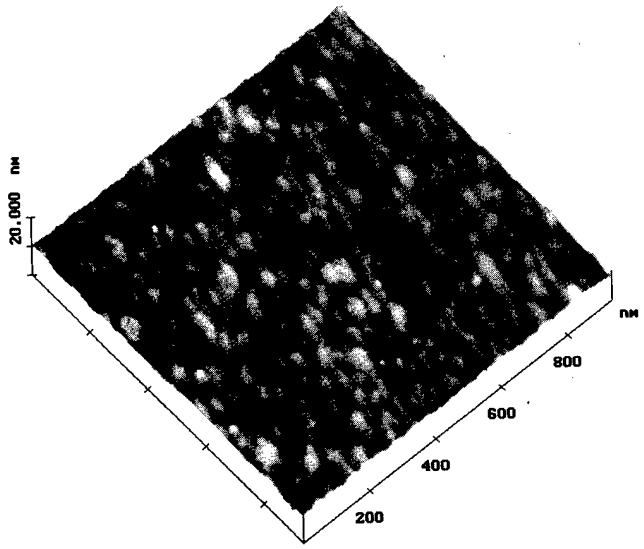


FIG. 1. AFM picture for nitrogen doped ta-C.

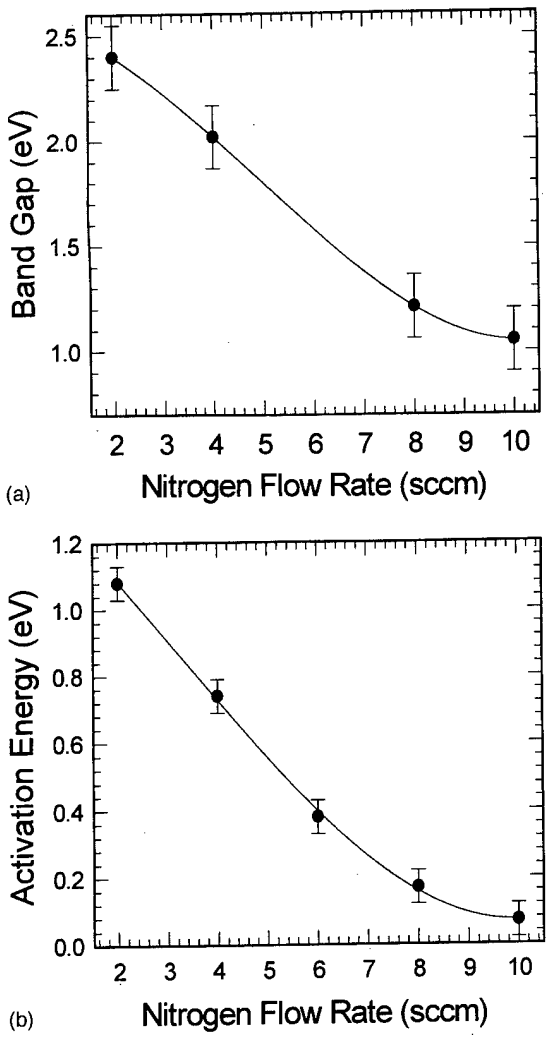


FIG. 2. (a) Optical band gap, and (b) activation energy as a function of nitrogen flow rate.

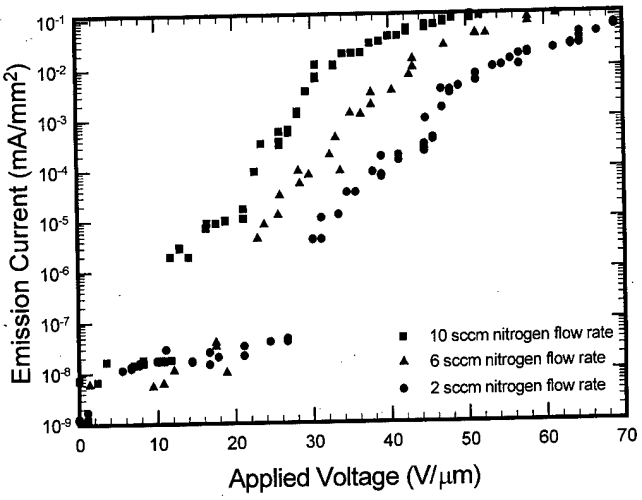


FIG. 3. $J-E$ curves for ta-C:N deposited on p^{++}-Si with various nitrogen flow rates.

in Fig. 3, no damage was observed on the sample surfaces under an optical microscope (50×). The highest onset field (~30 V μm^{-1}) was measured from weakly doped ta-C (2 sccm nitrogen flow rate). The lowest turn-on voltage (~10 V μm^{-1}) was obtained from the ta-C film with the highest nitrogen concentration (10 sccm nitrogen flow rate). Repeated emission current measurements exhibited current densities within 0.1 mA mm^{-2} of each other for all of the samples.

However, at some spots (mostly on the samples with low nitrogen flow rate), no emission current was observed as the voltage was increased. Then, at critical electric fields (normally around 55–60 V μm^{-1}), the current abruptly increased and was accompanied by an electric discharge, and subsequent ramping of the voltage yielded a much lower turn-on field. Similar results were reported by Talin and co-workers[5] for samples deposited by the laser ablation technique. After emission, examination of the film surface with a scanning electron microscope (SEM) shows crater formation. In some cases, protrusions were observed within the crater, which was due to melting of the cathode surface. The protrusions greatly modify the surface morphology and serve as field enhancement sites, which reduce the subsequent onset field.

IV. DISCUSSION

A. Energy band diagram construction

When two anisotype semiconductors with different band gaps are brought into contact, a heterojunction is formed.[6] The interface states of ta-C:N and Si have been ignored to a first approximation. The electron affinity of Si is taken to be 4 eV.[6] The band gap and activation energy of heavily doped p-type Si are taken to be 1.12 and 0.1 eV,[6] respectively. The doping concentration for p-type Si is in the range of $10^{18}-10^{19}$ cm^{-3}. The band gap energy and activation energy were determined from Fig. 2 for ta-C:N films grown at

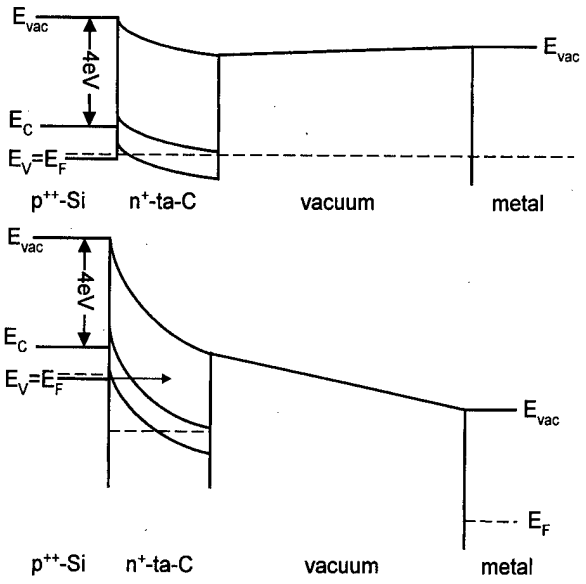


FIG. 4. Energy band diagram without the interfacial states for ta-C:N/p^{++}-Si heterojunction at (a) no applied field (b) an applied field between Si back contact and metal.

10 sccm nitrogen flow rate, the most heavily N doped ta-C, (n^+-ta-C), and the values are 1.0 and 0.07 eV, respectively. The electron affinity of n^+-ta-C is taken to be 2.5 eV.[7] The depletion region was calculated using a dielectric constant of 5–6 (determined from spectroscopic ellipsometer), built-in potential of 0.8 eV and assuming an ionized donor density of the order of 10^{16}–10^{17} cm^{-3}. A depletion region of a few tenths of nm is calculated based on the above parameters, and the band diagram is shown in Fig. 4.

B. Field emission mechanism

Related to the band diagram, under bias conditions, the fully depleted ta-C:N layer can be the origin of emission where the electrons that have accumulate at the Si near the space charge region enter into the conduction band of the ta-C:N. The consequence of this model is that the true cathode is the p^{++}-Si and the ta-C:N acts as a space charge region with a lower electron affinity to enhance electron emission. The proposed mechanism has been put forth by a number of workers. Bayliss and Latham[8] proposed that the injection of electrons into the conduction band of a dielectric occur by tunneling from the metal. Amaratunga and Silva[9] suggested this mechanism for a Si cathode deposited with DLC. With the semiconductor, the major difference is the ability to control the amount of band bending by varying the donor concentration and the width of the space charge region.

The lowest onset field from heavily doped ta-C:N deposited on p^{++}-Si (onset field of 10 V μm^{-1}) can be explained in terms of the band bending structure. The band bending is maximum at the p^{++}-Si/n^+-ta-C heterojunction structure, as shown in Fig. 4(b). The consequence of this structure is that the probability for tunneling electrons to enter the conduction band of the ta-C:N film increases.

V. CONCLUSION

The results imply that ta-C:N films are successful in producing the field emission with the sufficient current density of 0.1 mA mm^{-2}. The heavily doped heterojunction gives the lowest onset field and stable field emission current.

[1]X. Shi, B. K. Tay, H. S. Tan, L. Zhong, and Y. Q. Tu, J. Appl. Phys. **79**, 7234 (1996).
[2]V. S. Veerasamy, G. A. J. Amaratunga, C. A. Davis, A. E. Timbs, W. I. Milne, and D. R. McKenzie, J. Phys.: Condens. Matter **5**, 1169 (1993).
[3]J. Tauc and A. Menth, J. Non-Cryst. Solids **8–10**, 569 (1972).
[4]X. Shi, L. K. Cheah, and B. K. Tay, Thin Solid Films **312**, 166 (1998).
[5]A. A. Talin, T. E. Felter, T. A. Friedmann, J. P. Sullivan, and M. P. Siegal, J. Vac. Sci. Technol. A **14**, 2 (1996).
[6]V. S. Verasamy, G. A. J. Amaratunga, J. S. Park, H. S. MacKenzie, and W. I. Milne, IEEE Trans. Electron Devices **42**, 577 (1995).
[7]J. Robertson, Diamond Relat. Mater. **5**, 797 (1996).
[8]K. H. Bayliss and R. V. Latham, Proc. R. Soc. London, Ser. A **403**, 285 (1986).
[9]G. A. J. Amaratunga and S. R. P. Silva, Appl. Phys. Lett. **68**, 2529 (1996).

Field emission of nitrogen-doped diamond films

I. T. Han[a] and N. Lee
Electronic Materials Laboratory, Samsung Advanced Institute of Technology, P. O. Box 111, Suwon 440-600, Korea

S. W. Lee
Department of Physics, Myoung-gi University, Seoul 449-728, Korea

S. H. Kim
Department of Chemistry, Silla University, Pusan 617-736, Korea

D. Jeon
Department of Physics, Myoung-gi University, Seoul 449-728, Korea

(Received 30 September 1997; accepted 13 April 1998)

The authors investigated field emission characteristics of nitrogen-doped diamond films, which were grown using microwave plasma-enhanced chemical vapor deposition. Nitrogen-doped films showed low turn-on voltages below 2 V/μm. Secondary ion mass spectroscopy was used to compare nitrogen concentrations in the films. Morphologies, Raman spectra, the resistivities, and surface roughness of the films were changed as the nitrogen concentrations varied. The field emission properties of heavily nitrogen doped diamond films were related to the film resistivity, surface morphologies and Raman characteristics. © *1998 American Vacuum Society.*
[S0734-211X(98)04004-9]

I. INTRODUCTION

Diamond is a promising material for various electronic and mechanical applications. Among the potential applications of diamond films such as high temperature devices, power devices, radiation detectors, surface acoustic wave devices, optical windows, hard coatings, etc., a recent study on field emission is highlighted. The superiority of diamond for field emitters is based on low turn-on voltages, ease of fabrication, and inertness to sputtering.

Several mechanisms have been suggested to explain the excellent field emission characteristics of diamond films: negative electron affinity (NEA), graphitic carbon formation, impurity incorporation, and surface morphologies.[1–9] Nonetheless, field emission mechanisms of diamond films still remain to be understood. Field emission processes are such that electrons are injected into a film from an electrode, are then transferred to the film surface, and finally overcome a work function, escaping from the surface. An ohmic contact is to be maintained at the interface between the film and a substrate for an efficient supply of electrons. These electrons move through the film with sufficient conductivity which can be imparted with impurity incorporation and defect formation such as nondiamond carbon phases and crystalline defects. In addition, diamond films contain crystals with sharp corners or edges which could enhance electron emission.[10] As mentioned above, many experimental factors have effects on the field emission characteristics of diamond films, which can be controlled during film growth. These factors should be considered collectively at the same time rather than separately for tailoring for the optimal field emission. Different from most studies published to date which have examined these factors individually, however, this paper compared several experimental factors collectively to characterize their effects on field emission of diamond films, including surface roughness, resistivities, Raman spectra, and nitrogen concentrations. The transconductivities of chemical vapor deposited (CVD) diamond film increased with nitrogen incorporation, and the threshold voltage of field emission correlated with this. As the transconductivity increased, the threshold voltage decreased. This enhanced transconductivity seems to originate from nondiamond carbons which formed during film growth. However, surface roughness was not a critical factor for CVD diamond field emission.

II. EXPERIMENT

A low-power microwave plasma-enhanced chemical vapor deposition (MPECVD) system (ASTeX 1.5 kW) was used to deposit diamond films on (100)-oriented silicon wafers. Si wafers with low resistivities below 0.005 Ω cm were selected to enhance electron mobility through the substrates during field emission measurements. Prior to deposition, the substrates were agitated with 0.5 μm diamond powders in ethanol for 3 h, rinsed with pure acetone, ethanol, and de-

TABLE I. Deposition conditions of diamond films.

Sample no.	$N_2/CH_4/H_2$ (sccm)	Temperature (°C)[a]	Temperature (°C)[b]
1	0.0/ 6 /194	790	410
2	0.0/ 12 /188	790	410
3	1.5/ 6 /192	1070	900
4	3.0/ 6 /191	1070	900
5	6.0/ 6 /188	1070	900

[a]Temperature read by pyrometer.
[b]Temperature read by thermocouple.

[a]Electronic mail: ithan@saitgw.sait.samsung.co.kr

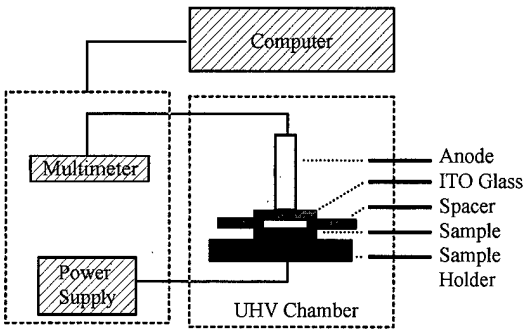


FIG. 1. Schematic of a field emission measurement system.

ionized water, and then blown with nitrogen gas. The substrate, loaded on a Mo holder, was heated with a rf heater. Following the evacuation of a reaction chamber up to 10^{-2} Torr using a mechanical pump, H_2 plasma pretreatment was carried out to clean the Si substrate with 800 W at 25 Torr for 5 min. The deposition conditions of diamond films are shown in Table I. Deposition temperatures were set to produce diamond films with high qualities under given conditions. Microwave power for deposition was maintained at 1000 W. Deposition for 18 h resulted in films as thick as 4 μm. Substrate temperatures during diamond deposition were monitored by optical pyrometry and by a thermocouple located just below the Mo substrate holder.

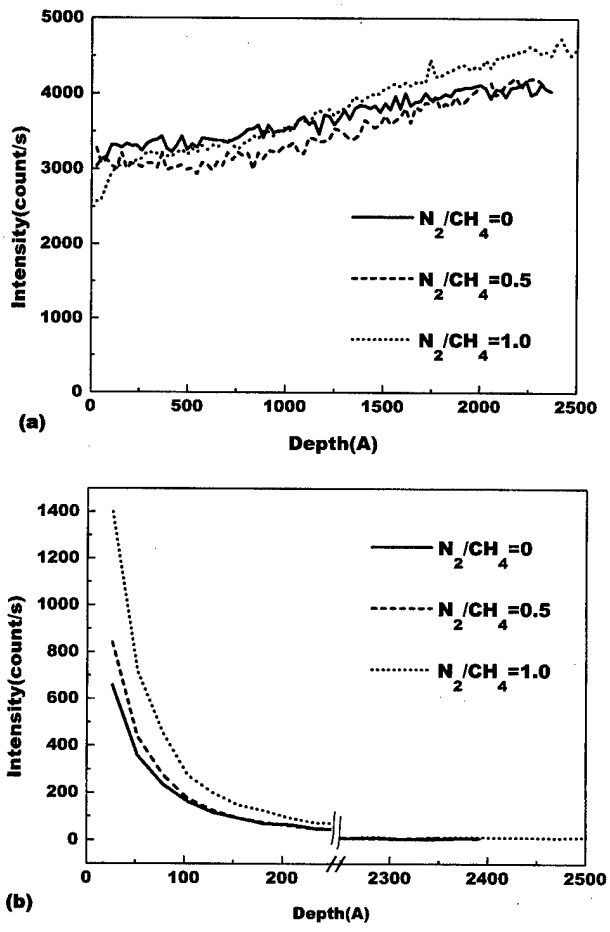


FIG. 2. Carbon and nitrogen concentrations of diamond films grown with a mixture gas of methane (3%) and nitrogen, measured by SIMS: (a) carbon depth profile, (b) nitrogen depth profile. Units are not absolute values.

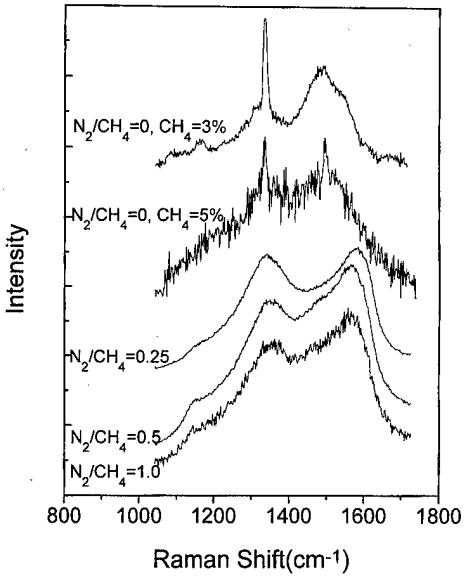


FIG. 3. Raman spectra of CVD diamond films grown with a mixture gas of methane (3%) and nitrogen: (a) $N_2/CH_4=0$, (b) $N_2/CH_4=0.25$, (c) $N_2/CH_4=0.5$, and (d) $N_2/CH_4=1.0$.

Diamond films were investigated using micro-Raman spectroscopy (Renishaw 3000) with a 514.5 nm wavelength laser, secondary ion mass spectroscopy (SIMS; FISON), and atomic force microscopy (AFM; PSI LS). Electrical transconductivities were measured for diamond films with Ti/Au electrodes of 100 μm diam and backcontacts of heavily N-doped Si substrates.

Field emission measurements were carried out under vacuum of 10^{-7}–10^{-8} Torr, as given in Fig. 1. An anode of indium tin oxide glass was biased with voltages between 0 and 5000 V using a dc power supply (Stanford Research Systems, P5350). The anode was connected to a Keythley 2000 multimeter to measure emission currents. All electronics were controlled with a computer interfaced through an IEEE 488.2 GPIB card. Voltages were ramped at 20 V/min. Currents were sampled ten times for every reading voltage, producing average values. Spacing between the sample and the anode was maintained as 150 μm with cover glasses.

III. RESULTS AND DISCUSSION

Figure 2 shows SIMS data for diamond films grown at different N_2 concentrations in the reaction gas mixture. While carbon concentrations in the films were maintained almost at the same level, nitrogen concentrations increased as more N_2 was added to the reaction gas mixtures. Nitrogen was, however, detected only near the surface regions, leading to the conclusion that it was concentrated within surface regions as deep as a few nanometers regardless of the nitrogen concentration. Based on the fact that the carbon concentrations revealed relatively small differences with depth, it appears that the variation of nitrogen concentrations with depth

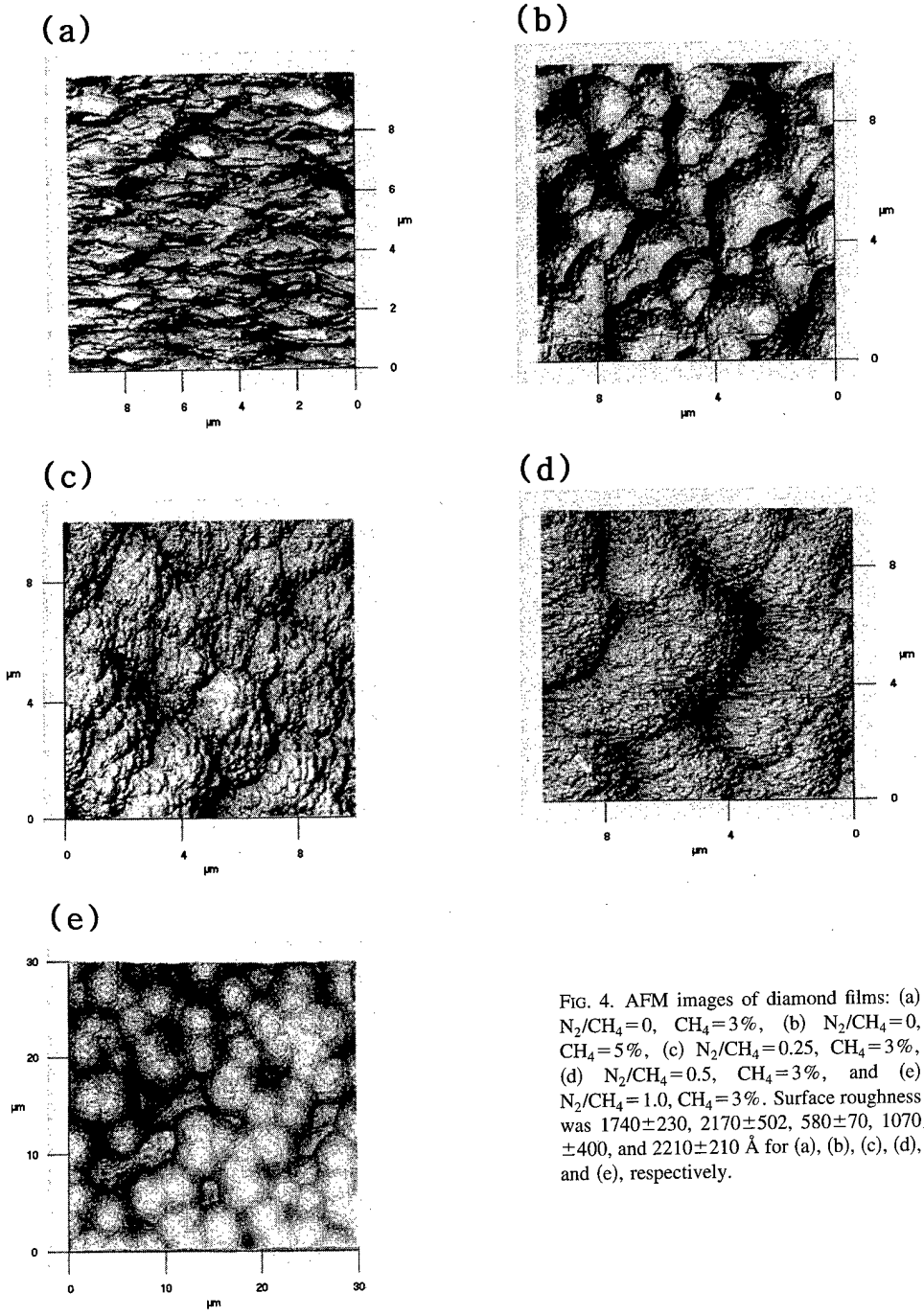


FIG. 4. AFM images of diamond films: (a) $N_2/CH_4=0$, $CH_4=3\%$, (b) $N_2/CH_4=0$, $CH_4=5\%$, (c) $N_2/CH_4=0.25$, $CH_4=3\%$, (d) $N_2/CH_4=0.5$, $CH_4=3\%$, and (e) $N_2/CH_4=1.0$, $CH_4=3\%$. Surface roughness was 1740 ± 230, 2170 ± 502, 580 ± 70, 1070 ± 400, and 2210 ± 210 Å for (a), (b), (c), (d), and (e), respectively.

apparently reflected real changes of nitrogen content of the films. This result seems to be similar to those of Li doping experiments reported by Nesladek *et al.*[11] They reported that *in situ* doped Li concentrations rapidly decreased from the surface at a high temperature. Since the deposition temperatures in this experiment were above 1000 °C, it is thought that the doping proceeded through a similar mechanism. Although nitrogen is incorporated into diamond films by introducing a nitrogen gas into the plasma during deposition, it is worth intensively studying how to achieve uniform doping concentrations of nitrogen in the depth direction of films.

The Raman spectra of diamond films are given in Fig. 3. The spectra are composed of two distinct peaks around 1332 and 1580 cm^{-1}, which are characteristic of diamond and graphitic carbon, respectively, in CVD-deposited diamond films. The film grown with the 3% methane concentration showed a full width at half maximum (FWHM) of 7.5 cm^{-1} for the 1332 cm^{-1} peak and an I_d/I_g ratio (the intensity ratio of the diamond peak to the graphitic carbon peak) of 1.46. As the methane concentration increased to 5%, on the other hand, the FWHM and the I_d/I_g ratio became 11.3 cm^{-1} and 1.0, respectively. With the variation of nitrogen concentration, the Raman spectra changed consistently like the SIMS data. For the N_2/CH_4 ratios of 0, 0.25, 0.5, 1.0 at 3% CH_4,

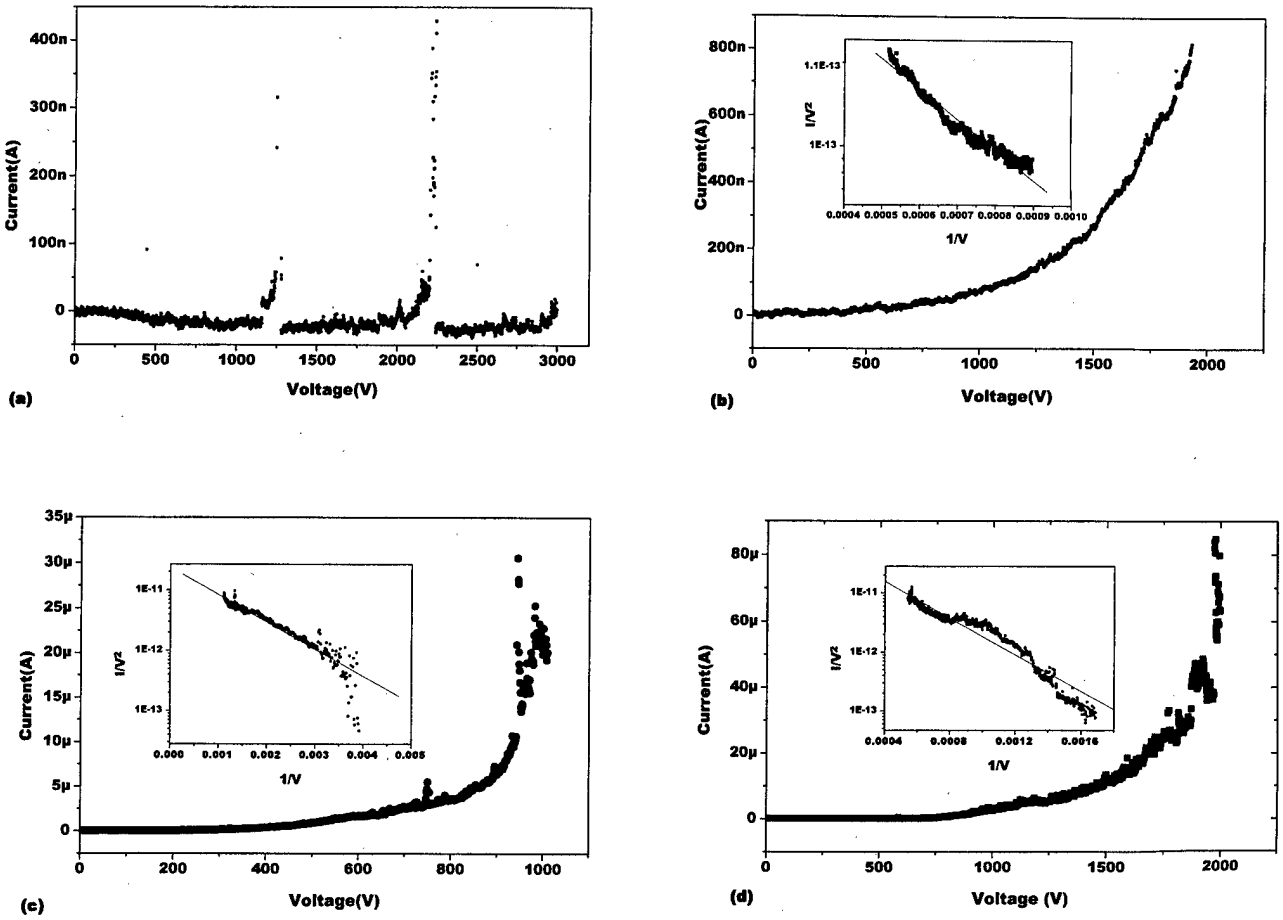


FIG. 5. *I-V* curves of field emission and their FN plots for diamond films: (a) $N_2/CH_4=0$, (b) $N_2/CH_4=0.25$, (c) $N_2/CH_4=0.5$, (d) $N_2/CH_4=1.0$. A methane concentration of the reaction gas mixture of all films was 3%. Insets represent FN plots of field emission.

the FWHM of the 1332 cm^{-1} peaks were measured to be 7.5, 55.5, 126.9, and 187.0 cm^{-1}, and the I_d/I_g ratios were 1.46, 0.93, 0.78, and 0.82, respectively. It is implied that the crystalline qualities of diamond films deteriorated with increasing amounts of nitrogen gas, as already reported by other studies.[12,13]

Figure 4 presents AFM images of as-grown diamond films. For a methane concentration of 3%, crystalline facets were observed to be preferentially (100) oriented. With increasing methane concentration or nitrogen concentration, crystalline facets disappeared, resulting in round-shaped growth features. The morphological changes seem to agree with the result of the Raman spectra that the crystalline qualities of films deteriorated at higher nitrogen concentrations. Surface roughness was averaged for five different measurements with the 80×80 μm^2 scans using AFM. The films grown at methane concentrations of 3% and 5% showed a surface roughness of 1740 ± 230 and 2170 ± 502 Å, respectively. As the N_2/CH_4 ratios at 3% CH_4 increased to 0.25, 0.5, and 1.0, the surface roughness became 580 ± 70, 1070 ± 400, and 2210 ± 210 Å, respectively. The higher the nitrogen concentrations were, the rougher the film surfaces were.

Transconductivities of diamond films deposited on heavily doped *n*-type Si substrates were measured with a MIM (metal-insulator-metal) capacitor structure. The Si substrate was used as a backcontact, and a 0.1-mm-wide, circular Ti/Au electrode was made on CVD diamond films using dc sputtering. Transconductivities were calculated with applied 10 V dc and saturated currents. Calculated resistivities were 6280, 3458, 1725, and 1909 Ω cm for N_2/CH_4 ratios of 0, 0.25, 0.5, and 1.0 at the 3% methane concentration, respectively.

Field emission characteristics were measured for the 3% CH_4 films with varying N_2/CH_4 ratios which revealed different diamond qualities, N concentrations, roughness, and resistivities. Figure 5 presents *I-V* curves and their Fowler–Nordheim (FN) plots. While electron emission did not occur for nitrogen-free films except for a few arcings, all nitrogen incorporated films showed *I-V* behaviors typical of field emission. Their threshold voltages of electron emission (V_{th}) were 7.5, 1.8, and 4.6 V/μm for the films with N_2/CH_4 ratios of 0.25, 0.5, and 1.0, respectively, at 3% CH_4. V_{th} was obtained at the point where the current density was 4×10^{-8} A/cm^2. The good linearity of the FN plots indicated that electrons were emitted through a field emission mechanism. To understand the effect of nitrogen content on field emission characteristics, the threshold voltages were compared

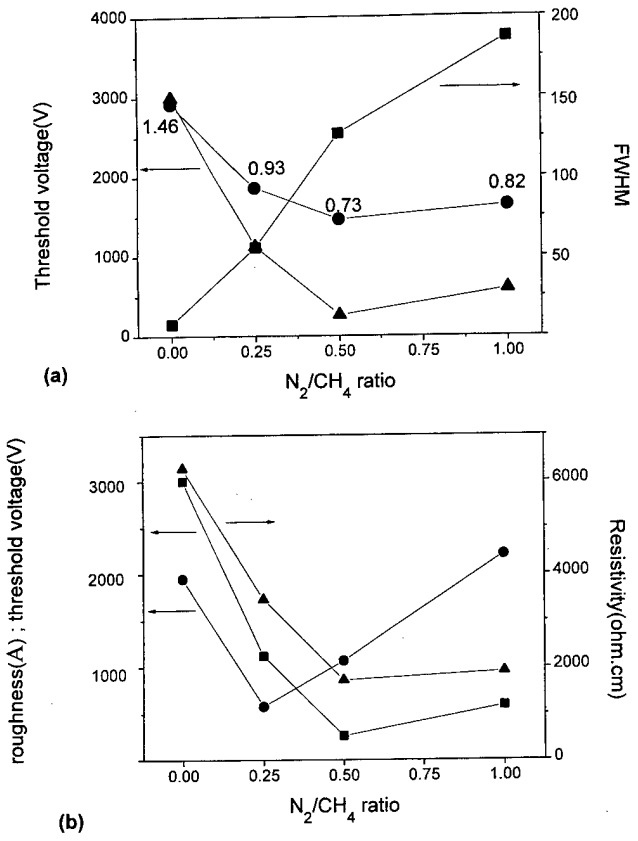


FIG. 6. Variations of threshold voltages, FWHM, I_d/I_g, surface roughness, and resistivities of diamond films with the N_2/CH_4 ratios: in (a), squares, circles, and triangles represent FWHM, I_d/I_g, and threshold voltages (V_{th}), and in (b), squares, circles, and triangles represent threshold voltages (V_{th}), surface roughness, and resistivities.

with FWHM and I_d/I_g in the Raman spectra, surface roughness and resistivities, as shown in Fig. 6. Threshold voltages drastically decreased as the diamond qualities deteriorated (FWHM and I_d/I_g) and with the resistivities of the films reduced. On the other hand, a relationship could not be established between the threshold voltages and the surface roughness. Although the film with no nitrogen addition had a very rough surface with sharp edges, it did not show any electron emission up to 20 V/μm. This result indicates that surface morphology is not crucial for the field emission of diamond films when the anode–cathode distance is much bigger than the surface roughness. Conclusively, film qualities and resistivities seem to play a crucial role in determining the field emission characteristics. However, film qualities and resistivities are considered to be correlated with each other rather than to be separated in their effects on the field emission characteristics. It is thought, therefore, that higher nitrogen concentrations resulted in the deterioration of film qualities due to increasing amounts of graphitic carbon, which led to the lower resistivities of the films. It seems that there is an optimal N_2/CH_4 ratio for field emission of diamond films. In this experiment, this optimal ratio appears to be 0.5 at 3% CH_4.

IV. CONCLUSION

We investigated the effect of nitrogen concentrations on field emission characteristics of MPECVD diamond films. Raman spectra revealed the deterioration of film qualities as the nitrogen concentrations in the reaction gas mixture increased, resulting in the lower resistivities of the films. Field emission measurements indicated that threshold voltages apparently became lower as nitrogen concentration increased, while films without nitrogen did not show electron emission. Field emission characteristics were compared with surface roughness, crystalline qualities, and transconductivities of diamond films. Whereas there is not an apparent relationship between field emission and surface roughness when the anode and cathode distance is sufficiently larger than the surface roughness, the threshold voltages of field emission decreased with lower crystalline qualities and resistivities of diamond films. It is considered, therefore, that the addition of nitrogen to the reaction gas mixture led to the formation of nondiamond carbon and consequently to higher conductivities of diamond films, enhancing field emission characteristics. It seems that the optimal N_2/CH_4 ratio of the reaction gas mixture at 3% CH_4 for the field emission of diamond films is 0.5.

[1]R. B. Jackman, B. Baral, C. R. Kingsley, and J. S. Foord, Diamond Relat. Mater. **5**, 378 (1996).
[2]D. Hong and M. Aslam, J. Vac. Sci. Technol. B **13**, 427 (1995).
[3]M. W. Geis, J. C. Twichell, N. N. Efremow, K. Krohn, and T. M. Lyszczarz, Appl. Phys. Lett. **68**, 2294 (1996).
[4]K. Okano, S. Koizumi, S. R. P. Siva, and G. A. J. Amaratunga, Nature (London) **381**, 140 (1996).
[5]W. Zhu, G. P. Kochanski, S. Jin, and L. Seibles, J. Appl. Phys. **78**, 2707 (1995).
[6]P. K. Baumann and R. J. Nemanich, Diamond Relat. Mater. **4**, 802 (1995).
[7]M. W. Geis, J. C. Twichell, J. Macaulay, and K. Okano, Appl. Phys. Lett. **67**, 1328 (1993).
[8]R. J. Nemanich, P. K. Baumann, M. C. Benjamin, S. W. King, J. van der Weide, and R. F. Davis, Diamond Relat. Mater. **5**, 790 (1996).
[9]J. Robertson, Diamond Relat. Mater. **5**, 797 (1996).
[10]K.-R. Lee, K. Y. Eun, S. Lee, and D.-R. Jeon, Thin Solid Films **290–291**, 171 (1996).
[11]M. Nesladek, K. Meykens, L. M. Stals, C. Quaeyhaegens, M. D'lieslaeger, T. D. Wu, M. Vanecek, and J. Rosa, Diamond Relat. Mater. **5**, 1006 (1996).
[12]H. Spicka, M. Grisser, H. Hutter, M. Grasserbauer, S. Bohr, R. Haubner, and B. Lux, Diamond Relat. Mater. **5**, 383 (1996).
[13]E. Rohrer, C. F. O. Graeff, R. Janssen, C. E. Nebel, and M. Stutzman, Phys. Rev. B **54**, 7874 (1996).

Field emission from ZrC and ZrC films on Mo field emitters

W. A. Mackie,[a] Tianbao Xie, M. R. Matthews, B. P. Routh, Jr., and P. R. Davis
Linfield Research Institute, McMinnville, Oregon 97128-6894

(Received 31 October 1997; accepted 6 April 1998)

We have been working for several years on field emission measurements from transition metal carbides. These studies have covered emission primarily from HfC and ZrC in the form of solid carbide emitters, thin film carbide overcoatings on single tip field emitters, and film coatings on field emitter arrays (FEAs). These carbide materials have work functions approximately 1 eV lower than Mo or Si, making them attractive candidates for low voltage microelectronic field emitter arrays. Uses for arrays or single emitters of these carbides include video displays, microwave applications, high current or small spot electron sources, and cold cathodes for operation in poor vacuums. Since Mo is one common FEA emitter material, we used it for an emitter substrate in this study. Our aim is to determine improvements in $I(V)$ characteristics, emission stability, capability of emission at high currents and in poor vacuum or tube conditions, and the ability to withstand exposure to atmospheric pressure without degradation. Individually, fabricated Mo field emitters were dosed via plasma vapor deposition from a high-purity ZrC_x source. The deposited film used for high currents were generally not heated but tested as deposited by field emission microscopy examination and determination of $I(V)$ characteristics. The results of these experiments indicate that work function reductions of the order of 1 eV can be achieved along with the ability to emit at higher current levels. Stable high current emission has also been obtained in the 0.5 mA range with ZrC films on Mo. This represents a tenfold increase in the obtainable emission levels compared to emission from a clean Mo substrate. These data are also compared to field emission data from bulk ZrC emitters. Some observations of high current emission in μTorr vacuum levels are also presented. The mechanisms for these improvements are discussed along with the possible role of stoichiometry.
© *1998 American Vacuum Society.* [S0734-211X(98)03904-3]

I. INTRODUCTION

High current density field emission from single crystal ZrC and HfC cathodes has been documented.[1–3] Work functions of these carbides are approximately 1 eV lower than Mo or Si which is attractive for field emitter array (FEA) applications. Beam confinement and enhanced stability has been observed from bulk carbide cathodes and from film coated cathodes.[4–8] For array applications, it is necessary to fabricate many identical field emitters. This can be done through the deposition of bulk carbide cones or through a carbide film overcoating on an existing array of Mo or Si emitters, areas we are currently investigating. The challenge in either approach is in achieving the desired properties (stability and low work function) in the deposited layer. We report here on experiments with the deposition of ZrC onto individual prefabricated Mo emitters and on $I(V)$ characterization of the resulting cathodes in ultrahigh vacuum (UHV) as well as in typical tube pressures. We also contrast data from ZrC films on Mo with field emission directly from single, bulk ZrC emitters in similar environments.

These carbides are unique in several respects. Along with their relatively low work functions, they have some of the highest melting points of any material. These properties seem to translate into robust field emitters, capable of withstanding high fields and/or high temperatures without field build-up or atom migration due to surface tension. Another unique aspect is their stoichiometry or carbon atom-to-metal atom ratio. While this family of transition metal carbides has a NaCl-type cubic crystal structure, they exist over a large range of stoichiometries. As we have seen and report here, high stoichiometry (near C/Zr=1.0) seems to yield increased resistance to surface contamination.

II. EXPERIMENT

An e-beam heated, high-purity ZrC evaporation source was used to deposit the film layer. The ZrC target used for plasma vapor deposition (PVD) was crystalline material prepared by arc float zone refinement.[9,10] These same materials were used for making the bulk carbide emitters also used in this study. Stoichiometries were checked by chemical analysis done in another laboratory. Most of the work reported here on bulk carbide emitters and evaporative targets was from substoichiometric ZrC_x (with $x \approx 0.92$), although a few bulk emitters were examined with lower stoichiometry (i.e., $x \approx 0.84$). We are currently working with bulk cathodes and evaporative source material which is approximately stoichiometric (i.e., $x \approx 1$). In this study, we also used Auger spectroscopy to ensure bulk and film purity and determine minimum film thickness. This calibration of film thickness was done by keeping the deposition parameters fixed and varying the deposition time. When the low energy Auger signal of the substrate fell below 1%, we assumed the coverage was equal to approximately three monolayers. These results were used as a guide for later ZrC deposition onto various planar

[a]Electronic mail: billm@linfield.edu

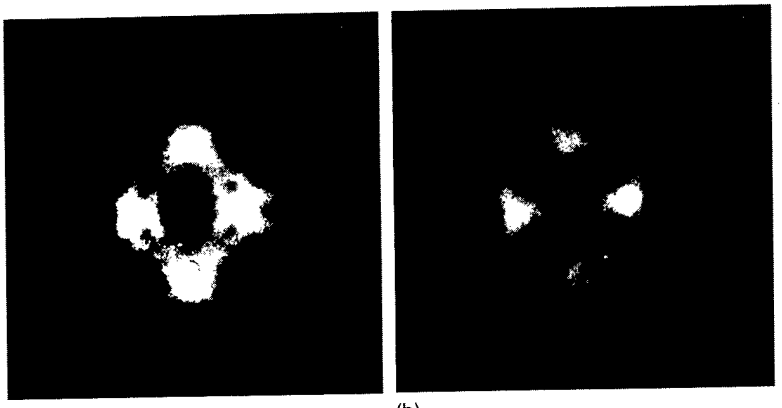


FIG. 1. Emission patterns from clean Mo and ZrC coated Mo field emitter. The Mo emitter has a (110) crystallographic orientation. Note the high emission areas on the (100) planes after ZrC deposition.

and field emitter substrates. For ZrC coatings on Mo field emitter tips, we assumed a thickness of ~5 nm based on the Auger study. After coating, we cannot see any tip geometry change, at least through scanning electron microscopy (SEM) analysis.

To summarize the data collection and emitter dosing process, we first configured a UHV system with a ZrC evaporative source, a field emitter holder, and a phosphor screen to view emission and obtain the necessary $I(V)$ data. The single emitters used in this study were individually fabricated via electrochemical etching of centerless ground single crystal rods of ZrC or polycrystalline drawn Mo wire. Generally, drawn Mo wire has a natural (110) orientation. The Mo emitters were etched in a phosphoric acid solution while those of ZrC were etched in a 10% perchloric/90% acetic acid solution.

Each single emitter was thermally cleaned, tested, and $I(V)$ data compiled. Cleaning was accomplished by heating up to 1900 and 2100 K for Mo and ZrC, respectively. Field emission microscopy (FEM) was used to verify the crystal orientation and condition of the emitter apex. The cleaning temperature for Mo was chosen so the emitters could be cleaned without being blunted too much due to the effects of surface tension. It is interesting to note that we have never detected similar surface tension blunting in any solid carbide field emitter. After a clean Mo tip was obtained and $I(V)$ data taken, ZrC film deposition followed immediately. Usually, the deposited ZrC film was several (5–20) monolayers thick. The film was either used as is or subjected to a variety of heating treatments. After deposition or treatment, FEM was used to examine the tip and $I(V)$ data were obtained.

For the high current experiments, currents were generally increased in 20 μA steps. $I(V)$ and $I(t)$ data were taken over typical times of 10–20 min. Both the short term noise and any current drop with time were noted. This process was usually continued until emitter failure. However, some emitters were held at 100 or 500 μA levels for additional time to better measure stability.

The tube pressure or poor vacuum experiments were made in a similar manner. Pressures were increased at decade intervals and currents increased to 100 μA levels before taking $I(V)$ and $I(t)$ data. When pressures reached 10^{-6}–10^{-5} Torr levels, emitting currents were increased in 50 μA intervals until failure.

III. RESULTS AND DISCUSSION

Many emitters, both bulk, single crystal ZrC tips and ZrC films on Mo tips, were prepared in this study and have been tested. The deposited film cathodes were subjected to heat

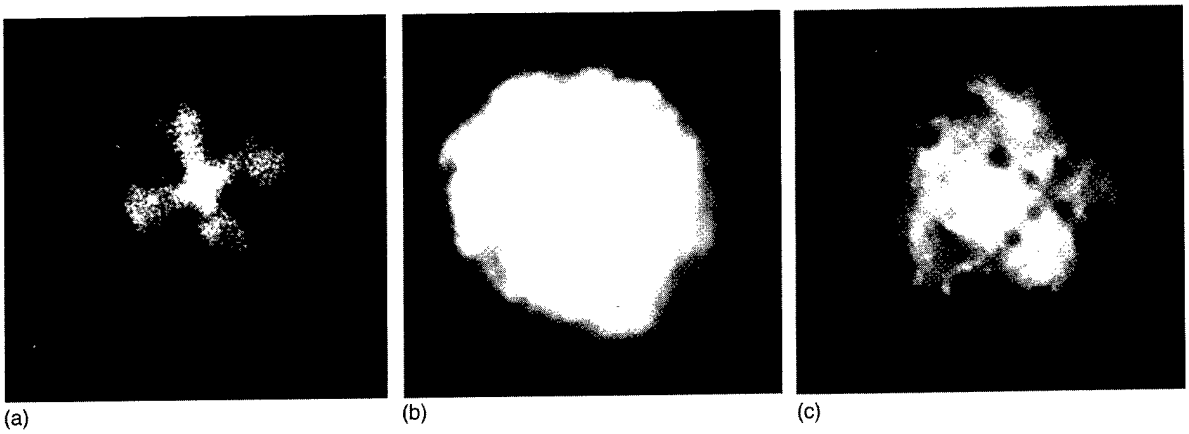


FIG. 2. $I(V)$ data from a crystalline Mo, single field emitter. The two curves are for clean and after deposition of ZrC.

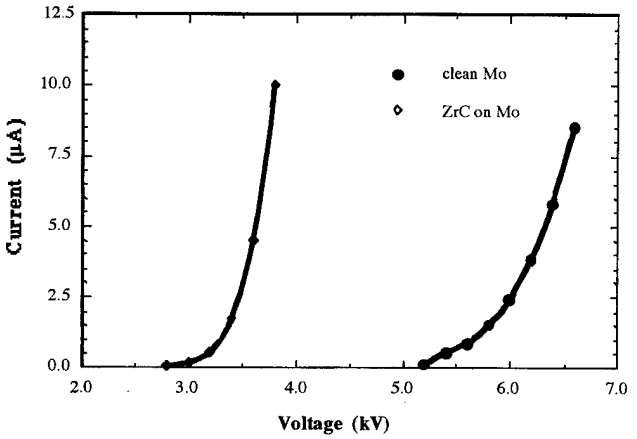


FIG. 3. Fowler–Nordheim curves from a single molybdenum field emitter. The two curves are for the same Mo emitter, clean and after deposition of a ZrC film, respectively.

treatments, each of which was followed by FEM examination and determination of $I(V)$ characteristics. It was found that 500–700 °C was adequate to achieve stable emission and a smooth, symmetrical FEM pattern.

Typical emission patterns obtained from a clean, single Mo emitter and after deposition of ZrC onto that same Mo emitter is shown in Fig. 1. Due to the nature of the wire used for the Mo emitters, we get a clean (110)-oriented pattern. After coating with ZrC the emission pattern changes with emission confined mainly to the (111) and (100) crystallographic planes of the underlying Mo substrate. For ZrC coating on Mo, significant decreases in operating voltage necessary to achieve a given emission current were observed. Figure 2 typifies $I(V)$ data from film-coated emitters. This marked lowering of the turn-on voltage is naturally accompanied by an increase in dI/dV or slope at turn-on. Similar results have been obtained with HfC as well on emitter substrates of both Mo and W, and additionally, we have previously demonstrated similar results on FEAs.

Base pressures during the UHV electron emission measurements were generally in the 10^{-10} Torr range. Fowler–Nordheim data were taken for each emitter, and the F–N slopes were calculated, as is shown in Fig. 3 for ZrC on Mo. By taking the ratio of ZrC film F–N slope to clean Mo F–N slope then taking this value to the 2/3 power, we arrive at the work function ratio. These data are summarized in Table I where the lowered work function assumes a constant β or tip shape. This method was used for all the data shown in Table I, with literature values assumed for the clean metal surfaces.

Figure 4 shows high current F–N data from the same emitter as shown in Fig. 3, overlaid on the data of Fig. 3. The dramatic change in slope of the F–N curve at high current (450 μA) is evident. This shift apparently is a result of both a lower value of the exponential factor in the F–N equation and a change in the F–N intercept as indicated by the analysis below.

In order to understand what is changing in the improved emission characteristics of ZrC coated Mo emitters, relative to clean Mo emitters, we use an analytical technique where

TABLE I. Results of ZrC and HfC film deposition on single emitters.

Film type and substrate	Voltage reduction at same current	Work function ratio from F–N	Work function clean (eV)	Work function with film (eV)
ZrC/W	38%	0.78	4.52	3.54
ZrC/Mo	56%	0.64	4.60	2.95
ZrC/Si	27%	0.87	4.82	4.19
ZrC/Mo FEA	44%	0.77	4.60	3.58
ZrC/Si FEA	~23%	~0.65	4.82	~3.15
HfC/W	60%	0.65	4.52	2.94
HfC/Mo	53%	0.65	4.60	2.99
HfC/Mo FEA	34%	0.88	4.60	4.05

the method and functional relationships have been discussed in detail elsewhere.[11,12] This method compares the slope and intercept values from F–N data superimposed on non-orthogonal families of curves which represent constant work function values and constant tip radius values, respectively. The F–N intercept depends strongly on both the field enhancement factor, β, and the emitting area of the tip. For smooth emitters, both of these parameters are related to the tip radius. However, emitters with rough surfaces, protrusions, etc., may have field enhancement factors which depend both on the gross emitter radius and on the local radii of curvature of protrusions. Furthermore, the effective emitting area represent only the sharpest protrusions (with the highest local fields) and not correspond at all to the gross emitter shape. Therefore, this slope-intercept analysis technique has meaning only for smooth emitters.

Figure 5 shows emission data plotted in this form. It *appears* that deposition of ZrC onto Mo, for coating thicknesses we have used in this study, initially results in a large reduction of work function with a corresponding large increase in tip radius. However, emission from the deposited film was measured without heating in this case. Thus, the surface had a rough geometry, probably with protrusions, producing an emission "pattern" of only a few bright spots, as observed in the FEM. Therefore, it is not possible to de-

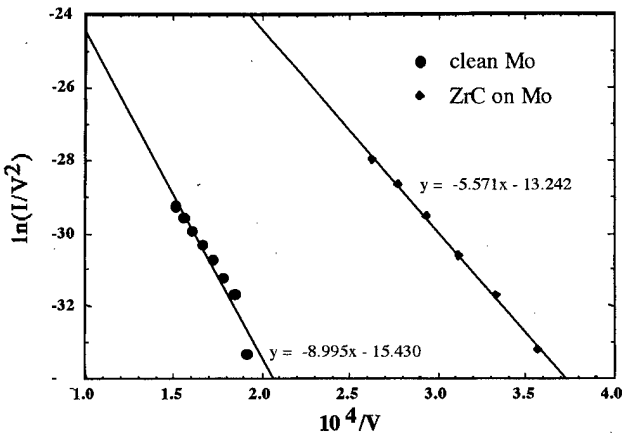


FIG. 4. Fowler–Nordheim curves from Mo emitter shown in Fig. 3 but with an additional curve from $I(V)$ data taken at high currents up to 450 μA.

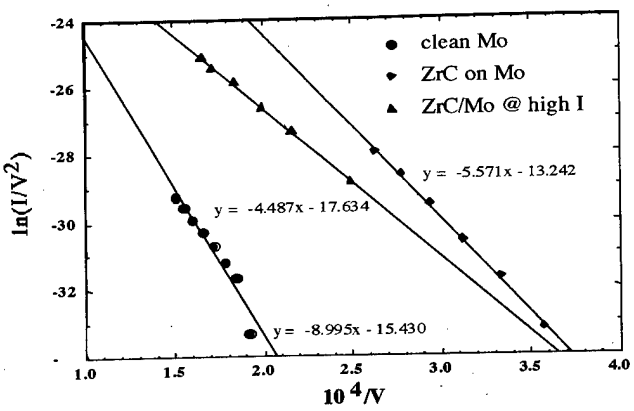


FIG. 5. Plotting Fowler–Nordheim slopes vs intercepts generates families of curves for constant ϕ, dashed lines, and constant β, solid lines. (The dashed lines of constant work function are lower to the right and the solid lines for constant β are lower to the left.) Superimposed on this are data from a smooth Mo emitter before and just after deposition of a ZrC thin film (without heating the substrate), and while operating at high current levels (450 μA) all shown as solid circles. (The absolute values of ϕ and β are not as relevant as the indicated trends.) Note, comparing the clean Mo point to the point operating at high currents indicate a work function lowering and a virtually unchanged tip radius.

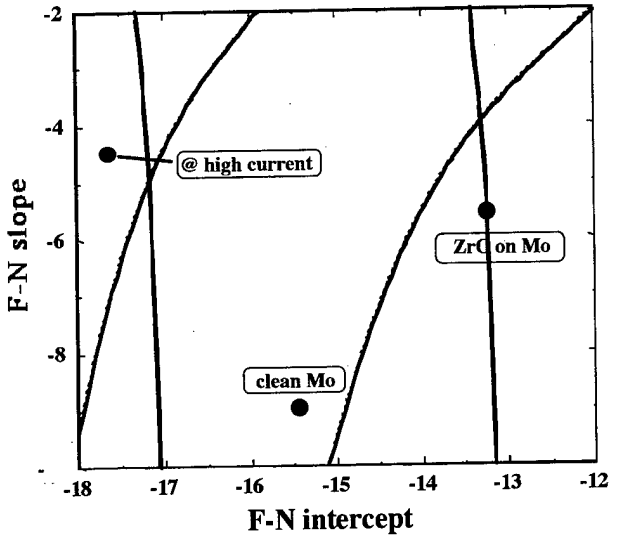


FIG. 6. Emission patterns from clean $ZrC_{0.92}$ field emitter and the same emitter running at high currents (170 μA) and in poor vacuum conditions (10^{-6} Torr air). The $ZrC_{0.92}$ emitter has a (100) crystallographic orientation. The third picture is from clean ZrC_1.

termine what has changed in this case. In Fig. 5, we have also plotted the data point corresponding to the high current emission shown in Fig. 4. From the general uniformity of the (very bright) FEM pattern observed for this high current situation, we infer that the emitting surface has been smoothed, probably by a combination of local emission heating of protrusions, field migration, and possibly localized field desorption. Therefore, we believe it is appropriate to apply the slope-intercept analysis technique to compare this coated, high current data point with the clean surface data point. From its position on the graph it seems that the work function has decreased without an appreciable change in tip radius or β factor from the clean Mo values.

Figure 6 shows emission patterns from solid ZrC emitters. The clean pattern is from an emitter with a (100) crystallographic orientation and is typical of carbides with stoichiometries somewhat less than 1. The four arms of emission extend from the (100) center toward the (110) planes with emission coming from the (210) and (310) areas. At high current and in poor vacuum, the emission pattern becomes unidentifiable due to the heavy ion bombardment. Even with such adverse conditions the emitter functions and in this case, when the pressure was again reduced and the tip heated to clean, the same $I(V)$ characteristics were obtained as before the operation at high current and pressure. The final photo in Fig. 6 shows a pattern obtained from a stoichiometric ZrC emitter. The orientation appears to be approximately (111). This surface-wide emission pattern has not been seen with substoichiometric emitters. Work is continuing with this type of emitter material.

We have investigated high current emission from a number of carbide emitters, both ZrC films on Mo emitters

TABLE II. High current field emission from individual emitters of Mo and ZrC film-coated Mo emitters. Assumptions: ϕ=constant (4.6 eV for Mo and 3.6 eV for ZrC) $\beta=2/[r\ \ln(4D/r)]$, $D=3$ cm, $a=(\pi r^2)/2$, where a is the emitting area, D is the tip to screen distance, and r is the tip radius.

Tip number	Vacuum condition[a]	Current (μA)	Voltage (kV)	Calculated r (nm)	Calculated J (A/cm^2)
Mo - 1	UHV	45	6.2	190	7.96×10^4
Mo - 2	UHV	145	7.6	192	2.51×10^5
$ZrC_{0.92}$/Mo-1	UHV	550	5.5	257	5.31×10^5
$ZrC_{0.92}$/Mo-2	UHV	150	2.9	105	8.66×10^5
$ZrC_{0.92}$/Mo-3	UHV	60	3.9	166	1.39×10^5
$ZrC_{0.92}$/Mo-4	UHV	650	6.8	272	5.59×10^5
$ZrC_{0.92}$/Mo-5	UHV	95	2.1	73	1.15×10^6
$ZrC_{0.92}$/Mo-6	UHV	400	5.3	208	5.91×10^5
$ZrC_{0.92}$/Mo-7	UHV	100	4.83	208	1.47×10^5
$ZrC_{0.92}$/Mo-8	UHV	100	2.4	85.7	8.67×10^5
$ZrC_{\approx1}$/Mo-9	2×10^{-9}	850	19	987	5.56×10^4
$ZrC_{\approx1}$/Mo-10	5×10^{-7} Torr air	100	2.5	90.3	7.81×10^5

[a]UHV is better than 1×10^{-9} Torr.

TABLE III. High current field emission from individual ZrC emitters. Assumptions: ϕ=constant (3.6 eV for ZrC). $\beta=2/[r\ \ln(4D/r)]$, D=3 cm, $a=(\pi r^2)/2$, where a is the emitting area, D is the tip to screen distance, and r is the tip radius.

Tip number	Vacuum condition[a]	Current (μA)	Voltage (kV)	Calculated r (nm)	Calculated J (A/cm^2)
$ZrC_{0.92}$-1	UHV	250	4.55	178	5.01×10^5
$ZrC_{0.92}$-2	UHV	580	6.1	240	6.40×10^5
$ZrC_{0.92}$-3	4×10^{-5} Torr Ar	175	6.2	272	1.50×10^5
$ZrC_{0.92}$-4	4×10^{-6} Torr air	130	6.4	290	9.82×10^4
$ZrC_{0.92}$-5	3×10^{-5} Torr air	300	8.9	411	1.13×10^5
$ZrC_{0.84}$-6	2×10^{-9} Torr	1290	6.5	242	1.39×10^6
$ZrC_{\approx1}$-7	2×10^{-9} Torr	220	8.4	392	9.07×10^4
$ZrC_{\approx1}$-8	UHV	4000	6.7	224	5.09×10^6
$ZrC_{\approx1}$-9	UHV	1500	8.8	353	7.49×10^5

[a]UHV is better than 1×10^{-9} Torr.

(Table II) and individual solid ZrC emitters (Table III). Currents from several hundred μA to a few mA from individual emitters have been observed in several cases. We have yet to run these tips for long periods of time and to conduct a quantitative study of emission noise. The time of tip operation represented in these two tables is generally 10–20 min at each of the current levels. As can be seen from these tables, a marked improvement is attainable with ZrC films on Mo over those of uncoated Mo. Note that, while most of the measurements were made under ultrahigh vacuum conditions, a few measurements were made under controlled pressures of gases, including Ar and air. These results and their interpretations are discussed later in this article.

IV. CONCLUSIONS

The increased emission at lower turn-on voltages is brought about by a work function lowering and not by a change in tip radius, or β factor, for the case of ZrC on Mo. This is demonstrated by viewing the plot of the Fowler–Nordheim slope versus intercept. It is not possible to determine accurate tip parameters from Fowler–Nordheim data due to multiple unknowns. However, we have observed that a plot such as is shown in Fig. 5 often suggests which of the variables is responsible for emission changes. From this information, we can go back to the Fowler–Nordheim slope data and find that the ratio of the slope while coated and running at high current to the slope of the clean Mo is 0.62. Taking this value to the power 2/3 we calculate a work function ratio of 0.72, assuming a clean work function of 4.6 eV for the molybdenum substrate, we obtain a coated work function value of 3.3 eV. This conclusion relies on the assumption that β has the same functional dependence on tip radius, that is, that the tip shape is congruent in all cases.

Our studies also suggest improvement in emission stability which we are currently investigating. Both the level of noise, so far measured qualitatively, and decrease in total emission current with time are significantly reduced in emission from ZrC_x when $x\approx1$ compared with $x\approx0.92$, and the extreme case where $x\approx0.84$. Furthermore, the structure of the FEM pattern for $x\approx1$ appears much more like a normal metal FEM pattern than do those observed for $x<1$. We believe that all these pieces of evidence suggest a more stable surface, more resistant to residual gas adsorption, for ZrC_1 than for substoichiometric ZrC.

Emission of high currents, of the order of 1 mA from individual emitters, is an interesting phenomenon. We do not yet understand the nature of changes at the emitting surface under these emission conditions, although ZrC_1 seems to be very stable relative to substoichiometric material. It seems likely that significant heating of the emitting region is occurring, an effect that is especially plausible for solid carbide emitters because of the relatively low thermal conductivity of the material.

Electron emission in controlled gas atmospheres is an area of study that has many practical applications. We supposed that ZrC would operate well under these conditions, because of its high sputtering resistance and relative chemical inertness, at least compared with pure metals. Our preliminary results seem to support this idea. The recent findings with ZrC_1, discussed above, would seem to imply that the best results would be obtained with this stoichiometric material.

This work, at high currents (around 1 mA) and in poor vacuum (in the 10^{-7} to 10^{-5} range), could be valuable in several practical areas. Uses might include e-beam lithography, accelerator sources, FEAs for video imaging, microwave devices, small ion thruster neutralizing beams, and other applications where high current density source are desired.

ACKNOWLEDGMENT

This work is supported by DARPA High Definition Systems Initiative under ONR Grant No. N00014-96-1-1011.

[1]W. A. Mackie, J. L. Morrissey, C. H. Hinrichs, and P. R. Davis, J. Vac. Sci. Technol. A **10**, 2852 (1992).
[2]W. A. Mackie, J. L. Morrissey, C. H. Hinrichs, R. L. Hartman, and P. R. Davis, in *Tri-Service/NASA Cathode Workshop*, 1992 Proceedings, edited by J. W. Gibson (Palisades Institute for Research Services, Washington, 1992).
[3]W. A. Mackie, R. L. Hartman, and P. R. Davis, Appl. Surf. Sci. **67**, 29 (1993).

[4]M. L. Yu, B. W. Hussey, T. H. Philip Chang, and W. A. Mackie, J. Vac. Sci. Technol. B **13**, 2436 (1995).
[5]M. L. Yu, H. S. Kim, B. W. Hussey, T. H. P. Chang, and W. A. Mackie, J. Vac. Sci. Technol. **14**, 2041 (1996).
[6]W. A. Mackie, T. Xie, J. E. Blackwood, S. C. Williams, and P. R. Davis, J. Vac. Sci. Technol. B **16**, 1215 (1998).
[7]W. A. Mackie, T. Xie, and P. R. Davis, "Field emission from HfC or Ni films deposited on single tip Mo field emitters." paper delivered at ICOPS-97, San Diego, CA, May 1997.
[8]W. A. Mackie, T. Xie, J. E. Blackwood, and P. R. Davis, "Field emission from HfC films on Mo field emitter arrays and from HfC arrays." paper delivered at ICOPS-97, San Diego, CA, May 1997.
[9]W. A. Mackie and C. H. Hinrichs, J. Cryst. Growth **87**, 101 (1988).
[10]High stoichiometry material obtained from Applied Physics Technologies, 5882 Booth Bend Rd., McMinnville, OR 97128.
[11]W. A. Mackie, R. L. Hartman, M. A. Anderson, and P. R. Davis, J. Vac. Sci. Technol. B **12**, 722 (1994).
[12]W. A. Mackie, T. Xie, and P. R. Davis, J. Vac. Sci. Technol. **13**, 2459 (1995).

Experimental evaluation of the extended Schottky model for ZrO/W electron emission

M. J. Fransen[a)]
Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

J. S. Faber
Philips Elecron Optics/FEI, Eindhoven, The Netherlands

Th. L. van Rooy and P. C. Tiemeijer
Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

P. Kruit
Department of Applied Physics, Delft University of Technology, Delft, The Netherlands

(Received 19 September 1997; accepted 29 May 1998)

The energy distribution of electrons emitted from a ZrO/W electron source with a radius of curvature of 0.9 μm was determined for extraction voltages ranging from 3000 to 6000 V and emitter temperatures from 1200 to 1900 K. Full width at half maximum values of the energy spectra between 0.3 and 0.8 eV were found. The experimental data are analyzed according to the extended Schottky model for electron emission, yielding the electric field at the emitter surface. The tunneling probabilities are evaluated numerically by integrating the one-dimensional Schrödinger equation, and analytically by employing the commonly used Wentzel–Kramers–Brillouin approximation. Both approaches give good agreement with the experimental data, except for small differences probably due to Coulomb interactions (Boersch effect). The same analysis is also applied to the experimental results for a Schottky emitter with a radius of curvature of 0.3 μm, taken from literature. © *1998 American Vacuum Society.* [S0734-211X(98)13104-9]

I. INTRODUCTION

The ZrO/W electron emitter,[1–3] a tungsten ⟨100⟩-surface covered with ZrO has become increasingly popular as electron source for high-performance electron beam instruments, like electron microscopes and electron beam pattern generators. The emitter is often referred to as "Schottky emitter," due to its use of the Schottky effect for electron emission. The Schottky emitter delivers a stable emission current comparable with thermionic electron sources,[4] and a brightness approaching field emitters.[5] The third important parameter for an electron source, the spread in energy of the emitted particles, has only been studied by a few groups in the case of the ZrO/W emitter.[3,6–8]

Accurate values for the energy spread of the Schottky emitter have become more necessary due to the rising importance of low-voltage scanning electron microscopy. The chromatic aberration of the objective lens has now become the resolution limiting factor,[9] so the energy spread of the electron source should be as low as possible while retaining a sufficiently high brightness.

The lack of recent data was recognized by other groups too, and recently the energy spectra of a 0.3 μm radius of curvature Schottky emitter were published.[10,11] In these articles the experimental data were compared with the field-assisted thermionic emission model, also known as *Schottky model*. However, for increasing field strengths, the Schottky model is only approximately valid, as electron tunneling through the potential barrier between the emitter and vacuum cannot be neglected anymore. In the present article we use a model in which the tunnel effect is included, the *extended Schottky model*.[3,12] We compare our experimental results on the energy distribution of a 0.9 μm radius of curvature Schottky emitter with the extended Schottky model. Due to the approximations in the derivation of this model, it also breaks down as the field increases. With a simple modification, energy spectra for fields above the extended Schottky limit can be described analytically. We use this method for the spectra published in Ref. 11. The extended Schottky model for electron emission makes use of the Wentzel–Kramers–Brillouin (WKB) approximation to the Schrödinger equation and estimates the potential barrier at the interface between the emitter and vacuum with a parabolic function. We will check the validity of these arguments by integrating the general equation for electron emission numerically and discuss the differences encountered between the analytical models and the numerical calculation.

II. MEASURING ENERGY SPECTRA

The energy spectra in this article were obtained using a commercially available Schottky emitter[13] placed in front of a slightly modified hemispherical analyzer.[14] The whole setup is operated in an ultrahigh vacuum system. During experiments, the pressure in the vacuum chamber was 5×10^{-8} Pa. The experimental system was mounted on air cushions to overcome the effect of mechanical vibrations. A schematic view of the setup is shown in Fig. 1.

The emitter is mounted in a compact module that can be shifted and rotated in front of the analyzer. Between module

[a)]Also at Department of Applied Physics, Delft University of Technology, Delft, The Netherlands; electronic mail: fransenm@natlab.research.philips.com

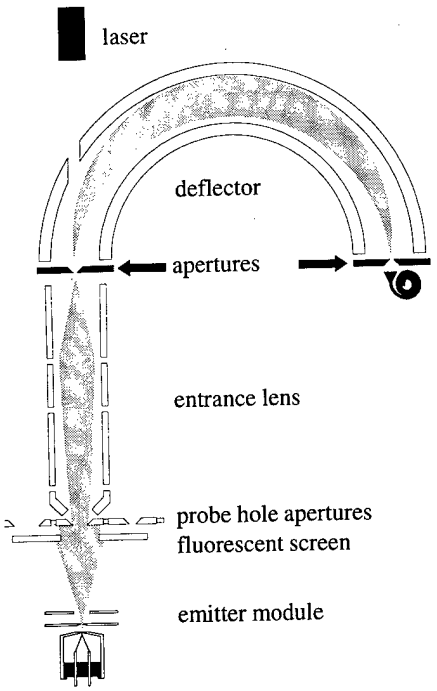


FIG. 1. Experimental setup: a Schotty emitter with hemispherical electron energy analyzer. The laser is used for alignment purposes.

and analyzer, a fluorescent screen with a hole in it is mounted for ease of alignment. The module is drawn schematically in Fig. 2. The emitter is spotwelded on a filament and can be heated resistively with a current supply. The heating current in the experiments ranged from 1.9 to 2.3 A, corresponding to emitter temperatures between 1200 and 1900 K, as determined by an optical pyrometer. Due to the asymmetric location of the current source, the potential of the emitter tip will change slightly with varying heating current. The emitter is surrounded by three electrodes: the suppressor, the extractor and the lens electrode. These electrodes are controlled with power supplies that float on the beam potential V_{beam}, -400 V in all experiments. During the experiments the suppressor was kept at a constant voltage of -500 V with respect to the emitter. The extractor voltage varied between 3000 and 6000 V, referred to the emitter. The upper value of the extraction voltage was set by the range of the high voltage supplies we used. The lens electrode is meant for keeping the illumination area on the fluorescent screen constant when changing the extraction voltage. In this setup, the extraction field on the emitter can be varied while the beam potential and thus the energy window of the analyzer remains constant. The emission current leaving the module (the beam current) is determined by measuring the voltage drop over a resistor R at the grounded side of the beam potential power supply. As a result of this, the potential of the module changes slightly when the emission current changes. In the experimental results presented in this article this effect is corrected.

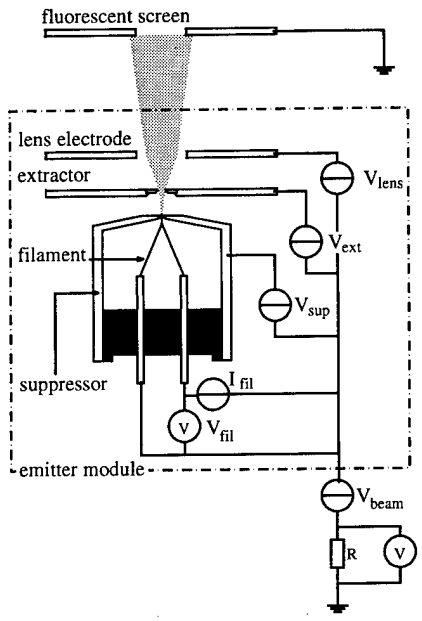


FIG. 2. Emitter module.

As energy-dispersive element, a commercially available hemispherical electron energy analyzer[14] was chosen. The resolution was improved considerably by decreasing the aperture size at the entrance and exit side of the half-spheres of the analyzer. For the experiments in this article aperture sizes of 0.1 and 0.03 mm were used. The pass energy of the electrons in the deflector was 9 eV. A channeltron is used to count the electrons that pass the second deflector aperture. In order to limit the current entering the analyzer, a set of probe hole apertures is placed between the fluorescent screen and the entrance lens of the analyzer. A typical probe hole aperture size in the experiments was 0.4 mm.

We checked the resolution of the analyzer by investigating field emission energy spectra from a tungsten ⟨111⟩-oriented electron emitter taken at 77 K. As suggested by Young and Kuyatt[15] the sharp falling edge of the energy distribution at the Fermi level is a sensitive indicator of analyzer performance. With approximately the same beam parameters as used for the ZrO/W emitter, an analyzer resolution of 0.06 eV is estimated, limited by the stability of the control electronics.

By computer control of the electron gun module and analyzer voltages, we collected energy spectra at seven extraction voltages and nine temperatures. We will not show all spectra individually, but show some examples and discuss the general trends.

III. REVIEW OF SCHOTTKY EMISSION THEORY

In this section we will review the theory of electron emission and the approximations made to derive expressions in closed form for the total energy spread and the emission current density. We will follow the approach outlined by Hawkes and Kasper,[16] resulting in equations as listed in Appendix I of Ref. 3.

The emission of electrons from a metal is a function of the temperature and the applied field. In order to find an expression for the energy spread, the effects of temperature

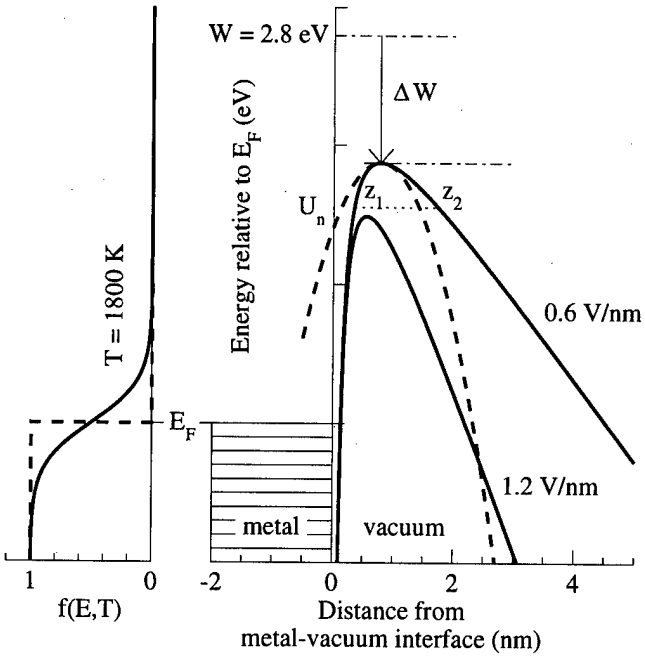


FIG. 3. Schematic diagram of the interface between a metal and vacuum. In the left part of the diagram the Fermi–Dirac electron energy distribution inside the metal at 1800 K is drawn. The right part of the drawing displays the interface itself. The effective potential barrier according to Eq. (10) is drawn for electric fields of 0.6 V/nm and 1.2 V/nm with solid lines. The work function lowering ΔW is indicated. The parameters z_1 and z_2 are integration limits for the calculation of the transmission coefficient $D(U_n)$ (17). The parabolic approximation (19) to the actual potential function is displayed with a dashed line.

and field are considered separately. The distribution of electron energies E in a metal is a function of the temperature T. The well-known Fermi–Dirac distribution function describes this:

$$f(E,T)=\frac{1}{1+\exp[(E-\zeta)/k_BT]}. \tag{1}$$

The constant k_B is Boltzmann's constant, ζ is the Fermi level. We neglect the temperature dependence of the Fermi level and use $\zeta=E_F+V_0$, in which V_0 determines the origin of the energy scale. The possibility of choosing V_0 freely will be used to ease subsequent calculations. The Fermi–Dirac distribution $f(E,T)$ is drawn on the left-hand side of Fig. 3 for 0 K with a dashed line and for 1800 K with a solid line. In this graph, the origin of the energy scale is chosen at the Fermi level E_F.

Electrons are kept inside the metal because of a potential difference $V(\boldsymbol{r})$ with the surrounding domains. The chance for an electron to overcome this potential barrier and enter the vacuum domain is described by a transmission coefficient D, being the ratio between the electron intensity after passing the potential barrier and the incident intensity. It is a function of the normal component U_n of the total energy E. The shape and height of the potential barrier, and thus the transmission coefficient, are functions of the applied field.

Analogous to Hawkes and Kasper,[16] the electric current density distribution $d^3\boldsymbol{j}(\boldsymbol{r})$ is expressed as

$$d^3\boldsymbol{j}(\boldsymbol{r})=\frac{2e}{mh^3}\boldsymbol{p}f(E,T)D(\boldsymbol{p},\boldsymbol{r})d^3p, \tag{2}$$

in which $\boldsymbol{p}$ is the electron impulse vector, e and m are the electron charge and mass, respectively, and h is Planck's constant.

For ease of calculation, the curving of the surface of the emitter is neglected and considered as plane, hence we can drop the $\boldsymbol{r}$ dependence in Eq. (2). A coordinate system is defined, such that the interface is in the xy plane at $z=0$. The electric potential $V(\boldsymbol{r})$ is now a function of z only, and the normal and transverse components of the energy U_n and U_t are conserved. In order to find expressions for the total energy distribution, we have to transfer to polar coordinates. The momenta are expressed in polar coordinates:

$$\begin{aligned} p_x&=\sqrt{2m[E-V(z)]}\,\sin\theta\cos\phi,\\ p_y&=\sqrt{2m[E-V(z)]}\,\sin\theta\sin\phi,\\ p_z&=\sqrt{2m[E-V(z)]}\,\cos\theta. \end{aligned} \tag{3}$$

With these equations, we can calculate the volume element in momentum space at each plane z as:

$$d^3p=p^2dp\,d\Omega=m\sqrt{2m[E-V(z)]}\,\sin\theta\,dE\,d\theta\,d\phi. \tag{4}$$

The total energy distribution can now be calculated by integration of Eq. (2) over all emission angles ($0<\phi<2\pi$, $0<\theta<\pi/2$), once the transmission coefficient is found. Several approximations to the transmission coefficient exist, depending on the strength of the applied field. We will discuss each approximation separately and start with the case in which the field is negligible.

For *thermionic* electron emission, the weak field that is applied to draw electrons from the surface of the metal can be completely neglected. Without this field, the potential barrier $V(z)$ between metal and vacuum can be approximated by a step function, with a step height W: the work function (2.8 eV for a ZrO/W emitter). The step height is indicated in Fig. 3 by a dash-dotted line at 2.8 eV. Only electrons at energy levels higher than the work function W can leave the metal directly. The origin of the energy scale is chosen at the top of the potential barrier. $V(z)$ becomes

$$V(z)=\begin{cases}-(E_F+W) & \text{for } z<0,\\ 0 \quad \text{for } z\geqslant 0.\end{cases} \tag{5}$$

The transmission factor $D(U_n)$ is given by

$$D(U_n)=\begin{cases}0 & \text{for } U_n<0,\\ 1 & \text{for } U_n\geqslant 0.\end{cases} \tag{6}$$

The current density d^3j_z is now

$$d^3j_z=\frac{ep\cos\theta}{m}\frac{2}{h^3}f(E,T)D(U_n)d^3p$$

$$\approx \frac{4me}{h^3} E \exp\left(-\frac{E+W}{k_b T}\right) \sin\theta \cos\theta \, dE \, d\theta \, d\phi, \tag{7}$$

in which $f(E,T)$, describing the energy distribution of electrons in the metal, is simplified by omitting the unit term in the denominator of Eq. (1). This approximation is allowed, because the only energies of interest are located around the top of the potential barrier, far away from the Fermi level. After writing the transmission coefficient in terms of θ and E using Eq. (3), the total energy distribution for thermionic emission dj_0 can be calculated:

$$dj_0 = \frac{4\pi me}{h^3} E \exp\left(-\frac{E+W}{k_B T}\right) dE, \tag{8}$$

and a subsequent integration over all energies $(0<E<\infty)$ yields the emission current density:

$$j_0 = \frac{4\pi me}{h^3} (k_B T)^2 \exp\left(-\frac{W}{k_B T}\right), \tag{9}$$

the Richardson–Dushman equation for thermionic emission. The thermionic energy distribution is plotted in Fig. 4(a). It has a sharp leading edge starting at the top of the potential barrier W and a falling edge determined by the Fermi–Dirac distribution function $f(E,T)$.

If the field is increased, its influence on the effective work function cannot be neglected anymore. In order to calculate the potential distribution $V(z)$ it is necessary to include the potential of image charges in the model. This yields a potential function made out of three components: the work function W, the effect of the electric field F and the image charge potential:

$$V(z) = \begin{cases} 0 & \text{for } z<0, \\ W - eFz - e^2/16\pi\epsilon_0 z & \text{for } z>0, \end{cases} \tag{10}$$

with ϵ_0 the permittivity of free space. In Fig. 3, this equation is plotted for two fields: 0.6 V/nm and 1.2 V/nm. It can be seen that the effective barrier height is lowered as the field increases. The top of the potential barrier is located at

$$z_m = \sqrt{e/16\pi\epsilon_0 F} \tag{11}$$

and

$$V_m = V(z_m) = W - \Delta W = W - \sqrt{e^3 F/4\pi\epsilon_0}. \tag{12}$$

The origin of the energy scale for the calculation of the total energy distribution is again chosen at the top of the barrier. The transmission coefficient is given by

$$D(U_n) = \begin{cases} 0 & \text{for } E<-(E_F+V_m), \\ 1 & \text{for } E>-(E_F+V_m). \end{cases} \tag{13}$$

The calculation of the total energy distribution and the emission current density is completely analogous to the thermionic emission case considered earlier:

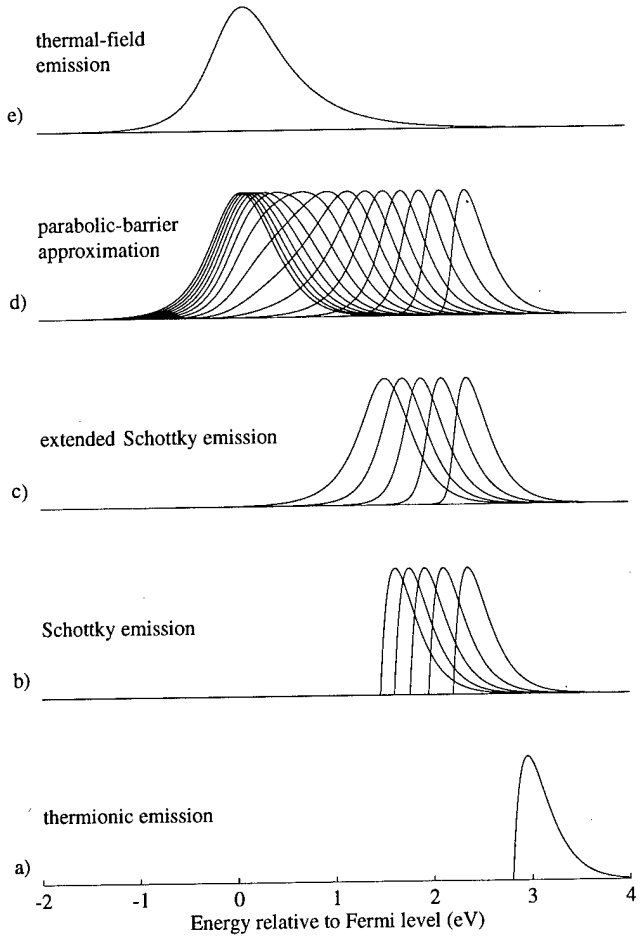


FIG. 4. Different electron emission models and the corresponding energy spectra for several values for the electric field, as mentioned in the text. The work function in these calculations is 2.8 eV, the emitter temperature is 1800 K.

$$dj_S = \frac{4\pi me}{h^3} E \exp\left(-\frac{E+V_m}{k_B T}\right) dE. \tag{14}$$

If we choose the origin of the energy distribution (14) at the Fermi level E_F, it can be seen that the energy distribution will shift towards lower energies for increasing F. In Fig. 4(b) this effect is plotted for five fields ranging from 0.25 to 1.25 V/nm.

The total emission current is given by

$$j_S = \frac{4\pi me}{h^3} (k_B T)^2 \exp\left(-\frac{W-\Delta W}{k_B T}\right) = \exp\left(\frac{\Delta W}{k_B T}\right) j_0. \tag{15}$$

The enhancement of the emission current with respect to the thermionic case due to the lowering of the potential barrier is known as the *Schottky effect*. Plotting $\ln(j)$ as a function of $\sqrt{F}$ should result in a straight line: the Schottky plot. The Schottky model is generally used for the description of ZrO/W electron emission. However, it does not describe the contribution of electron tunneling to the energy spectrum for increasing electric fields.

IV. EMISSION THEORY FOR INCREASING FIELDS

A better description of ZrO/W electron emission can be found by incorporating the effect of electron tunneling through the potential barrier in the transmission coefficient of the Schottky model. The transmission coefficient can now be found by solving the Schrödinger equation for the potential barrier given by Eq. (10). In order to do so, we adopt two approaches.

In the first one, we simply integrate the one-dimensional Schrödinger equation numerically. We choose the starting condition of the numerical integration such that the wave function for large positive z can be written as $\Psi(z) = a_{\text{out}} \exp(ik_{\text{out}}z)$. For negative z the calculated wave function can be decomposed as $\Psi(z) = a_{\text{in}} \exp(ik_{\text{in}}z) + b_{\text{in}} \times \exp(-ik_{\text{in}}z)$. The transmission coefficient follows as

$$D(U_n) = \frac{k_{\text{out}}}{k_{\text{in}}} \left| \frac{a_{\text{out}}}{a_{\text{in}}} \right|^2. \tag{16}$$

The advantage of this approach is that no approximations are made. A drawback is that it gives no closed expression for the total energy distribution and the emission current density. As far as we know, a numerical evaluation of $D(U_n)$ as in Eq. (16) has not yet been done.

In the second approach, which is commonly used, the transmission coefficient is evaluated according to the WKB approximation as

$$D(U_n) = \frac{1}{1+\exp(G)}, \tag{17}$$

with G the Gamov exponent:

$$G(U_n) = \frac{2}{\hbar} \int_{z_1}^{z_2} \sqrt{2m(V(z) - U_n)}\, dz. \tag{18}$$

The limits z_1 and z_2 of the integral correspond to the intersection of U_n with $V(z)$, as indicated in Fig. 3.

The Gamov exponent can be integrated analytically. The resulting expression however, contains elliptic integrals which is not convenient for subsequent integrations. Because most electrons will tunnel near the top of the barrier, it is sufficient to use an approximation that is only valid in this region. The barrier function (10) is Taylor-approximated with a second-order polynomial:[12]

$$V_a(z) \approx V_m - \frac{e^2}{16\pi\epsilon_0 z_m^3}(z - z_m)^2. \tag{19}$$

This parabolic function is drawn in Fig. 3 with a dashed line. With this approximation, Eq. (18) can be written in closed form, yielding:

$$G(U_n) = \frac{(V_m - U_n)}{\kappa} \tag{20}$$

with

$$\kappa = \frac{\hbar}{\pi\sqrt{m}}(4\pi\epsilon_0 e F^3)^{1/4} = c_1 F^{3/4}. \tag{21}$$

As stated by Hawkes and Kasper[16] the resulting transmission coefficient may also be used above the top of the potential barrier, although Eq. (18) does not hold there.

With the choice of the origin of the energy scale at the top of the potential barrier, the current density d^3j can be integrated over all emission angles, yielding the total energy distribution in the parabolic-barrier approximation dj_{PBA}:

$$dj_{\text{PBA}} = \frac{4\pi me}{h^3} \frac{\kappa}{1+\exp((E+W-\Delta W)/k_BT)} \times \ln\left[1+\exp\left(\frac{E}{\kappa}\right)\right] dE. \tag{22}$$

The total energy distribution in the *extended Schottky model* dj_{ES} is found after omission of the unit term in the denominator of the Fermi–Dirac distribution function. This simplification is necessary in order to find an expression for the emission current density in closed form. Integration of dj_{ES} from $E = -\infty$ to ∞, using a tabulated integral,[17] yields:

$$j_{\text{ES}} = \frac{4\pi me}{h^3}(k_BT)^2 \exp\left(-\frac{W-\Delta W}{k_BT}\right) \frac{\pi q}{\sin \pi q} = \frac{\pi q}{\sin \pi q} j_S, \tag{23}$$

with $q = \kappa/k_BT = c_1 F^{3/4}/k_BT$. For low fields q approaches zero and the current density calculated with the extended Schottky model reduces to the Schottky model. As the field increases and/or the temperature decreases q will rise and the current density is increased due to the contribution of the tunneling effect. When q approaches unity Eq. (23) breaks down, due to the omission of the unit term in the denominator of the Fermi–Dirac distribution function $f(E,T)$. The energy distribution calculated with the extended Schottky model is shown in Fig. 4(c) for the same fields as the Schottky model. For increasing fields, the slope of the leading edge of the spectrum decreases, due to the additional contribution of electrons tunneling through the potential barrier.

If we use $f(E,T)$ in its exact form (22), energy spectra can be calculated for larger q values. Energy spectra calculated with the parabolic-barrier approximation are plotted in Fig. 4(d) for fields ranging from 0.25 to 4 V/nm with an increment of 0.25 V/nm. It can be seen in Fig. 4(d) that the width of the energy distribution increases considerably and decreases again as the field rises. The spectrum shifts towards the Fermi level and stays close to the Fermi level once a certain value for the field is reached. With the parabolic-barrier approximation we can describe the gap between the regions of validity of the extended Schottky model and the thermal-field emission model[3] at least qualitatively.

The *thermal-field regime* is an extension to the cold field emission theory for non-zero temperatures (see, e.g., Ref. 16). A comparison of the last plot of Fig. 4(d) with the energy distribution in the thermal-field regime [Fig. 4(e)] for the same field of 4 V/nm using Eq. (44.31) of Ref. 16 shows that the parabolic-barrier approximation has a reasonable similarity with the spectra in the thermal-field regime, at least for low work function electron emitters.

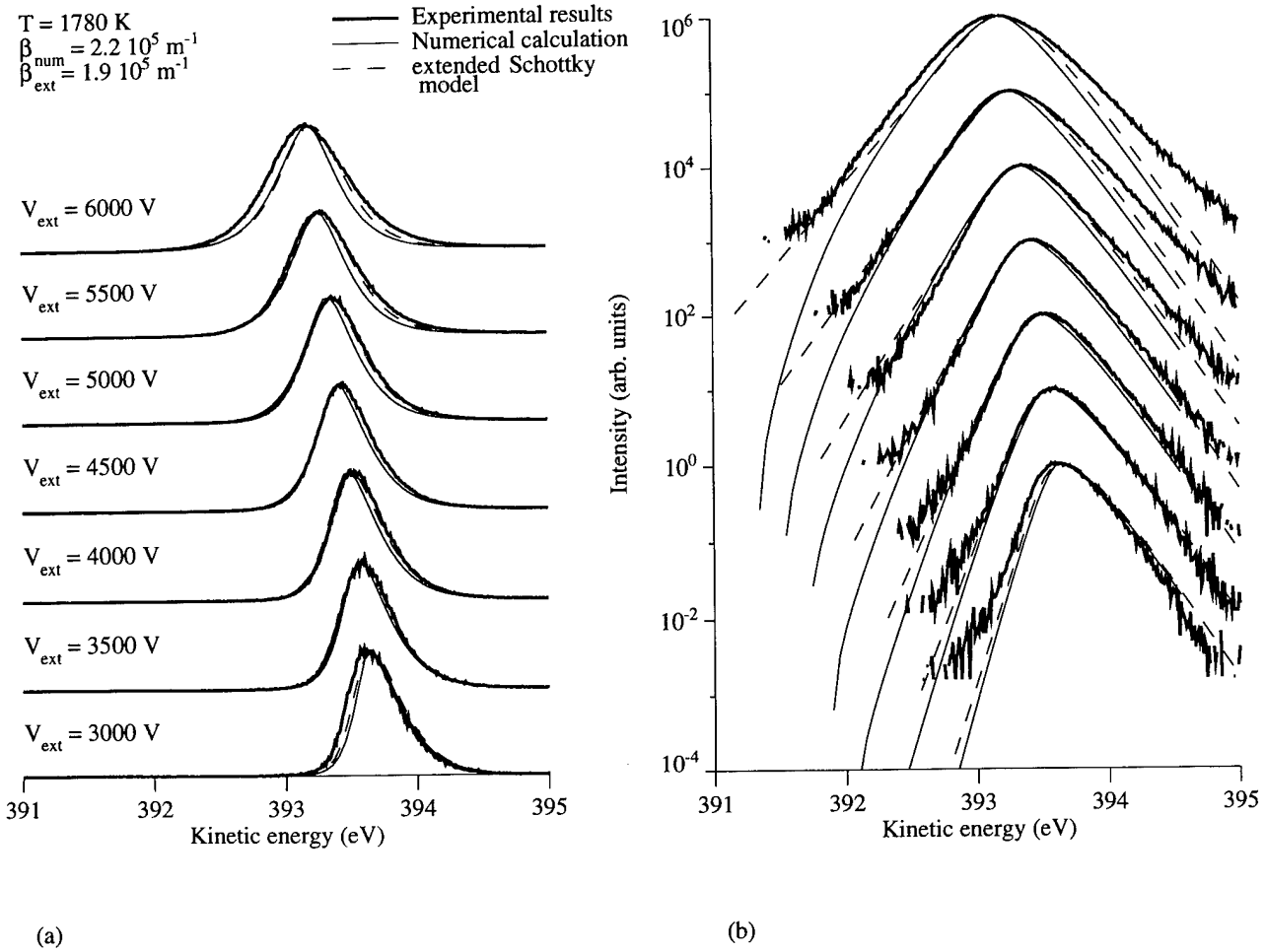


FIG. 5. Experimental results taken at 1780 K (thick solid line) and fitted spectra using Eq. (22) (dashed line), with the unit term in the denominator of the Fermi–Dirac electron energy distribution function omitted, and the numerical model (thin solid line). In (a) the spectra are plotted on a linear y scale, and normalized to their maximum value. In (b) the same normalized spectra are plotted on a logarithmic scale, with arbitrary vertical offsets.

V. DISCUSSION OF EXPERIMENTAL RESULTS

We will discuss the applicability of the extended Schottky model for ZrO/W electron emission by comparing experimental spectra with Eq. (22). We will investigate the influence of the omission of the unit term in the denominator of the Fermi–Dirac distribution term in Eq. (22) and use the spectra obtained by the numerical method as well.

Before doing so we have to find a value for the electric field at the surface of the emitter. The linear relation between the extraction voltage V_{ext} and the electric field at the surface of the emitter F is expressed as $F=\beta V_{ext}$, in which the factor β is a complicated function of several geometrical and electrical parameters. It is known in literature that the electric field changes over the surface of the emitter due to the facet at the end of the tip, causing a ring-shaped emission pattern.[18] This effect can be estimated by a numerical calculation of the potential distribution in the emitter module, and especially the varying electric field at the tip, but instead of this we will assume that β is constant by taking a small angle from the middle part of the emission pattern that can be observed on the fluorescent screen in front of the energy analyzer. With this constant β value, we fit calculated energy distributions to experimental spectra taken at different temperatures and extraction voltages. We obtained energy spectra for nine different temperatures, ranging from 1200 to 1900 K, and for seven extraction voltages between 3000 and 6000 V. We fitted all 63 spectra using two fit parameters. The factor β is assumed to stay constant, in other words: we do not expect a significant change in emitter geometry during the experiment. For each emitter temperature a different offset was necessary to account for the shift in voltage induced by the asymmetric location of the heating supply of the emitter, and the beam potential itself. For the extended Schottky model the best fit was found for $\beta=1.9\times10^5\ \mathrm{m}^{-1}$. With the numerical model a somewhat higher value of $2.2\times10^5\ \mathrm{m}^{-1}$ accompanied with a slightly different offset was necessary for a good fit. The difference between the two models is caused by the fact that the WKB transmission coefficient is overestimated below the top of the potential barrier in the extended Schottky model, as compared to the transmission coefficient found directly by numerical integration.

As an example of the quality of our fits we show Fig. 5(a), where the spectra are plotted on a linear scale and Fig. 5(b) with plots on a logarithmic scale. The temperature of 1780 K

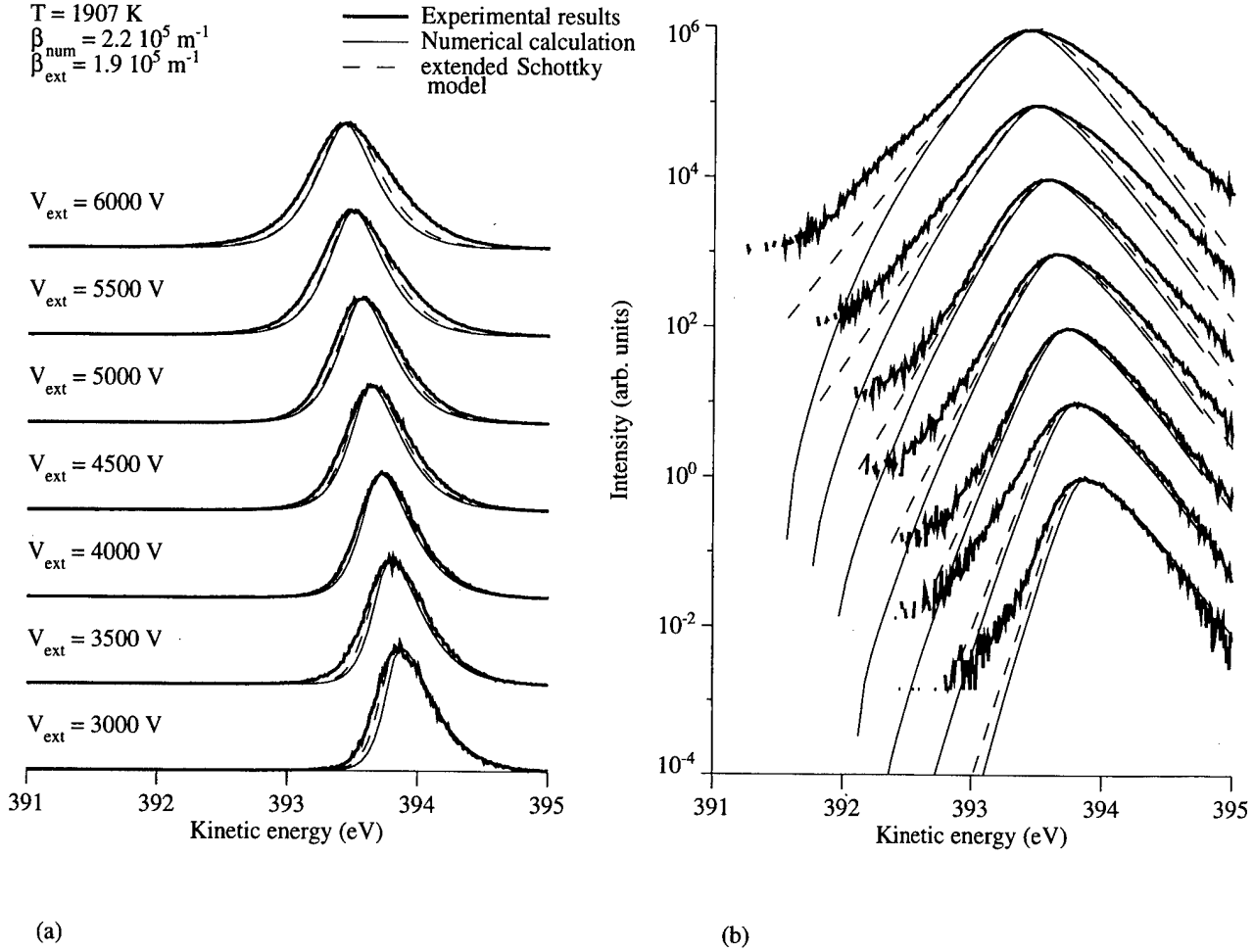


FIG. 6. Same as Fig. 5, but now taken at 1907 K. Note the symmetrical broadening of the experimental spectra with respect to the calculated distributions occurring already at medium extraction voltages.

is close to the optimum operating temperature of the Schottky emitter (1800 K). It is clear from the graphs that the fit of the simulations to the experimental data is good, especially at low extraction voltages. At higher voltages however, the experimental energy distribution is broadened with respect to the simulated spectrum. We suspect that electron–electron interactions (Boersch effect) come into play at high emission current densities. Similar spectra taken at 1907 K with the same range of extraction voltages (Fig. 6) support this: the broadening of the experimental spectra with respect to the calculations becomes evident at medium extraction voltages already.

For increasing q, i.e., for low temperatures and high fields, the accuracy of the extended Schottky model for electron emission should decrease. In Fig. 7 we show energy spectra obtained at 1210 K. It can be seen that the fit of the extended Schottky model is reasonable, even with $q=0.85$. A closer inspection of the logarithmically plotted spectra [Fig. 7(b)] reveals that the extended Schottky model deviates from the experimental results at the highest extraction voltages. For higher fields it is necessary to use the parabolic-barrier approximation.

In Fig. 8 the full width at half maximum (FWHM) energy spread of each experimental spectrum is plotted as a function of the emission current density, calculated with the extended Schottky model, Eq. (23). The solid lines in Fig. 8 represent FWHM values for the same conditions as in the experiment, as calculated with the total energy distribution in the extended Schottky model dj_{ES}. At low current densities, the experimental FWHM values correspond reasonably to the calculated energy spreads. For current densities exceeding $j_{ES}=10^8$ A/m^2 the experimental spread deviates from the calculated FWHM values considerably, which we assume to be due to the Boersch effect.

Let us estimate the magnitude of the Boersch effect in our electron-optical system, in order to see whether this is a viable assumption. For this we use Kruit and Jansen's theory[19] for the Boersch effect. A quick calculation of the regions in our system shows that a significant energy broadening due to the Boersch effect only occurs in the region between tip and extractor. We model the source as a crossover followed by a drift space in which the beam diverges. We thus assume that the electrons gain their full kinetic energy directly after leaving the source. In order to use the analytical approximations for the energy broadening due to the Boersch effect in a

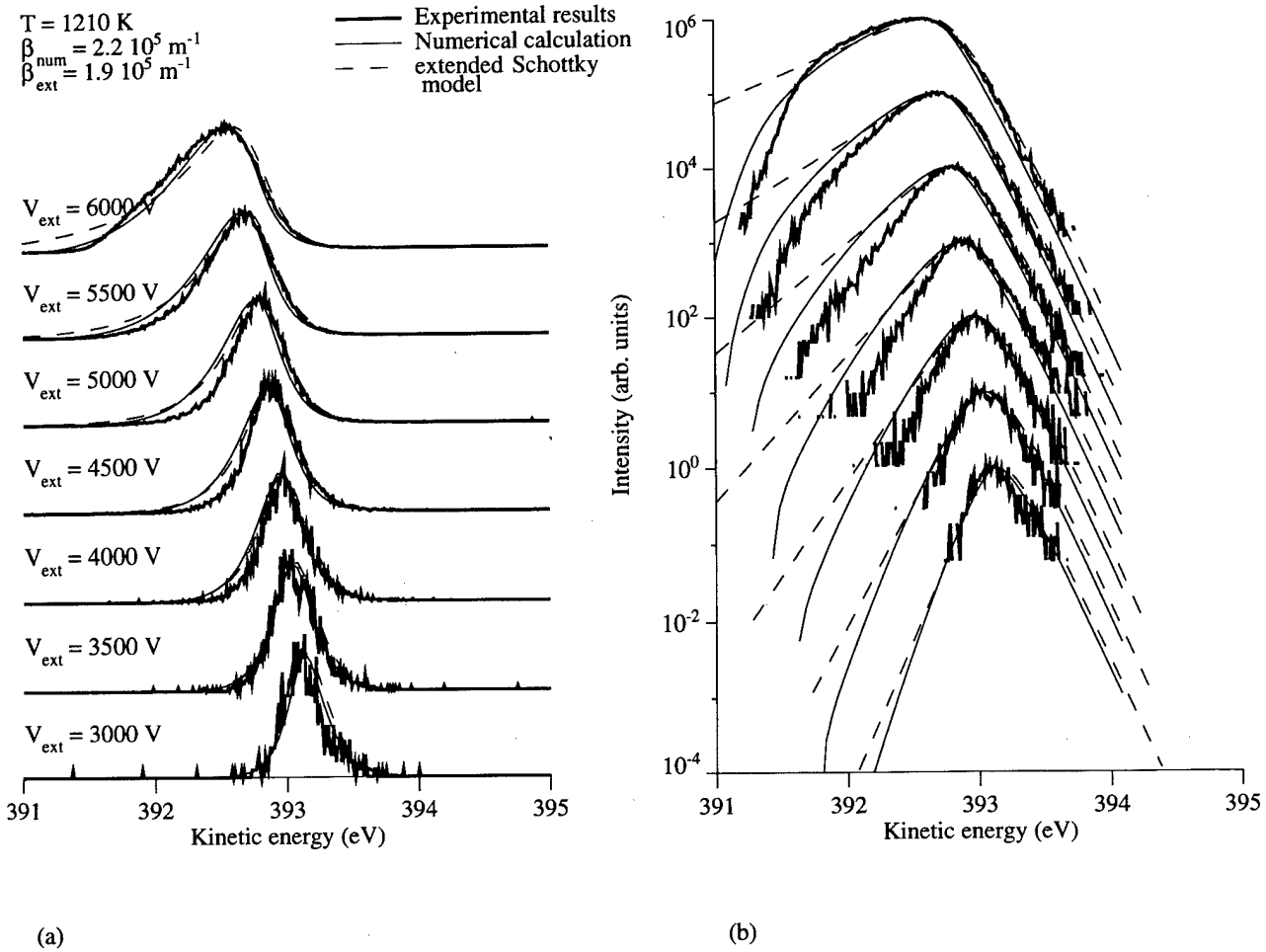


FIG. 7. Same as Fig. 5, but now taken at 1210 K. Even for high q values the fit of the extended Schottky model to the experimental spectra is reasonable.

narrow crossover given by Kruit and Jansen, we need to know the current from the emitter facet I_f, the half-opening angle of the beam α and the radius of the crossover r_c. We calculate the facet current I_f from the current leaving the emitter module (see Fig. 2), using the additional assumption that the half-opening angle α of the beam emitted from the central facet is 7°.[7] The radius of the crossover r_c can be found from the emitting area I_f/j_{ES}.

We estimate the contribution of the additional energy broadening in our experimental spectra by subtracting quadratically the energy spread calculated with Eq. (22) from the experimental energy spread. For the three spectra from Fig. 6 taken at 6000, 5500, and 5000 V, we find experimental broadenings of 0.44, 0.44, and 0.41 eV, respectively. The theoretical values, derived from Kruit and Jansen's theory, are 0.3, 0.2, and 0.2 eV. Hence, we consider it likely that the

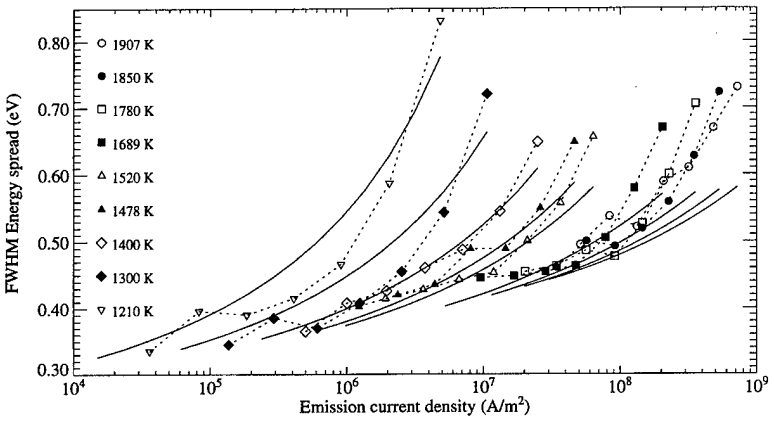


FIG. 8. FWHM values of the experimental energy distributions as a function of the emission current density, calculated with Eq. (23) using $\beta = 1.9 \times 10^5$ m^{-1}. The solid lines are calculated FWHM values using Eq. (22).

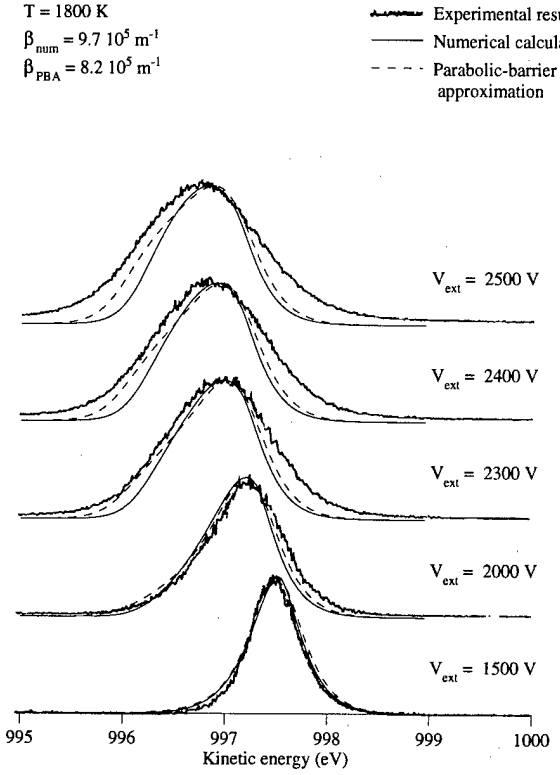


FIG. 9. Experimental energy spectra taken from Fig. 3 of Ref. 11. Energy spectra, calculated with Eq. (22) in its exact form, and the numerical model, are fitted to the experimental spectra. A slightly better fit can be obtained by small individual variations of the β values for each spectrum.

additional broadening in our experimental spectra is due to the Boersch effect.

VI. COMPARISON WITH OTHER WORK

Recently published energy spectra of a Schottky emitter with a radius of curvature of 0.3 μm[11] show large FWHM values accompanied by a strong change in the shape of the spectra for increasing fields. In Fig. 9 we show experimental energy spectra, taken from Fig. 3 of Ref. 11, at an emitter temperature of 1800 K. We fitted the parabolic-barrier approximation and the numerical model to these experimental spectra, using β and an offset to the graphs accounting for the beam potential as the only fit parameters. In this case, it was necessary to use Eq. (22) in its exact form. For the parabolic-barrier approximation, the best fit was found for $\beta = 8.2 \times 10^5$ m^{-1}, and for the numerical model $\beta = 9.7 \times 10^5$ m^{-1} yielded the best results. By fitting each spectrum with an individual β, the fit becomes slightly better, the necessary variation in β being below 10 %.

At the lowest fields in Fig. 9 the fit of both models is in good agreement with the experimental results. At the highest extraction voltages in Fig. 9 the accuracy of the parabolic-barrier approximation decreases, as expected, when compared to the full numerical calculation.

For increasing fields, a symmetrical broadening of the experimental energy distribution with respect to the numerical model occurs. We did not succeed in explaining this effect with a different choice for β. The broadening may be due to the Boersch effect. Again, we can use the Kruit and Jansen equations[19] to determine the magnitude of the Boersch effect, by modeling the source as a crossover followed by a diverging beam. We calculate the current emitted by the facet I_f from the angular current density given in Ref. 11, assuming the same half-angle of 4° accepted by the extractor as used in our own experiments. Furthermore, the half-angle of the beam emitted from the tip apex is again taken as 7°. The radius r_c of the crossover is estimated from the emitting area I_f/j_{PBA}. The current density j_{PBA} is calculated by numerical integration of Eq. (22). We estimate additional energy broadenings due to the Boersch effect for the upper three spectra from Fig. 9 (taken at 2500, 2400, and 2300 V) of 0.70, 0.55, and 0.45 eV. By subtracting quadratically the FWHM energy spread calculated with the parabolic-barrier approximation (22) from the experimental spectra we find additional broadenings of 0.4, 0.3, and 0.3 eV. This indicates that also in these spectra the additional energy broadening is due to the Boersch effect, as stated by the authors.

VII. CONCLUSIONS

We have shown experimental energy distributions, obtained from a Schottky emitter with a radius of curvature of 0.9 μm, for emitter temperatures ranging from 1200 to 1900 K and extraction voltages between 3000 and 6000 V. The experimental results can be explained adequately with the extended Schottky model, as verified by fitting total energy distributions, calculated with this model, to the experimental results. The difference between the extended Schottky model and a numerical evaluation of the general equation for electron emission is small. In the numerical model a 10% higher extraction field has to be chosen in order to obtain a good fit. The reason for this is that the parabolic approximation to the actual potential barrier overestimates the tunnel probability for an electron.

For field strengths exceeding the range of validity of the extended Schottky model, energy spectra can be calculated by using the "parabolic-barrier approximation." In this model the Fermi–Dirac distribution function is used in its exact form. With the parabolic-barrier approximation the energy spectra published by Kim *et al.*[11] for a Schottky emitter with a radius of curvature of 0.3 μm can be explained qualitatively. The numerical calculations fit slightly better to the experimental energy spectra. Again the best fit is found for a somewhat larger electric field strength.

At high emission current densities, the experimental energy distributions from both emitters are broadened with respect to the calculated spectra. We have shown that it is likely that this broadening is caused by the Boersch effect.

[1]V. N. Shrednik, Sov. Phys. Solid State **3**, 1268 (1961).
[2]L. W. Swanson and L. C. Crouser, J. Appl. Phys. **40**, 4741 (1969).
[3]L. W. Swanson and A. E. Bell, *Advances in Electron. El. Physics* (Academic, New York, 1973), Vol. 32.
[4]L. W. Swanson and N. A. Martin, J. Appl. Phys. **46**, 2029 (1975).
[5]L. W. Swanson, J. Vac. Sci. Technol. **12**, 1228 (1975).
[6]A. E. Bell and L. W. Swanson, Phys. Rev. B **19**, 3353 (1979).

[7]L. W. Swanson and G. A. Schwind, in *Handbook of Charged Particle Optics*, edited J. Orloff (CRC, Boca Raton, 1997), Chap. 2, pp. 77–102.
[8]P. Zinzindohoué, Optik (Stuttgart) **74**, 131 (1986).
[9]D. C. Joy and C. S. Joy, Micron **27**, 247 (1996), part VI.
[10]H. S. Kim, M. L. Yu, E. Kratschmer, B. W. Hussey, M. G. R. Thomson, and T. H. P. Chang, J. Vac. Sci. Technol. B **13**, 2468 (1995).
[11]H. S. Kim, M. L. Yu, M. G. R. Thomson, E. Kratschmer, and T. H. P. Chang, J. Appl. Phys. **81**, 461 (1997).
[12]S. G. Christov, Phys. Status Solidi **17**, 11 (1966).
[13]FEI Company, 7425 N.W. Evergreen Parkway, Hillsboro, Oregon 97124-5830.
[14]VSW (Vacuum Science Workshop) Ltd. Unit 4, Heather Close, Lyme Green Business Park, Macclesfield, Cheshire, SK11 0LR, UK.
[15]R. D. Young and C. E. Kuyatt, Rev. Sci. Instrum. **39**, 1477 (1986).
[16]P. W. Hawkes and E. Kasper, *Principles of Electron Optics Vol. 2: Applied Geometrical Optics* (Academic, New York, 1989), Chap. 44.
[17]W. Gröbner and N. Hofreiter, *Integraltafel. II Teil: Bestimmte Integrale*, 3rd ed. (Springer, Wien, 1961), p. 83, Eq. 4(a).
[18]D. Tuggle, N. K. Kang, and L. W. Swanson, Optik (Stuttgart) **33**, 313 (1983).
[19]P. Kruit and G. H. Jansen, in *Handbook of Charged Particle Optics*, edited by J. Orloff (CRC, Boca Raton, 1997), Chap. 7, pp. 275–318.

Surface work function studies in porous silicon

Tejashree M. Bhave and S. V. Bhoraskar[a)]
Department of Physics, University of Pune, Pune 411 007, India

(Received 29 October 1997; accepted 24 April 1998)

Band structure studies in porous silicon provide useful information about the operative phenomenon responsible for its room temperature photoluminescence. We have measured the average surface work function, using a retarding field diode method in ultrahigh vacuum conditions, for porous silicon having different crystalline columnar dimensions. The average crystallite size was determined by grazing angle x-ray diffraction measurements; whereas the band gap was estimated from the photoluminescence measurements. Based on these results, different empirical band structural models are reviewed. Photoconductivity measurements in porous silicon also insinuate that the electrical resistivity of the surface of porous silicon is manipulated by the silicon complexes present on the surface. It has been concluded that, in addition to the quantum confinement, the surface molecular species dominantly control the behavior of photoluminescence and average surface work function of porous silicon. © *1998 American Vacuum Society.*
[S0734-211X(98)03404-0]

I. INTRODUCTION

On account of the visible light emission at room temperature,[1–4] an extensive research work has been carried out in the silicon nanostructures; namely, porous silicon (PS). Integration of optical components with silicon technology is the possible technological outcome of this research. In order to use this material more effectively in commercial optoelectronics; study of its fundamental electronic properties is a prerequisite.

In depth analysis has proved that the preparation parameters of PS affect the physical and optical properties to a large extent.[5–7] There are reports[5] about the blueshift in the photoluminescence (PL) peak position with the reducing crystallite size which is obtained by increasing formation current densities. Lehmann *et al.*[8] have proposed a band structure modification, associated with the blueshift, with increasing formation current densities. To explain this, they propose a mechanism in which holes are necessary for the electrochemical dissolution and they are depleted from the rest of the material where the pores are formed. This causes a reduction in the hole concentration, thereby, leading to the shift in the Fermi level position. However, PS is usually named as *n* type or *p* type according to the type of wafer of which it is formed. A direct experimental measurement is, hitherto, not carried out for determining the electronic behavior of the surface, but as to understand the band structure at the surface. Moreover, there is no report about the evidence of the movement of the Fermi level with varying crystallite size. Xue *et al.*[9] have shown from the quantum mechanical treatment to the PS that the chemical potential plays an important role in the quantum confined state. Second, in spite of the consistent efforts in theoretical and experimental research, the question whether the PL in PS is related to quantum confinement in the Si crystallites or to surface states or to various silicon compounds is still debatable.[10–12]

Following these considerations we have carried out a comprehensive measurement of the average surface work function of PS layer prepared with different anodic current densities by making use of retarding field diode method. Based on these results, we have proposed the restructured band model for the surface and interface of PS.

Grazing angle x-ray diffraction studies have helped in confining the structural measurements over the porous layer of silicon as well as facilitating the grain size determination. Measurements of dark conductivity and photoconductivity have further assisted in assuring the nature of porous layer. Results obtained with work function measurements, along with those obtained with x-ray diffraction (XRD) and photoconductivity studies indicate that quite a prominent role is played by the surface chemical species in controlling the properties of PS.

II. EXPERIMENT

Porous silicon was prepared by anodic etching of *p*-type, 285 μm thick, Si(111) wafers with resistivity of 7–12 Ω cm. Ohmic contacts were established on one of the surfaces of 10 mm$\times$10 mm square pieces of the wafer. Synthesis of PS was performed in a solution containing HF (48%) and ethanol (in 1:1 proportion); with anodic current densities ranging between 5 and 100 mA/cm^2 for 20–30 min. Graphite was used as the cathode. The resulting PS samples were of reddish brown appearance.

The PL studies were performed on Perkin–Elmer luminescence spectrometer, LS-50 B, with xenon discharge lamp as a source. The excitation wavelength was 300 nm. All the PL measurements were carried out at room temperature. The work function measurements were performed by retarding field technique (modified diode method), using a low energy electron beam collimated by an axial magnetic field. The low energy electron gun was similar to that described by Klauser and Bas.[13] Work function measurements using this method has been published in our earlier communications for variety

a)Electronic mail: svb@physics.unipune.ernet.in

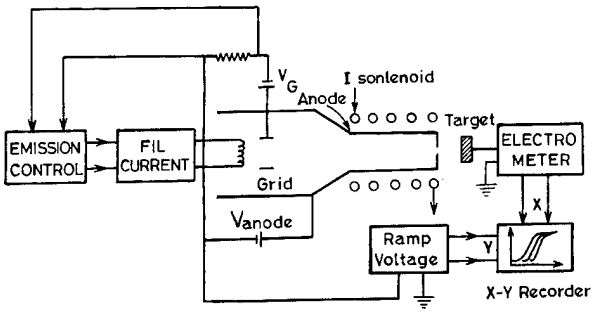


FIG. 1. Schematic diagram of the work function measurement setup. Retarding potential is applied on the sample.

of samples.[14] This method provides consistent results and is applicable for metals, semiconductors, and insulators.

The diode method, in principle, consists of a filament (the electron source) and the sample (the target). The target current of a diode operated in either the retarding field or space charge limited mode can be expressed as

$$I_t = f(\Phi_t - V_t),$$

where I_t is the target current, V_t is the retarding potential, f is a monotonic function and Φ_t is the target work function. The relationship between the target current and target voltage is given by[13]

$$I_t = AT_f^2 \exp[(V_t - \Phi_t)e/kT_f],$$

where A is a constant and T_f is the filament temperature.

The ln I_t plotted as a function of the retarding potential V_t exhibits a straight line. Such plots for samples varying in ϕ will therefore be parallel to each other. At points on any two curves equidistant from V_t axis, we have

$$I_{t1} = I_{t2}$$

and correspondingly for the two samples,

$$V_{t1} - \phi_{t1} = V_{t2} - \phi_{t2}$$

and

$$\phi_{t2} - \phi_{t1} = V_{t2} - V_{t1}.$$

If ϕ_{t1} relates to the known reference sample, ϕ_{t2} for the unknown sample can be calculated.

A schematic diagram of the experimental arrangement is shown in Fig. 1. The beam energy was kept at 10 eV with a typical beam current of the order of 10^{-10} A , measured on the target, when no retarding voltage was applied. The accuracy in measuring the work function was determined to be ± 0.01 eV. The measurements were carried out in an ion pumped ultrahigh vacuum system at pressure less than 5×10^{-9} Torr. Polycrystalline gold ($\phi = 4.9$ eV) was used as a reference sample. Fourier transform infrared spectra (FTIR) were taken in transmittance mode using NICOLET - 60 SXB model FTIR spectrometer, with 2 cm^{-1} resolution, by averaging over 500 scans. The $I-V$ characteristics were recorded by establishing planer ohmic contacts on the surface of PS using two-probe method and subsequently photoconductivity

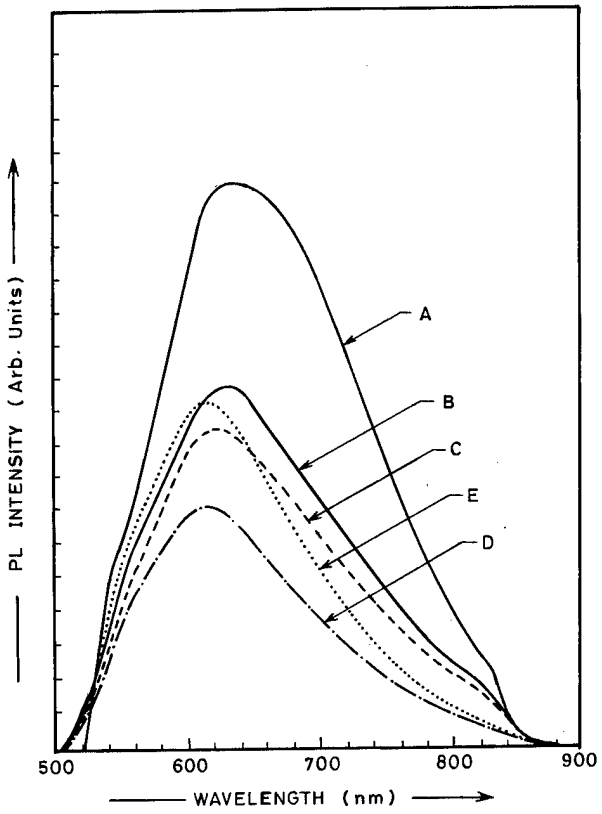


FIG. 2. Photoluminescence spectra for the PS samples prepared at various anodic current densities. (a) 10 mA/cm^{-2}, (b) 25 mA/cm^{-2}, (c) 40 mA/cm^{-2}, (d) 60 mA/cm^{-2}, (e) 100 mA/cm^{-2}.

was measured by illuminating the surface with a 500 W sungun. Grazing angle x-ray diffraction (XRD) measurements were carried out using Rigaku (Rotalex RV 200B) at a grazing angle of 0.5°.

III. RESULTS AND DISCUSSION

Samples of PS were prepared at different current densities, varying between 5 and 100 mA/cm^2. They showed broad room temperature PL, centered in the red region, when excited by xenon arc discharge lamp with an excitation wavelength of 300 nm (Fig. 2). A blueshift in the PL peak position, along with an increasing anodic current density, was observed in the PL peak position; as can be inferred from Fig. 2. It is observed that the blueshift is smaller for higher current densities as also the variation in crystallite size which is evidenced by x-ray diffraction results.

Structural investigation was carried out using grazing angle x-ray diffraction studies for glancing angle of 0.5°; which corresponds to a depth of 1.243 μm from the surface in silicon. The inset of Fig. 3 is a typical XRD pattern for PS, prepared at an anodic current density of 25 mA/cm^2, which shows a prominent peak at $2\theta \approx 28°$; representing the (111) plane in silicon. The average grain size (crystallite size or the average column diameter) was determined using the Scherrer formula.[15] The crystallite size deduced from these measurements showed a variation from 11.38 to 4.8 nm, as the current density was varied from 5 to 100 mA/cm^2 (Fig. 3). It

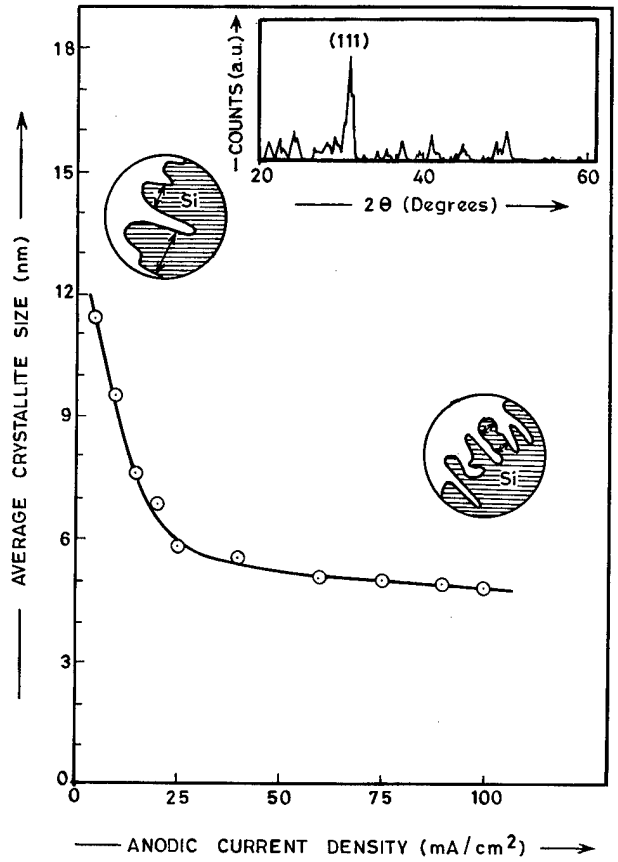


FIG. 3. Variation of average crystallite sizes deduced from grazing angle x-ray diffraction pattern with anodic current density. Inset shows a typical x-ray diffraction spectrum of PS recorded at a glancing angle of 0.5° which exhibits peak at $2\theta \approx 28°$ correlating to (111) planes.

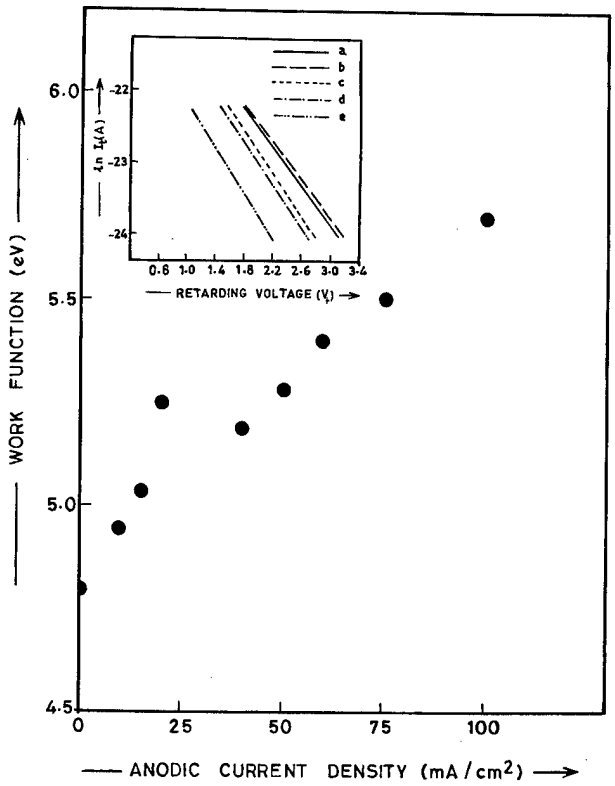


FIG. 4. Variation of work function with anodic current density determined from retarding field measurements. Inset shows the variation of ln I_t with the retarding voltage for (a) Gold (reference sample) and some PS samples prepared at anodic current densities, (b) 10 mA/cm^2, (c) 40 mA/cm^2, (d) 50 mA/cm^2, (e) 100 mA/cm^2.

can be clearly seen from the plot that reduction in the average crystallite size with the anodization current density is not a linear phenomenon. The rate of reduction of the average crystallite size has slowed down at higher current densities. This could be explained as a consequence of a diffusion limited process involved in the formation of PS using electrochemical etching.[16] The decrease in the average crystallite size also indicates an increase in porosity of the layer as is evident in previous reports by Yon *et al.*[17] who have estimated the porosity by gravimetric measurements.

Relative values of the surface work function for the samples of PS, prepared with different current densities, were estimated by retarding field diode measurements in ultrahigh vacuum conditions. Relative assessment of this parameter is emphasized because PS is not a pure single crystal with a unique band gap and there is a size distribution of nanocrystallites. Moreover, this material is constituted of more than one phase, i.e., microcrystalline and amorphous silicon. The electron beam diameter is approximately 1 mm. This would cover around 10^4 crystallites which is large enough for one to safely assume that the value of surface work function so obtained from our experiment is fairly an average value for the surface. The variation in the sample current with the retarding field for porous silicon samples, prepared with different anodic current densities, was used to measure the values of work function ϕ. Inset of Fig. 4 shows a plot of ln I_t as a function of retarding potential V_t for few samples of porous silicon, plotted along with that for gold chosen as a reference. The calculated values of work function are plotted as a function of the anodic current density in Fig. 4.

The effects are discussed in view of the existing models in literature with porous silicon. The variation in the work function ϕ can be thought to arise from three different reasons.

(1) Anodic etching of crystalline silicon during the preparation of PS involves hole dissolution. On account of this, the remaining material becomes less p type, causing the Fermi level to move towards the midgap. The value of the average work function should then show a decrease as shown in Fig. 5(a).

(2) As a consequence of quantum confinement of the electronic wave function, the band structure should get modified and consequently the band gap will increase. This is reflected as a blueshift in the peak position of the PL. Subsequently, the work function would show an increase as shown in Fig. 5(b).

(3) The crystalline silicon structure in the columnar morphology, in most probable cases, is covered by a complex silicon derivative containing Si–H_x, O–Si–H, and Si–O_x kinds of bonds. This would change the surface band structure and a totally new value of the work function may appear at the surface which may be different than the bulk. The work

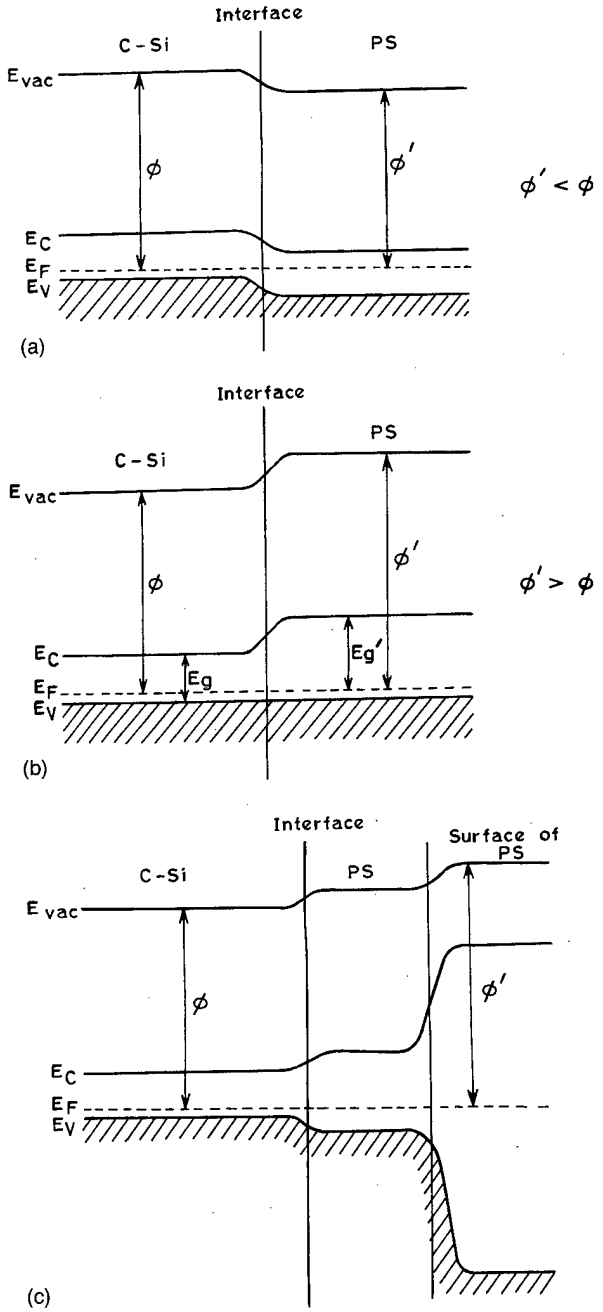


FIG. 5. (a) Schematic of the proposed band diagram for the interface of PS with crystalline silicon as a result of hole dissolution. (b) Schematic of the proposed band diagram for the interface of PS with crystalline silicon as a result of quantum confinement. (c) Schematic of proposed band diagram for PS as a result of hole dissolution, quantum confinement, and presence of surface species.

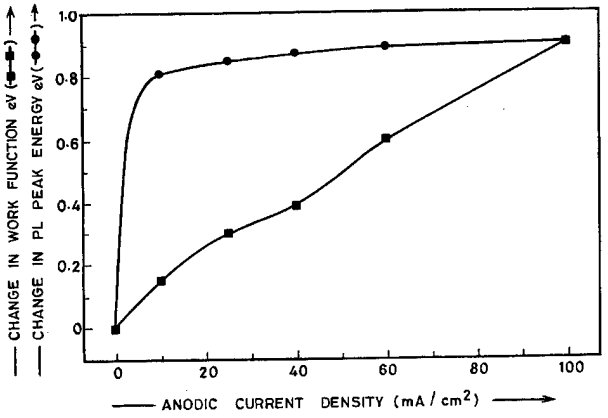


FIG. 6. Plot of differential work function and change in the PL peak energy with anodic current density for PS.

function would alter, either due to the different electron affinity of the surface layer, or due to the presence of a dipolar space charge layer at the surface which causes the bands to bend. Figure 5(c) shows the combined effect of the above three assumptions.

Each of these effects would certainly be affected by the anodic current density, since the pore density and the column dimension of nanocrystallite is decided by the current density. If the variation in ϕ is solely due to quantum confinement, the results should be in close agreement with the variation of blueshift in PL. However, Fig. 6 shows that the rate of change of the work function, with the anodic current, is larger as compared to that of the PL peak energy derived from the PL measurements. In fact, actual PL peak energies will be lower than the band gap values by the exciton binding and localization energies.[18] On the other hand, if the hole dissolution effects are accounted, ϕ should have shown a decrease. The observed variation in ϕ therefore seems to be dominated by the molecular species having different chemical structure than that of the bulk, in addition to the first two effects, as shown in Fig. 5(c). Figure 5(c) therefore depicts the proposed band structure of the surface of PS derived from the observed results.

Presence of different silicon complexes is also supported by the infrared measurements. Table I presents the data obtained from the FTIR spectrum of PS which confirms the presence of Si–H, O–Si–H, and Si–O–Si complexes.[19] In fact we have shown in our previous publication that presence of O–Si–H kinds of bonds on the surface of PS is favorable for improving the luminescence efficiency and its stability.[20]

It is worth citing here the references where in simple quantum confinement model alone has not been able to fit into the observed results. Kanemitsu *et al.*[21] have compared the size dependent luminescence peak energy in surface oxidized Si nanocrystals with the theoretically derived band gap energy. The observed blueshift is negligible in comparison to the theoretical variation. Andersen and Veje[22] have con-

TABLE I. Data obtained from FTIR spectrum of PS.

Wave number (cm^{-1})	Assigned bonds
460	Si–O–Si bend
640	Si–H bend
835	$Si–H_2$ wagging
870	$Si–H_2$ scissors
1010–1160	Si–O–Si stretch
2100	Si–H stretch
3420–3580	O–H stretch

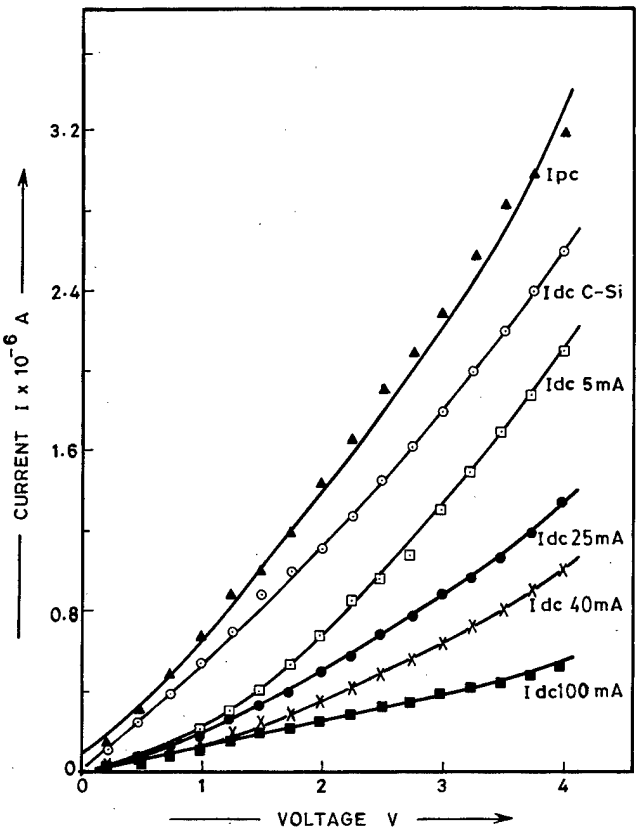


FIG. 7. Variation of dark conductivity of PS samples prepared with different anodic current densities.

cluded that the observed luminescence from their PS samples, prepared at various current densities, is of molecular nature.

Additional information about the surface layer was supplemented by the electrical resistivity and photoconductivity measurements which was carried out by recording the planer $I-V$ characteristics. Proper precautions were taken to ensure that the contacts were perfectly ohmic in nature and not rectifying, since in analyzing the photocurrent mechanisms in any semiconducting structure, it is important to distinguish between the contact (including depletion layer) and volume effects, for example, diffusive transport. More so, especially in PS, this is of crucial importance because the nature and extent of contact is unknown. Ohmic contacts were therefore confirmed by varying the geometrical parameters like separation between the electrodes as well as by varying the contact area. Moreover, the steady state conditions were ensured for eliminating the diffusion behavior. The surface resistivity was seen to increase as seen in Fig. 7, for the PS samples prepared by variation of anodic current from 5 to 100 mA/cm^2. The photoconductivity remained approximately identical in all the PS samples prepared with different current densities. The behavior of photoconductivity was also identical to that obtained with a virgin crystalline Si sample. The gain in photocurrent, defined by Ipc/Idc, is seen to increase with the increasing anodic current density, i.e. (decreasing crystallite size). This means that the photosensitivity increases with the decreasing crystallite dimensions. This effect is correlated to the decrease in the dark current alone, which is quite obvious, since the planer surface resistivity of porous silicon is mainly governed by the surface layer covering the crystallite columns of silicon. As has already been inferred from our work function measurements, the surface layer consists of complex silicon derivatives. The surface resistivity of this film ought to be high compared to the crystalline silicon, but lower than hydrogenated amorphous silicon a-Si:H.[23] Thus we may conclude that the surface layer is "intermediate" in properties and structure between a crystalline system and an a-Si:H like system. Similar conclusions have also been drawn by Burstein *et al.*[24] from their surface photovoltage spectroscopy. The constant photocurrent can be explained on the basis of the absorption coefficient of PS. It has been reported[25] that the absorption coefficient of PS is very small ($\sim 10^3$ cm^{-1} at 500 nm), on account of which even at 500 nm, the penetration depth of light is over 10 μm. The average thickness of the layer of PS was measured using the scanning electron microscopic imaging and was found to be 5–10 μm. If we assume the nature of PS to be consisting of filamentary type of arrays, the thickness of the porous layer and therefore the columnar length was observed to be less affected by the current density. Consequently, the photoconductivity observed in PS is mainly a contribution from the crystalline bulk which remains unaffected.

IV. CONCLUSION

Summarizing, we find that the increase in the work function with the decreasing crystallite size is much larger compared to what is expected if the blueshift in the band gap or the extent of hole dissolution is accounted. In conclusion, we have found that the surface electrical and electronic properties of PS are largely governed by the complex molecular structure which covers the nanocolumnar arrays and pore walls. Our previous observations,[19] in regards to the enhanced PL in ion-irradiated porous silicon, had also revealed that the molecular complexes consisting of O–Si–H, Si–O–Si, and Si–H kind of species are important in controlling the PL behavior of PS.

ACKNOWLEDGMENTS

One of the authors (T.M.B.) wishes to acknowledge Council for Scientific and Industrial Research (CSIR) for financial assistance. Thanks are due to Dr. S. G. Hegde, National Chemical Laboratory, Pashan 411 008, for his help in recording the FTIR spectra of PS.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]L. Brus, Nature (London) **353**, 301 (1991).
[3]R. W. Fathauer, T. George, A. Ksendzov, and R. P. Vasquez, Appl. Phys. Lett. **60**, 995 (1992).
[4]N. Noguchi and I. Suemune, Appl. Phys. Lett. **62**, 1492 (1993).
[5]S. M. Prokes, W. E. Carlos, and V. M. Bermudez, Appl. Phys. Lett. **61**, 4447 (1992).
[6]M. Thonissen, S. Billat, M. Kruger, H. Luth, M. Berger, U. Frotschen, and U. Rossow, J. Appl. Phys. **85**, 2990 (1996).

[7]R. T. Collins, P. M. Fauchet, and M. A. Tischler, Phys. Today **50**, 24 (1997).
[8]V. Lehmann, B. Jobst, T. Muschik, A. Kux, and V. Petrova-Koch, Jpn. J. Appl. Phys., Part 1 **32**, 2095 (1993).
[9]F. Xue, X. Bao, and F. Yan, J. Appl. Phys. **81**, 3175 (1997).
[10]X. Wang, D. Huang, L. Ye, M. Yang, P. Hao, H. Fu, and X. Y. Hou, Phys. Rev. Lett. **71**, 1265 (1993).
[11]M. Stutzmann, M. S. Brandt, M. Rosenbauer, J. Weber, and H. D. Fuchs, Phys. Rev. B **47**, 4806 (1993).
[12]P. N. Saeta and A. C. Gallagher, Phys. Rev. B **55**, 4563 (1997).
[13]S. J. Klauser and E. B. Bas, J. Phys. E **12**, 841 (1979).
[14]S. B. Dake, N. R. Rajopadhye, and S. V. Bhoraskar, J. Phys. D **20**, 1631 (1987); N. R. Rajopadhye and S. V. Bhoraskar, J. Non-Cryst. Solids **105**, 179 (1988); A. H. Bhuiyan, N. R. Rajopadhye, and S. V. Bhoraskar, Thin Solid Films **161**, 187 (1988).
[15]B. D. Cullity, *Elements of X-ray Diffraction* (Addison–Wesley, Reading, MA, 1956), p. 99.
[16]R. L. Smith, S. F. Chuang, and S. D. Collins, J. Electron. Mater. **17**, 533 (1988).
[17]J. J. Yon, K. Barla, R. Herino, and G. Bomchill, J. Appl. Phys. **85**, 2990 (1987).
[18]A. J. Read, R. J. Needs, K. J. Nash, L. T. Canham, P. D. J. Calcott, and A. Qteish, Phys. Rev. Lett. **69**, 1232 (1992).
[19]M. S. Brandt, H. D. Fuchs, M. Stutzmann, J. Weber, and M. Cardona, Solid State Commun. **81**, 307 (1992).
[20]T. M. Bhave, S. V. Bhoraskar, S. Kulkarni, and V. N. Bhoraskar, J. Phys. D **29**, 462 (1996).
[21]Y. Kanemitsu, H. Uto, Y. Masumoto, T. Matsumoto, T. Futagi, and H. Mimura, Phys. Rev. B **48**, 2827 (1993).
[22]O. K. Andersen and E. Veje, Phys. Rev. B **53**, 15643 (1996).
[23]S. V. Bhoraskar, T. Bhave, and T. A. Railkar, Bull. Mater. Sci. **17**, 523 (1994).
[24]L. Burstein, Y. Shapira, J. Partee, J. Shinar, Y. Lubianiker, and I. Balberg, Phys. Rev. B **55**, R1930 (1997).
[25]J. P. Zheng, K. L. Jiao, H. P. Shen, W. A. Anderson, and H. S. Kwok, Appl. Phys. Lett. **64**, 459 (1992).

Computer-controlled fabrication of ultra-sharp tungsten tips

Young-Guon Kim, Eun-Ha Choi, Seung-Oun Kang, and Guangsup Cho[a)]
Charged Particle Beam and Plasma Laboratory, Department of Electrophysics, Kwangwoon University, Seoul, Korea 139-701

(Received 27 August 1997; accepted 15 May 1998)

In order to monitor the end-point current during electrochemical etching, an analog to digital converter circuit aided by a personal computer has been setup. At the moment the lower part of the needle drops off during the etching process, a maximum current change across the reference resistor is detected by the PC interface card and the applied voltage is then cut off within a few milliseconds. Using these circuits to control the etching time, our experiment has been able to fabricate an ultra-sharp tip of ~200 Å radius with a higher reproduction rate and reliability than the conventional method. © *1998 American Vacuum Society.* [S0734-211X(98)07004-8]

Many ways to fabricate ultra-sharp tips of various materials for field emission have been devised such as electrochemical etching,[1] ion milling,[2] chemical etching,[3] and cathode sputtering.[4] These techniques, however, can be used only for the preparation procedure, and an auxiliary step is needed to obtain an ultra-sharp tip and optimum aspect ratio near the apex region, both of which are controlled by microzone electropolishing.[5] Recently, it has been reported that the oxidation process[6] can be used to shape a tip radius accurately. In this brief communication, the fabrication of an ultra-sharp tip for the preparation of a field emission electron source has been studied as a one-step procedure. The etching power supply is computer-controlled together with the analog to digital converter (ADC) in the interface card. The etching current is monitored during the etching process, and the end-point etching time is determined by cutting off the power supply when the change of current across the reference resistor has a maximum value, which occurs at the moment the lower part of the needle drops off. This technique is primarily intended to control both the tip shape structure and the apex radius with good accuracy and reproducibility. Characteristics of tips prepared with this technique were evaluated by field ion microscopy of the tip geometric structure, current–voltage (I–V) characteristics, emission stability, and Fowler–Nordheim (FN)[7] plots from field emitted electron currents. We set up the electrochemical etching system so that the etching voltage can be automatically cutoff by a computer system when the lower part of the needle is etched to drop off. Figure 1 is a schematic diagram of etching circuit. The power supply unit is controlled by the computer to provide a stable and accurate voltage. In this circuit, we can measure the etching voltage applied to the tip electrode with 0.5 mV resolution and 0.1 ms period from two channels of the ADC. Figure 1 schematically depicts a tungsten electrode of 0.15 mm diameter and ring-type counterelectrode, which is made of 1 mm diameter copper wire with a ring diameter of 10 mm. After applying voltage to the tungsten electrode which is dipped 3 mm into electrolyte, we can see the etching current in the reference resistor R_f. The shaping of the etched tungsten during the whole etching process through an optical microscope with 60× magnification objective lens. Figure 2 shows the etching process of drop-off. The specimen near the air-liquid interface region is etched faster than any other submerged region. When the lower section of the needle drops off, we observe the maximum change in etching current across the reference resistor R_f. At that moment the voltage is switched off nearly instantaneously, in order to avoid further etching or blunting of the tip. After etching is finished the tip is rinsed in distilled water by using an ultrasonic cleaner. With these circuits the power supply has been controlled by the computer to cut off the voltage automatically with a time delay between drop-off and switch-off of less than 0.1 ms. In this experiment, NaOH electrolyte solution is employed, and the concentration of electrolyte solutions is varied from 1 to 3 N.

Experiments have been performed to measure the geometric structure and field emission properties of tips fabricated by the computer-controlled circuits. Figure 3(a) shows the plots of taper length versus applied voltages under NaOH electrolyte concentrations of 1, 2, and 3 N. Tip photographs for applied voltages of 2, 5, and 8 VDC at a fixed concentration of 2 N NaOH, and those for NaOH concentrations of 1, 2, and 3 N at a fixed etching voltage of 5 VDC are shown in Figs. 3(b) and 3(c), respectively. In Fig. 3, we find that taper length is shortened as either the applied voltage or the NaOH concentration is increased. The taper length is varied from 250 to 170 μm as the voltage increases from 2 to 8 VDC for fixed 2 N NaOH electrolyte [Figs. 3(a)–3(b)]. The

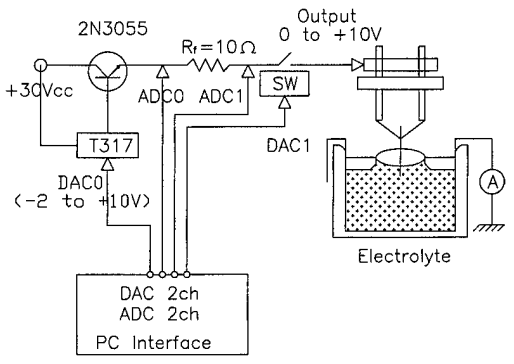


FIG. 1. Schematic of the etching circuit.

[a)]Electronic mail: gscho@litholab1.kwangwoon.ac.kr

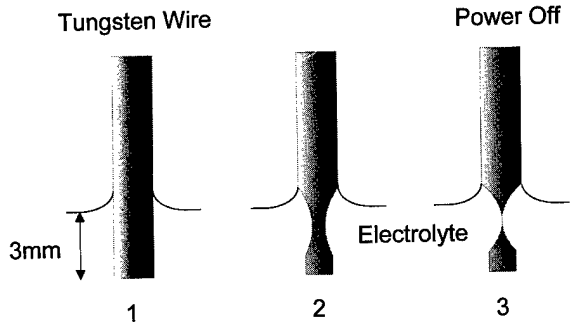


FIG. 2. Etching process of drop-off.

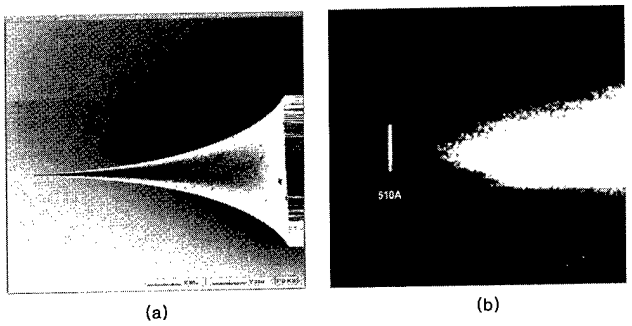


FIG. 4. Image of a tungsten tip from scanning ion microscopy.

taper lengths with fixed 5 VDC are 220, 200, and 190 μm for 1, 2, and 3 N NaOH concentrations, respectively [Fig. 3(c)]. For the NaOH-2N concentration total etching time required for needle drop-off is measured to be 364.1, 340.9, and 308.2 s for the applied voltage 2, 5, and 8 V, respectively. From these data the etching rate can be estimated radially to be about 0.2 μm/s since the etching time is about 350 s for half of the radius. Figure 4 shows a tungsten tip image obtained from scanning ion microscopy. It is shown that a smooth surface and well-formed cylindrical symmetry in the etched tip are achieved as shown in Fig. 4(a). The apex tip image is shown in Fig. 4(b), where the radius of curvature of the tip is estimated and shown to be about 200 Å. Even using various NaOH concentrations and several etching voltages, we have obtained field emission electrodes of ultra-sharp tip apex radius of about 200 Å by controlling the etching drop-off time with ms accuracy. In order to investigate the emission characteristics of an ultra-sharp tip, a simple experimental system is set up as shown in Fig. 5. The distance from cathode to anode is 0.5 mm, and the anode hole size is 0.13 mm. Also, a Faraday cup is located 10 mm downstream from the anode. We measured the beam current by Faraday cup, and recorded the emission current in the PC using the ADC. Figure 6 shows the emission stability of the field-emitted electron beam at 2×10^{-5} Torr, which appears reasonably stable during 12 000 s of operation. The beam stability, $\Delta I/I$, is found to be less than 5% of the ~5 μA total emission current at an extraction voltage of 700 V. The current–voltage curve of the field-emitted electron beam is shown in Fig. 7. The ignition voltage is about 300 V with an emission current of ~10 nA, and the current increases to 9 μA with extraction voltage 1.4 kV, as shown in this figure. With $I-V$ measurements near the ignition voltage, Fowler–Nordheim (FN) plots can be drawn. The effective radius, r_{eff}, of the tip can be derived from the FN formula[7]

$$\frac{I}{V^2}=a\exp\left(-\frac{b\phi^{3/2}}{V}\right), \tag{1}$$

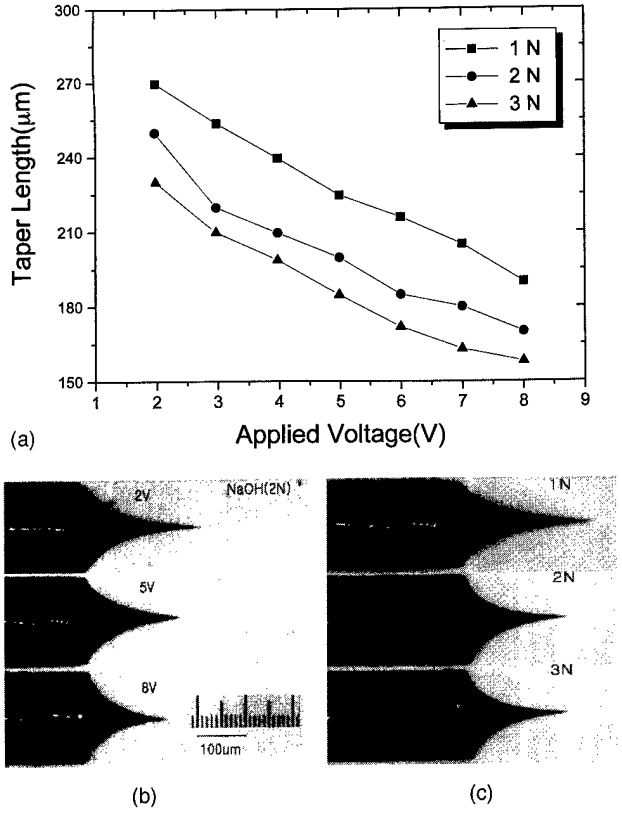


FIG. 3. (a) Plots of taper length vs applied voltages under several concentrations of NaOH electrolyte of 1, 2, and 3 N. (b) Photographs of tips etched at the applied voltages of 2, 5, 8, VDC for fixed concentration of 2 N NaOH. (c) Photographs of tips etched at the NaOH concentrations of 1, 2, and 3 N for fixed etching voltage of 5 VDC.

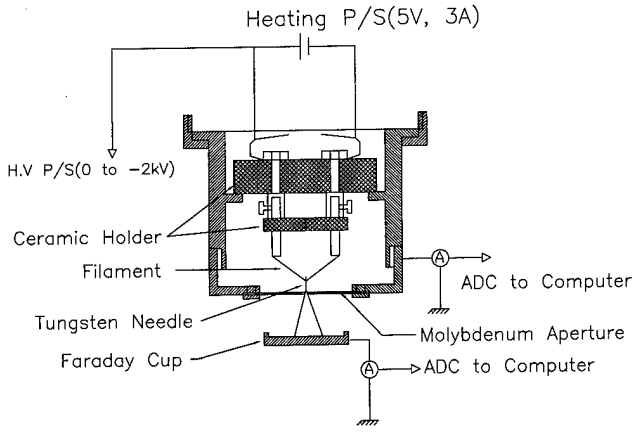


FIG. 5. Experimental system setup for investigating emission characteristics of an ultra-sharp tip.

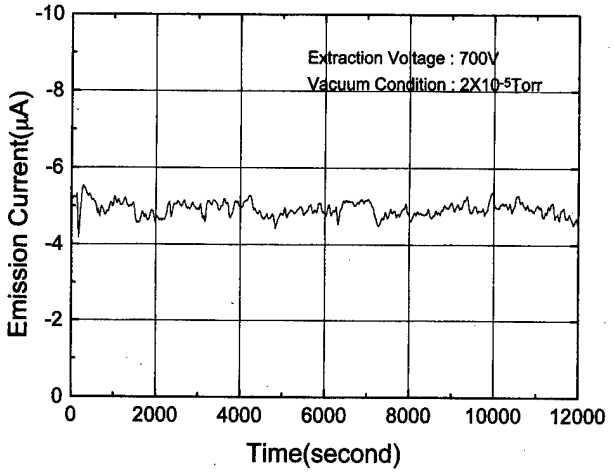


FIG. 6. Electron beam stability emission current as a function of time.

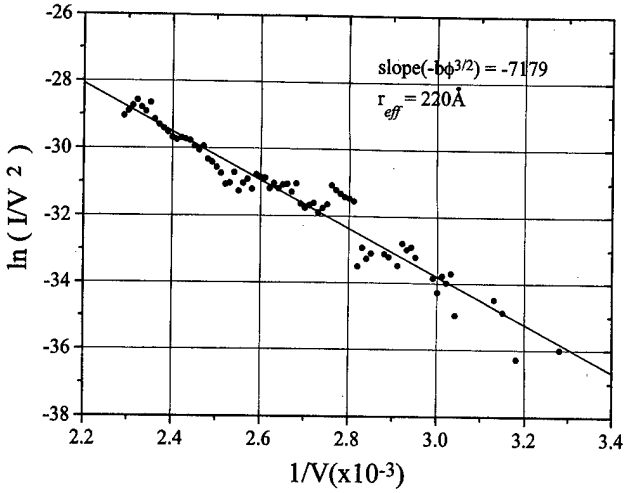


FIG. 8. Fowler–Nordheim plot of data shown in Fig. 7. An effective tip radius of ~200 Å is calculated from the slope of the line.

where $a=6.2\times10^6A(\mu/\phi)^{(1/2)}(\mu+\phi)^{-1}+(\alpha kr_{\text{eff}})^2$, $b=6.8\times10^7\alpha kr_{\text{eff}}$, in which α is the Nordheim image correction term ($\simeq1$), ϕ is the work function of the tungsten tip (~4.5 eV), k is the form factor which depends on tip-anode geometry and emitter shape ($k\simeq5$ for a large separation between the tip and anode), μ is Fermi level, I is the total emission current, V is the extraction voltage, and A is the total emitting area. A FN plot of $\ln(I/V^2)$ versus $1/V$ is shown in Fig. 8, where the slope, $-b\phi^{3/2}$, of a least squares fitting line is estimated to be about -7179. From this value, the effective radius r_{eff} has been measured to be ~200 Å. It is noted that this result is remarkably in good agreement with that of image size from the scanning ion microscopy. In summary, an ultra-sharp tungsten electrode can be fabricated with good reproducibility and reliability by computer controlling the applied etching voltage. The etching process can be terminated within the time delay between tip drop-off and switching-off of less than 0.1 ms. Electrolyte employed in this experiment is usually NaOH, in which taper lengths of the tip are changed according to either the applied voltages or concentrations of an electrolyte. The etching rate, the taper geometry and length of the tip are mainly influenced by the applied voltages and concentrations of the electrolyte, while the apex structure and tip radius are affected by delay time between tip drop-off and switching-off etching voltage, rather than the electrolyte. It is worth while mentioning that we can make an ultra-sharp tip with good reproducibility in a one-step etching process using computer-controlled circuits. The electron beam stability at 2×10^{-5} Torr was measured to be within 5% for total emission current of 5 μA during 12 000 s of operation. The ignition voltage was found to be about ~300 V. In the FN plots obtained from $I-V$ emission characteristics, the effective tip radius is found to be ~200 Å, which is in good agreement with the direct image size from a scanning ion microscope. Also, the etching rate is estimated to be 0.2 μm/s in this experiment. Here the drop-off etching time is computer-controlled to obtain an ultra-sharp tip with smoothing surface and well-defined cylindrical symmetry, by monitoring the abrupt change in etching current across the reference resistor. As soon as the lower part of the needle falls into the electrolyte, the etching voltage is cutoff by computer control to prevent further etching. Thus an ultra-sharp apex of radius ~200 Å is successfully fabricated throughout the experiments, with good reproducibility. Electrolyte concentration and applied etching voltage influence the taper geometry and length.

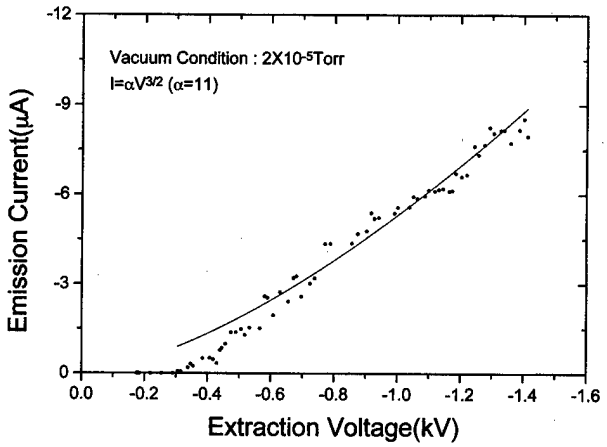


FIG. 7. Characteristics of current–voltage (I–V) measured emission currents at given extraction voltages.

This work is partially supported by the Korean Ministry of Education through the Basic Science Research Institute Program (BSRI-96-2451).

[1]A. J. Melmed, J. Vac. Sci. Technol. B **9**, 601 (1991).
[2]M. G. Burke, D. D. Sieloff, and S. S. Brenner, J. Phys. (France) **47**, 459 (1986).
[3]*Thin Film Process*, edited by J. L. Vosson and V. Kern (Academic, New York, 1978).
[4]E. W. Muller, Z. Phys. **106**, 132 (1937).
[5]R. J. Morgan, Rev. Sci. Instrum. **44**, 808 (1967).
[6]H. S. Kim, M. L. Yu, U. Staufer, L. P. Muray, D. P. Kern, and T. H. P. Chang, J. Vac. Sci. Technol. B **11**, 2327 (1993).
[7]R. Gommer, *Field Emission and Field Ionization* (Harvard University Press, Cambridge, 1961).

Electrophysical characteristics and low-energy cathodoluminescence of vacuum fluorescent display and field emission display screens

S. A. Bukesov,[a] N. V. Nikishin, A. O. Dmitrienko, S. L. Shmakov, and J. M. Kim[b]
Saratov State University, 83 Astrakhanskaya Str, Saratov 410071, Russia

(Received 19 August 1997; accepted 1 June 1998)

Special methods are proposed to estimate the conductivity of a thin phosphor layer deposited onto the screen of vacuum fluorescent display and field emission display. Some phosphor compositions have been studied with the aid of these methods. The conductivity of screens has been found to be subject to the voltage applied to a display and to alter significantly under the influence of electron bombardment. Electron-induced conduction is considered as the main cause of the observed conductivity. A correlation between some features of cathodoluminescence and electron-induced conduction has been revealed. Recommendations on how to calculate the optimal mode of display operation are given. © *1998 American Vacuum Society.* [S0734-211X(98)13604-1]

I. INTRODUCTION

There are a number of special requirements to be specified for the low-voltage cathodoluminescent screens of vacuum fluorescent displays (VFDs) and field emission displays (FEDs), which operate in the excitation voltage range below 1 kV. The most important of these requirements concerns the necessity to provide charge removal from the screen through the phosphor layer, as the secondary emission coefficient is much less than 1 at low voltages.[1] The problem of improving the conduction of phosphor compositions deposited on such screens is therefore of practical significance.

Of all the commercially available VFD and FED phosphors, only ZnO:Zn (bluish–green) possesses a rather high conductivity (10^{-6}–10^{-3} S/cm),[2] but its application is very restricted by its chromaticity coordinates. The semiconductor phosphors used in multicolor screens of modern displays have low conductivities (10^{-14}–10^{-8} S/cm) and need a conductive (nonluminescent) additive that, unfortunately, reduces the emitting fraction of the display screen.[3]

Usually, the conductivity of luminescent compositions is estimated by measuring the resistance of polycrystalline samples in vacuum. However, such values cannot be assumed to be equivalent to the screen of an actual display, as a consequence of processing steps (encapsulation, potting, pumping, getter sputtering, etc.), which follow deposition, and which most likely affect the electrophysical and luminescent properties of the luminescent coating.[4]

In order to determine the optimum conductive additive content in phosphor compositions to optimize the performance of displays, we have worked out three methods for estimating the voltage drop across the phosphor layer. The conductivity of these layers has been estimated in actual VFDs and pilot samples of FEDs made by the Samsung Advanced Institute of Technology (Korea) and R&D Volga Institute (Russia).

[a]Electronic mail: dmi3enko@scnit.saratov.su
[b]Also with: Samsung Advanced Institute of Technology, Suwon, Korea.

II. EXPERIMENT

The methods proposed are based on measuring the dynamic parameters of an equivalent circuit of the VFD (FED), where the phosphor layer is represented by an extra resistor R_L in the anode subcircuit. This seems correct as the thickness of the layer (about 10 μm) is much less than the anode plate-to-cathode distance and must not affect the field distribution in the display. Independent methods were employed to estimate the voltage drop across the phosphor layer.

A. First method

In a vacuum triode, due to the linearity of the equations describing the field distribution, the electron trajectories are unambiguously determined by the ratio of the anode and grid voltages: U_{as}/U_g, U_{as} being the anode surface-to-cathode voltage, U_g the grid voltage. The current distribution ratio $k=J_a/J_g$ (J_a being the anode current density, J_g the grid current one) is therefore an increasing function of one variable, U_{as}/U_g, not of two separate ones, U_{as} and U_g. This is important as it enables one to extract valuable information from volt–ampere curves.

At a sufficiently low current density, the voltage drop $\Delta U=U_a-U_{as}$ across R_L is negligible, i.e., the anode surface and the background have the same potential. k measured under such conditions was accepted as the initial value k_0.

As the anode current density J_a is raised (at constant U_a and U_g), ΔU increases, U_{as} decreases, and k decreases. After every increment of J_a we tried to restore the initial value k_0 by means of increasing U_a, to compensate ΔU. This increment of anode voltage must be equal to ΔU at the corresponding J_a. Volt–ampere curves were plotted.

B. Second method

The J_a versus U_g dependence is known to go through a maximum where $U_g=U_{as}$. However, as $\Delta U\neq 0$, U_a will differ from U_g by this value which can therefore be calculated by simple subtraction. J_a versus U_g curves were recorded at several anode voltages (typical ones are presented in Fig. 1).

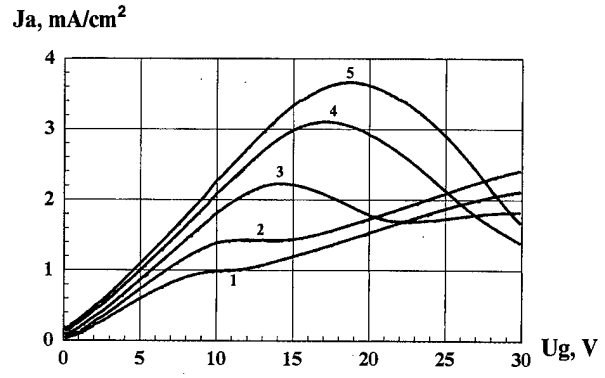


FIG. 1. Anode current as a function of grid voltage at several voltages applied to the anode: 1–10 V; 2–12 V; 3–14 V; 4–18 V; 5–20 V.

C. Third method

This method requires manufacturing special displays (either triode or diode) with both phosphor-covered and bare metal segments. Their volt–ampere curves were recorded. (Figure 2 shows typical ones for ZnS:Cu,Al with various In_2O_3 content as the conductive additive.) The presence of the phosphor layer causes a shift towards higher voltages, equal to ΔU.

The described methods enable the voltage drop across the phosphor layer to be determined at a given current density. We estimated the conductivity of the phosphor layer according to Ohm's law.

The method of choice is generally governed by the design of a given display. The third method implies making special displays, which may prove inconvenient. In spite of the methods giving estimations only, they provide sufficient accuracy and good reproducibility.

III. RESULTS AND DISCUSSIONS

A number of VFDs with a green ZnS:Cu,Al-based phosphor with varying In_2O_3 content were studied to select the optimal content of the additive. Figures 3 and 4 show the dependences of the phosphor layer conductivity on the current density and anode voltage. Each of them has a minimum.

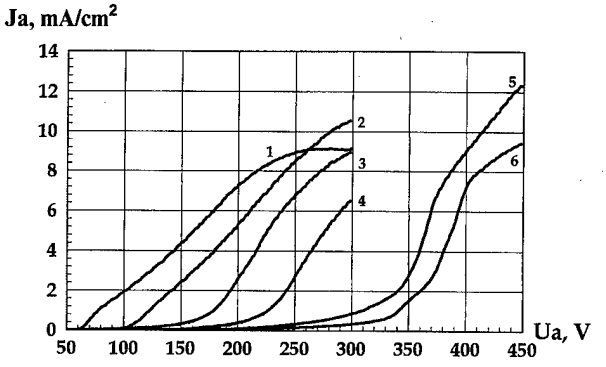


FIG. 2. Volt-ampere characteristics of cathodoluminescent screens with: 1–bare metal segments and ZnS:Cu, Al phosphor with various contents of In_2O_3: 2–5 wt %; 3–10 wt %; 4–15 wt %; 5–25 wt %; 6–30 wt %.

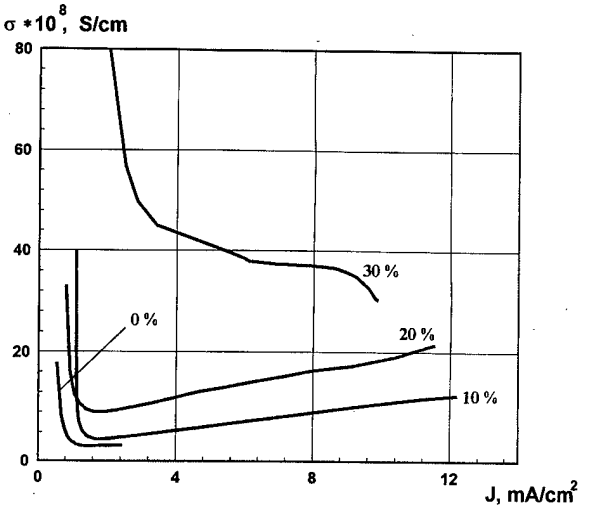


FIG. 3. Conductivity of phosphor layer of cathodoluminescent screen covered with ZnS:Cu,Al phosphor with different wt % of In_2O_3 as a function of anode current.

As is seen from these figures, the position of this minimum is practically independent of current density and subject to anode voltage only. The maximum voltage drop across the phosphor layer (70%–80% of the total anode voltage) was observed for the sample without the additive.

The initial decrease in conductivity (Fig. 4) is related to the small concentration of free carriers in the matrix. As the energy of the bombarding electrons rises, the conductivity increases. This effect is obviously due to electron-stimulated conduction in the matrix, as ZnS, like the majority of phosphors, exhibits good cathodoluminescence. The location of the minimum, at which bombardment-induced charge carriers start to appear, depends on the anode voltage (Fig. 5) and correlates well with the cathodoluminescence threshold, also shown there.

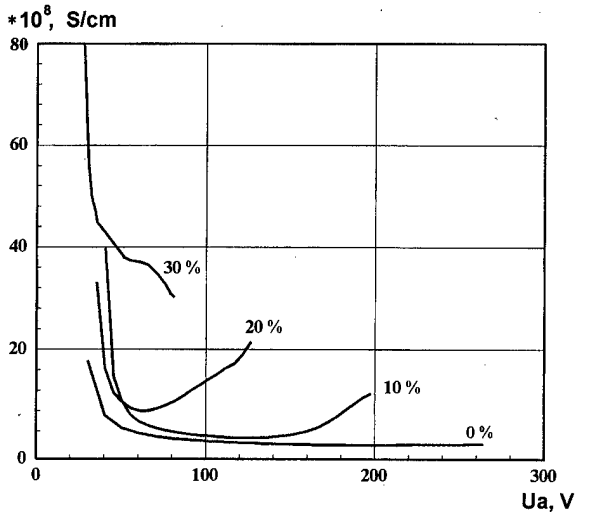


FIG. 4. Conductivity of phosphor layer of cathodoluminescent screen covered with ZnS:Cu,Al phosphor with different wt % of In_2O_3 as a function of anode voltage.

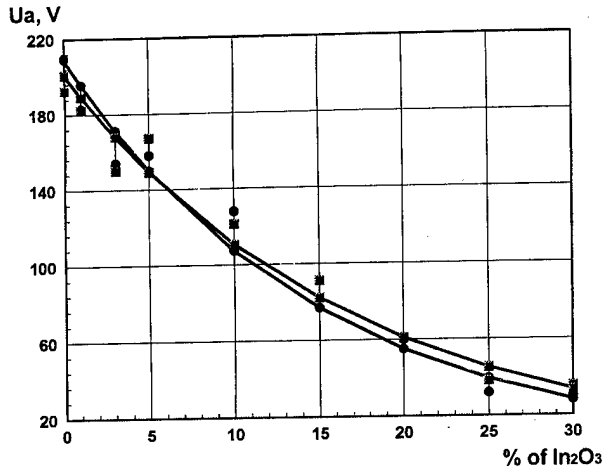


FIG. 5. Influence of the In_2O_3 content on the luminance threshold (●) and on the anode voltage at which the conductivity of the layer is minimal (■) for ZnS:Cu,Al phosphor with different wt % of In_2O_3.

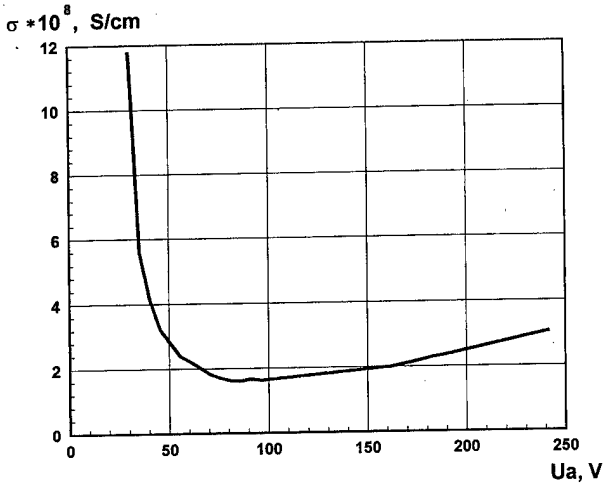


FIG. 7. Conductivity of phosphor layer of cathodoluminescent screen covered with Y_2O_2S:Eu phosphor with 10% In_2O_3 vs anode voltage.

This correlation confirms our supposition of electron-stimulated conduction of the phosphor itself, although such an effect in In_2O_3 should also be taken into account. Owing to the similarity between cathodoluminescence and induced conduction, we suppose that the rate of rise of the electron-induced conductivity must correlate with the efficiency of cathodoluminescence.

The described behavior is characteristic not only of sulphide phosphors but also of Y_2O_2S:Eu based phosphors as well (Figs. 6 and 7).

The results obtained were used to optimize the performance of displays. Figures 8 and 9 present brightness–power curves and the power distribution under several anode voltages.

The power distribution is seen to be a strong function of the conductivity of the phosphor composition. For example, in the case of the ZnS:Cu,Al composition with a small amount of In_2O_3 at low anode voltages and high current densities, a significant part of the power supplied is converted into heat in the phosphor layer, leading to a decrease in the screen brightness due to temperature quenching and to fast degradation of the screen.

The use of the additive at the optimal content (20%–25% at $U_a < 200$ V) allowed us to lower working voltages. Attempts to increase power at the expense of raising current density only have not provided the desirable efficiency level due to saturation.[5] At higher working voltages, the amount of the additive can be decreased noticeably, and above 500 V it is not needed.

Attention was paid to the power distribution as a function of the phosphor layer conductivity. This enabled improved estimations of the phosphor efficiency regardless of the working mode and design of a specific phosphor. The existing values are obviously underestimated.

In the absence of the additive, the optimal working mode of displays with ZnS:Cu,Al (the maximal efficiency and life-

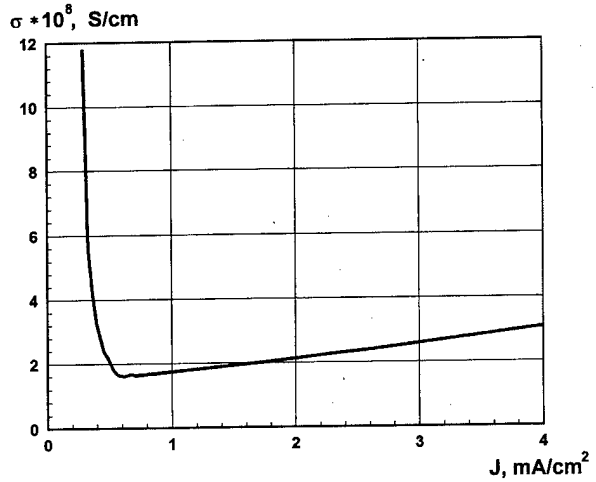


FIG. 6. Conductivity of phosphor layer of cathodoluminescent screen covered with Y_2O_2S:Eu phosphor with 10% In_2O_3 vs anode current.

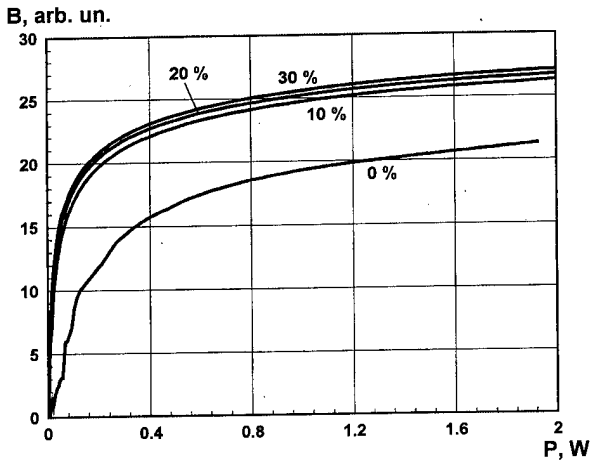


FIG. 8. Brightness as a function of the power dissipated on the display at several different wt % In_2O_3 in the phosphor composition.

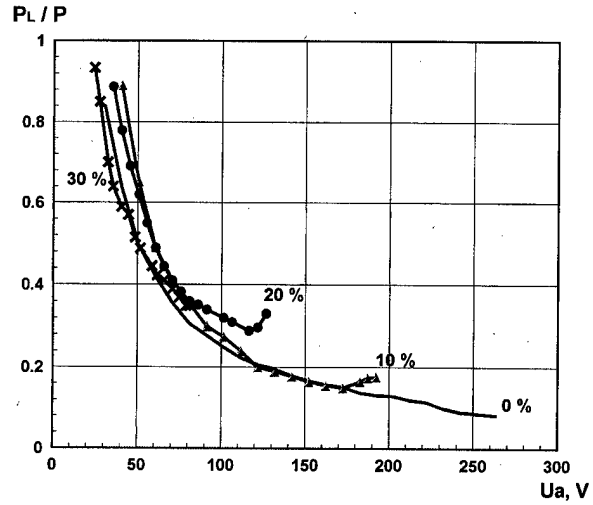


FIG. 9. Electron beam/supplied power ratio as a function of anode voltage at several wt % of In_2O_3 in the phosphor composition.

time) involves high (above 500 V) voltages and low current densities.

The results obtained on VFD can be completely extended to FED screens.

IV. CONCLUSIONS

Up to 70%–80% of the voltage applied to the anode drops across the phosphor layer if made of a phosphor (such as ZnS:Cu,Al or Y_2O_2S:Eu) with a low conductivity without a conductive additive. The use of a conductive additive is obligatory for anode voltages below ~500 V and not required above this value. Electron-induced conduction plays a significant role in the electrophysical properties of phosphors under electron bombardment. The dissipated/supplied power ratio for the phosphor layer in displays can be calculated and used to estimate display performance. The optimal mode of display operation can be calculated for prescribed values of anode voltage and current density.

Presented at the IVMC '97 Conference, Kyongju, Korea, 17–21 August 1997.

[1]O. Hachenberg and W. Brauer, Adv. Electron. Electron Phys. **11**, 413 (1959).
[2]E. G. Bylander, J. Appl. Phys. **49**, 1188 (1978).
[3]M. Hiraki *et al.*, J. Lumin. **12/13**, 941 (1976).
[4]A. O. Dmitrienko, S. L. Shmakov, S. A. Bukesov, and N. V. Nikishin, in Proceedings of 2nd International Conference on the Science and Technology of Display Phosphors, San Diego, CA, 18–20 November 1996, pp. 325–328.
[5]C. Stoffers, S. Yang, S. M. Jacobsen, and C. J. Summers, in Proceedings of the 1st International Conference on the Science and Technology of Display Phosphors, San Diego, CA, November 1995, pp. 225–227.

Resistivity effect of $ZnGa_2O_4$:Mn phosphor screen on cathodoluminescence characteristics of field emission display

Shin-Sung Kim, Sung Hee Cho,[a] Jae Soo Yoo,[a),b)] Sung Ho Jo, and Jong Duk Lee
School of Electrical Engineering, Seoul National University, Shinlim-Dong San 51, Kwanak-Gu, Seoul 151-742, Korea

(Received 18 July 1997; accepted 26 May 1998)

The cathodoluminescent characteristics of phosphor screens excited by cold electrons from field emitter arrays can be influenced by the resistivity of the phosphor screen. The resistivity of $ZnGa_2O_4$:Mn phosphor screens was determined experimentally and the correlation between the screen resistivity and the luminance was obtained. Addition of WO_3 to $ZnGa_2O_4$:Mn phosphor was one way of controlling the resistivity, and subsequently the luminance. In field emission displays (FEDs), the resistivity of the phosphor screen gives rise to a voltage drop on the anode side and modifies the emission characteristics of emitters, which leads to a negative effect on the luminance. The luminance has been generally expressed as $L=\eta I_{\text{anode}}V_{\text{anode}}$. But with the resistivity of the phosphor screen taken into consideration, it should be expressed as $L=\eta I_{\text{anode}}V_{\text{phosphor}}$. For either reason, the resistivity of the phosphor needs to be held as low as possible to get the maximum luminance out of FED. Experiments show that the resistivity of $ZnGa_2O_4$:Mn phosphor is typically 10^{10} Ω cm. © *1998 American Vacuum Society.* [S0734-211X(98)11204-0]

I. INTRODUCTION

The selection of the operating voltage for phosphor screens is still controversial. No need for focusing and low-cost spacer technology are attractive features of low voltage field emission display (FED) construction, which have been demonstrated for 10 in. VGA displays.[1] To date, there has been no report of color red–green–blue (RGB) phosphors with good color coordinates, and the development of low voltage phosphors equivalent to ZnO:Zn still remains a big challenge.

High current densities are utilized to excite the phosphor in a low voltage FED to overcome the low luminous efficiency at low voltage excitation. Therefore, charge accumulation on the anode plate can have negative effects on the phosphor lifetime and the emission performance of field emitters. The solution to this problem may be found by understanding the electrical performance of a phosphor screen.

The resistivity of a phosphor screen depends not only on the phosphor but also the screen condition. Conductive materials are added to the phosphor in some cases to increase the conductivity of the phosphor screen. In our previous works,[2–4] metallic oxides such as WO_3, V_2O_5, and In_2O_3 were added to the $ZnGa_2O_4$:Mn phosphors to reduce the resistivity. Here, the resistivity of $ZnGa_2O_4$:Mn phosphor screens was determined through experimental measurements in a vacuum chamber. Cathodoluminescence (CL) brightness was also measured simultaneously. Finally, the effect of the resistivity of the phosphor screen on CL as well as emission current of field emitters are discussed.

[a)]Also with: Dept. of Chemical Engineering, Chung-Ang University, Huksuk-Dong 221, Dongjak-Gu, Seoul 156-756, Korea.
[b)]Author to whom correspondence should be addressed; electronic mail: jsyoo@cau.ac.kr

II. EXPERIMENT

$ZnGa_2O_4$:Mn powder phosphors were prepared and further treated by mixing with a wide band-gap material such as WO_3 to control the optical and electrical characteristics of phosphor screen. The anode current and conductivity of phosphor screens with different WO_3 weight percentages were measured.

Indium–tin–oxide (ITO) coated glass, which is a good conducting material, was used as a reference to estimate the conductivity of phosphor screens. Four $ZnGa_2O_4$:Mn phosphor screens and an ITO coated glass were mounted on a rotator located 2 cm away from the electron gun in a vacuum chamber. Use of a rotator makes it possible to characterize each sample under the same experimental conditions.

Also, the $ZnGa_2O_4$:Mn phosphor screens were excited by a field emission array (FEA) in a vacuum chamber. The correlation between the emission characteristics of FEA and the resistivity of phosphor screen was examined.

III. RESULTS AND DISCUSSION

Figure 1 shows the bias dependence of the anode current for phosphors with various WO_3 contents. To obtain an anode current of 6.8 μA, a bias voltage of 4 V is needed for the ITO sample while additional voltage (ΔV) of approximately 20 V is needed for the screen with 0.2% of WO_3 weight percentage and about 40 V for the others. The effective resistance can be calculated by the relation, $R=\Delta V/I$.

By considering that the light-emitting area and the thickness of the phosphor are, in our system, $(0.5)^2\ \pi$ cm^2 and 10 μm, respectively, the conductivity for each sample can be calculated as shown in Fig. 2. Typically, they are several $10^{-10}\ \Omega^{-1}$ cm^{-1}.

To check the validity of this value, a sample was made with the phosphor layer sandwiched between two metal lay-

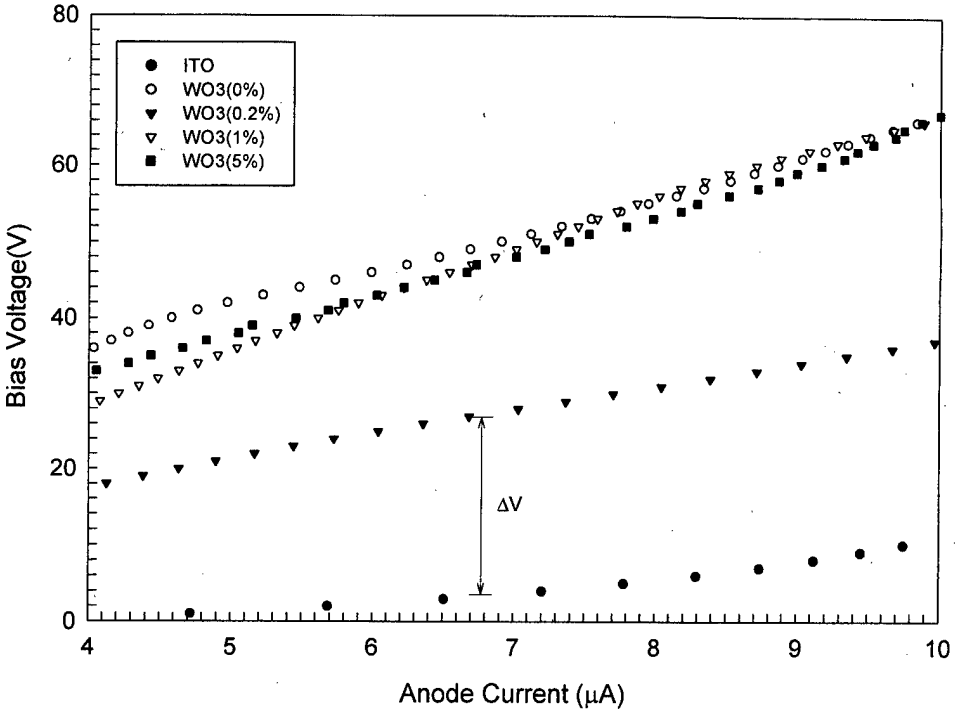


FIG. 1. Bias dependence of the anode current for $ZnGa_2O_4$:Mn phosphor screen with different mixing levels of WO_3.

ers. Aluminum with a thickness of 0.5 μm was deposited on a silicon wafer by E-gun evaporation. Then phosphor was deposited electrophoretically on the Al, followed by another evaporation of Al with 0.5 μm of thickness. From the current–voltage ($I-V$) characteristic of this sample, the calculated resistance is 32.3 MΩ. The top metal layer was $(0.25)^2$ π cm^2 in area and the phosphor layer was 15 μm thick. So, we can estimate the conductivity of the sample to be 2.37×10^{-10} Ω^{-1} cm^{-1}. This gives the same order of magnitude for the conductivity obtained above.

The conductivity increases as the WO_3 percentage is increased. It reaches its maximum at 0.2% by weight of WO_3 and decreases at higher contents. A similar tendency is found when CL intensity is plotted as a function of WO_3

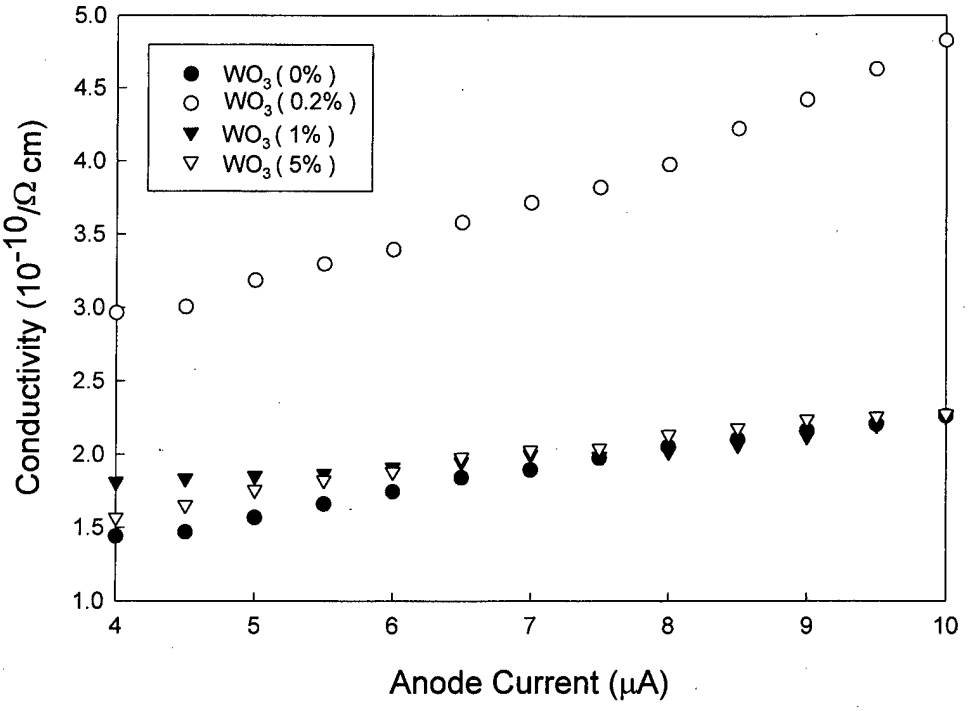


FIG. 2. Conductivity for $ZnGa_2O_4$:Mn phosphor screen with different mixing levels of WO_3.

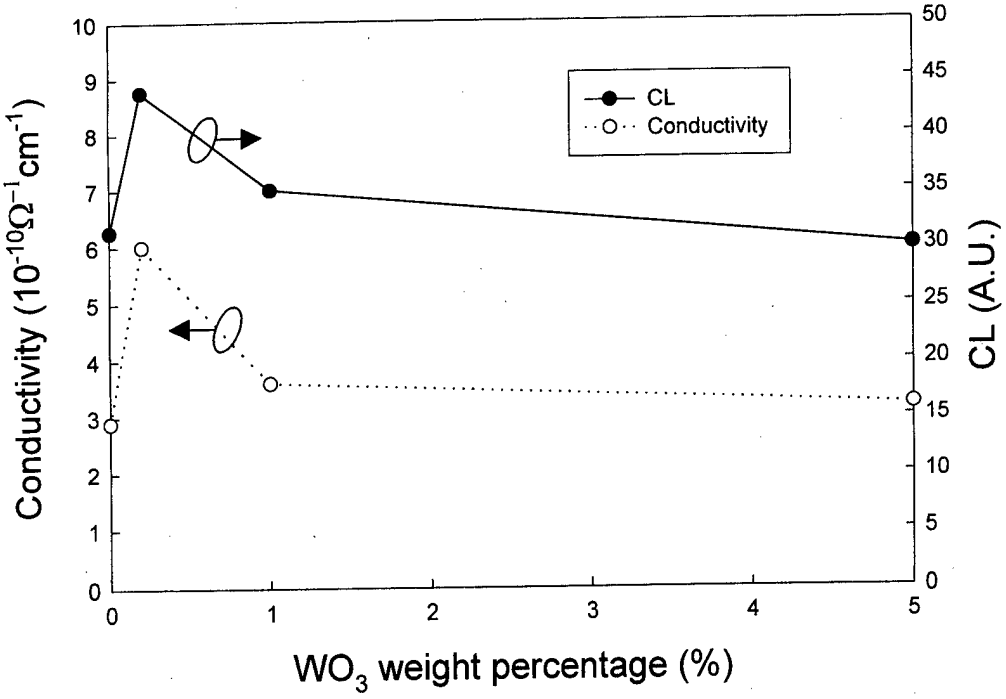


FIG. 3. Correlation between the conductivity and CL when the anode current is 4 μA.

percentage.[3,4] Thus, as shown in Fig. 3, a strong correlation exists between the conductivity and CL of the phosphor. Modifiers such as WO_3 and V_2O_5 are chosen generally because these wide-band oxides are capable of providing a high concentration of positively charged oxygen vacancies.[5] But what their role is with regard to oxygen vacancies and conductivity as a whole has not yet been clearly understood. Based on our experimental results, we suggest a plausible answer to this question.

We start with the basic formula for conductivity, $\sigma = (e^2/m^*)n\tau$, where τ is the relaxation time, n is the free carrier concentration, and m^* is the effective mass of electron. As the percentage of WO_3 is increased, oxygen vacancies can be also increased along with the free carrier concentration. It was confirmed that free carrier concentration increases almost linearly with respect to oxygen vacancies.[6,7] But the relaxation time, τ, decreases as the oxygen vacancies increase, and more so when the oxygen vacancy concentration is higher. Since the conductivity is given as a product of these two counteracting parameters, it should have a concave form when plotted with respect to the oxygen vacancies or WO_3 percentage. This would explain why we have the conductivity curve shown in Fig. 3 and see a good correlation between conductivity and CL of the phosphors studied.

When a phosphor screen is employed in FED, the effect of phosphor resistivity becomes more evident. In FED, sufficient anode voltage (typically 400–500 V) is required to accelerate the electrons enough for substantial luminance and to direct most of the emitted electrons toward the anode where the phosphor target is located. But since there is always an intrinsic resistance to phosphor screens, the surface potential of the phosphor screen V_{phosphor} is reduced from V_{anode} to $V_{\text{anode}} - I_{\text{anode}}R_{\text{phosphor}}$, where V_{anode}, I_{anode}, and R_{phosphor} denote anode voltage, anode current, and resistance of phosphor screen, respectively. The voltage drop across the phosphor screen also induces a decrease of I_{anode} because a smaller portion of the emitted electrons are directed toward anode, again reducing the amount of the voltage drop. Through this process, I_{anode} and V_{phosphor} are determined. It is generally accepted that cathodoluminescence satisfies the relation, $L = \eta I_{\text{anode}} V_{\text{anode}}$, where η is the efficiency of the phosphor. With the resistivity of the phosphor screen taken into consideration, it is properly expressed as $L = \eta I_{\text{anode}} V_{\text{phosphor}}$.

Figure 4 shows the anode current as a function of gate voltage for different values of phosphor resistance. As the resistance is increased, I_{anode} is decreased. For resistance over 10 MΩ, saturation of the anode current is observed. V_{phosphor} is calculated from the simple relation, $V_{\text{phosphor}} = V_{\text{anode}} - I_{\text{anode}}R_{\text{phosphor}}$. By multiplying I_{anode} and V_{phosphor}, we can see how the brightness is changed in terms of R_{phosphor} and V_g as shown in Fig. 5. As the resistance increases, the luminance decreases for a given gate voltage and so does the width of the luminance curve. As the phosphor resistance increases, the maximum luminance decreases while the gate voltage required for that maximum luminance decreases due to the high resistivity of the phosphor screen. These data show how significant an influence the resistivity of the phosphor screen has on the emission characteristics of FEA. Figure 6 shows the CL patterns of the phosphor screen for different levels of anode current. First, the luminance increases as the anode current is increased, but for anode currents beyond 1 mA the pattern becomes darker from the

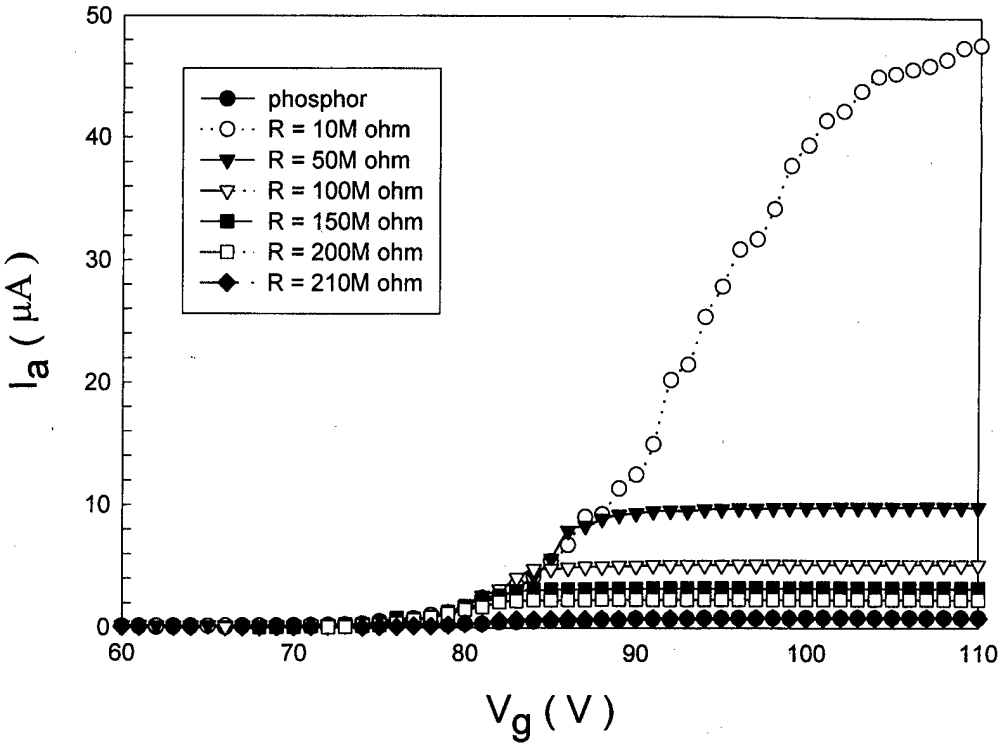


FIG. 4. Anode current vs gate voltage for different values of phosphor resistance.

center. This is thought to be a direct manifestation of the resistivity effect. Since the area nearer to the center of the CL pattern has higher current density, it experiences a larger voltage drop. So darkening from the center is observed.

IV. CONCLUSION

When WO_3 was added to phosphor as a modifier, the luminance increased by 20%. By examining the dependence

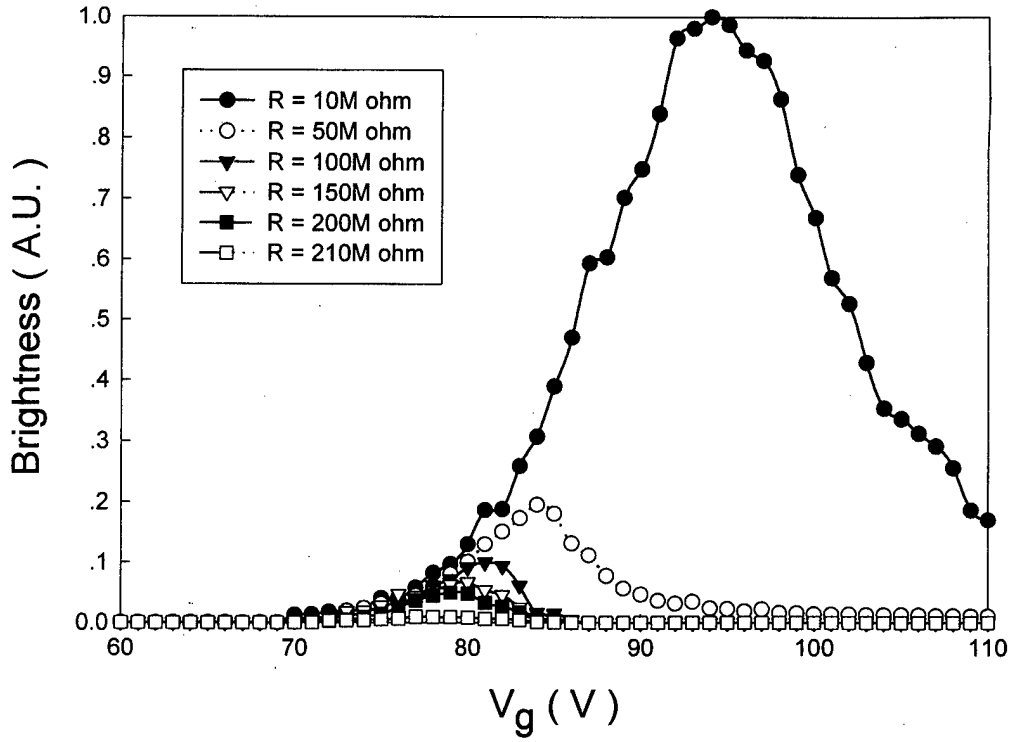


FIG. 5. Brightness vs gate voltage for different values of phosphor resistance. Brightness was given in arbitrary unit and normalized with respect to the maximum brightness when R equals 10 MΩ.

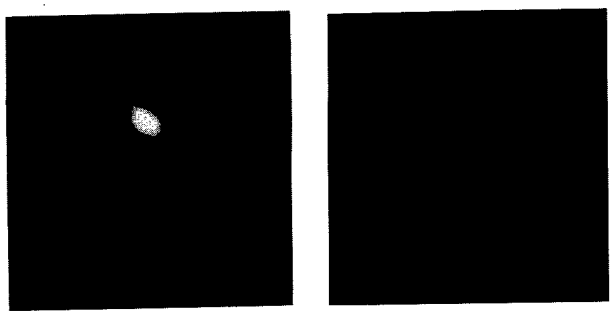


FIG. 6. CL patterns when anode current is (a) 0.5 mA and (b) 1.5 mA. ZnO:Zn phosphor was used and anode voltage was held at 400 V.

of anode current on bias voltage, the conductivity of $ZnGa_2O_4$:Mn phosphors with different WO_3 contents was calculated to be several $10^{-10}\,\Omega^{-1}\,cm^{-1}$. We found that enhancement of the luminance by incorporation of WO_3 into $ZnGa_2O_4$:Mn phosphors could be explained in terms of conductivity.

The resistivity of the phosphor needs to be held as low as possible in order to reduce the voltage drop across the phosphor screen and thus to get the maximum luminance out of FED. As we have some required level of luminance, an increase in phosphor resistance results in a narrow range of V_g for use and may subsequently restrict the design of FED systems. Also, direct evidence was presented that high current density operation of FED may cause low luminance due to the increased voltage drop across the phosphor screen.

ACKNOWLEDGMENTS

This work was supported by ETRI (Electronics and Telecommunications Research Institute) through ISRC (Inter-University Semiconductor Research Center). Also, the authors were financially supported by KICC (Korea Information and Communications Co., Ltd.).

Presented at the IVMC'97 Conference, Kyongju, Korea, 17–21 August 1997.

[1]G. Courreges, IVMC'96 Proceedings, Russia, 1996, p. 669.
[2]S. H. Cho, J. S. Yoo, and J. D. Lee, Extended Abstracts of 2nd International Conference on the Science and Technology of Display Phosphors, San Diego, CA, 1996, p. 211
[3]S. W. Kang, B. S. Jeon, J. S. Yoo, and J. D. Lee, J. Vac. Sci. Technol. B **15**, 520 (1997).
[4]S. H. Cho, J. S. Yoo, J. D. Lee, and J. W. Lee, J. Korean Inst. Surf. Eng. **30**, 1 (1997).
[5]A. O. Dmitrienko and S. L. Shmakov, Inorg. Mater. **30**, 572 (1994).
[6]M. Lubecka, A. Wegrzyn, and D. Sendorek, Thin Solid Films **131**, 15 (1985).
[7]K. Vanheusden, C. H. Seager, and W. L. Warren, Appl. Phys. Lett. **68**, 403 (1996).

Critical role of degassing for hot aluminum filling

J. Proost,[a] E. Kondoh,[b] G. Vereecke, M. Heyns, and K. Maex
IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

(Received 16 October 1997; accepted 17 April 1998)

The purpose of this work is to relate desorption studies on spin-on-glass with electrical data on via resistance and yield. This will allow for a more detailed understanding of the *in situ* degas process preceding hot aluminum filling. It is shown that the degas temperature and time need to be tuned for both physical and chemical water desorption. Insufficient degassing leaves the chemisorbed components unreleased and has a large impact on the electrical behavior of dense via structures, while the isolated ones are unaffected. More evidence for this density effect in the degas characteristic is provided by focused ion beam techniques. © *1998 American Vacuum Society.* [S0734-211X(98)03104-7]

I. INTRODUCTION

In today's microelectronics considerable research is executed in establishing hot sputtered aluminum (hot Al) as a robust and cost effective technology for contact and via filling in the deep submicron region. Although first related literature goes back to the beginning of the decade,[1] vast efforts still continue in both new process[2,3] and hardware development.[4,5] They not only report on the actual Al sputtering but also on the conformal deposition of mostly Ti-based wetting and barrier layers. Relatively few papers deal with the direct impact of degassing on the overall performance of the hot Al process. Degassing becomes particularly important with the introduction of spin-on-glasses,[6] which are now widely used in the semiconductor industry as partially planarizing dielectrics between either the device and first metal level and/or as part of a multilevel metallization. One failure mode which occurs when spin-on-glass (SOG) is used is usually referred to as a "poisoned" via. This term attributes the electrical failure at the via to further outgassing from the SOG during the thermal cycle of the metal deposition process. Today the poisoned via problem is still considered to be one of the key issues in comparing the performance of advanced SOGs, besides their dielectric constant and planarizing capabilities.[7]

Its exact failure mechanism however has never been described. Initial efforts to eliminate the problem were mainly empirical, prebaking the wafers and inserting them into the metal deposition system within a specified time.[8] More profound desorption studies identified H_2O as the main desorption product[9] and showed the time for a significant amount of reabsorption upon exposure to air to be in the order of 1 h.[10] It was concluded that a superior approach would be to complete the degassing bake *in situ* in the metal deposition system, if allowed by the processing equipment and throughput considerations. To our knowledge no data have been published since then relating specific degas parameters to electrical data on via resistance or yield. This article combines electrical results with quantitative thermal desorption studies (TDS) by means of atmospheric pressure ionization mass spectroscopy (APIMS). This will allow for a more detailed understanding and optimization of the *in situ* degas process in order to eliminate the poisoned via problem.

II. EXPERIMENT

A. Material

In this study SOG will be considered both in blanket films and integrated in double-level metal device wafers. In both cases a commercially available precursor, derived from a tetraethylorthosilicate (TEOS)/alcohol solution, was spun in two sequences of 300 nm onto a 250 nm plasma enhanched chemical vapor deposition (PECVD) SiO_2 film. The cure cycle for each sequence consisted of 3 hotplate bakes in air at resp. 80, 150, and 250 °C, followed by an O_2 plasma treatment. Finally a 450 °C anneal was carried out for 70 min in O_2.

For the blanket films the SOG was left uncapped, without any further patterning. These films were used for the APIMS-TDS desorption studies. For the device wafers, used for the electrical evaluation, additional processing was done in order to define via test structures, as shown schematically in Fig. 1. Conventional *I*-line lithography and plasma etching was used to define the vias, with bottom size of 0.4 μm. After wet and dry stripping the vias were left open for several weeks, enabling the exposed SOG at the sidewalls of the via to reabsorb H_2O from the clean-room environment (relative humidity 45% at 22 °C). For the metal deposition a Balzers CLC9000 cluster tool was used, which allowed for an *in situ* degassing, sputter cleaning, conventional Ti/TiN and Al(0.5 wt % Cu) sputtering without vacuum break. The base pressure for this system was 10^{-8} Torr, both in the process chambers and the transport module. For the second metal level a 50 nm Ti wetting layer was sputtered at 300 °C before the actual deposition of 850 nm Al(Cu) at a chuck temperature of 500 °C. During the hot Al(Cu) sputtering part of the Ti wetting layer was transformed into $TiAl_3$. All process parameters in the cluster tool were kept constant throughout this study, except for the degassing. Details of the exact degassing process will be discussed below. Finally the

[a]Electronic mail: proost@imec.be
[b]Present address: Kyushu Institute of Technology, Center for Microelectronics Systems, 820-8502 Fukuoka, Japan.

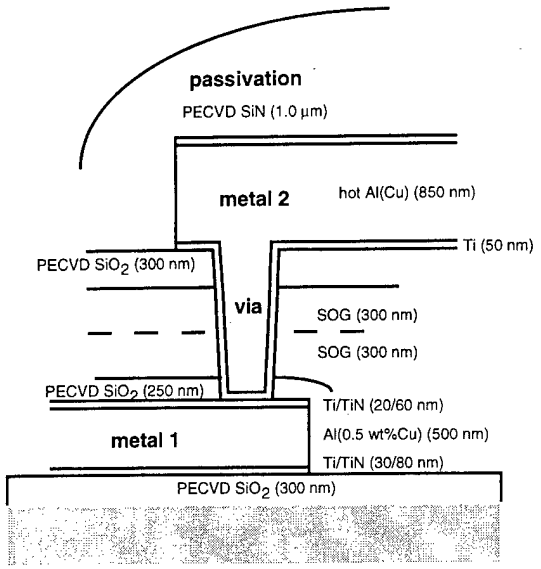


FIG. 1. Schematic cross section of the double level metal device wafers.

device wafers were passivated with 1 μm of PECVD SiN and annealed at 420 °C for 30 min.

B. Analysis

In the degas chamber of the cluster tool the wafer is heated under vacuum conditions by thermal radiation from several halogen lamps, with a total available power of 2.6 kW. No backside Ar is used. The chamber is connected to a turbomolecular pump with a pumping speed of approximately 100 ℓ/s. During degassing the wafer is positioned a distance away from the wafer table block by individual quartz pins. The absence of a direct contact with the wafer table clearly improved the within-wafer uniformity and wafer-to-wafer reproducibility by minimizing the memory effect from this additional radiation source. It also allows for degassing from the backside of the wafer. After degassing the wafer is transferred directly under vacuum for subsequent processing.

For the detailed desorption studies, thermal desorption spectroscopy (TDS) measurements were carried out on blanket SOG films in a separate, commercial single-wafer rapid thermal annealing tool (RTP), connected to an atmospheric pressure ionization mass spectrometer (VG Trace+ APIMS). The RTP used in this study (AST SHS2800 Epsylon) has a rectangular quartz tube, and the wafer temperature is measured and controlled by a thermocouple molded in a small SiC sheath in contact with the wafer backside. The actual wafer temperature had been calibrated with a wafer with a thermocouple embedded at the wafer center. This setup permits a wide range of temperature control from about 50 to over 1000 °C. In our TDS experiments, the wafer temperature was raised at a constant heating rate. Semiconductor-grade nitrogen gas, purified with an active-type purifier, was used as a carrier gas. The ambient gas inside the chamber was sampled and diluted, and then introduced into the APIMS system. The blanket SOG wafers were kept in the

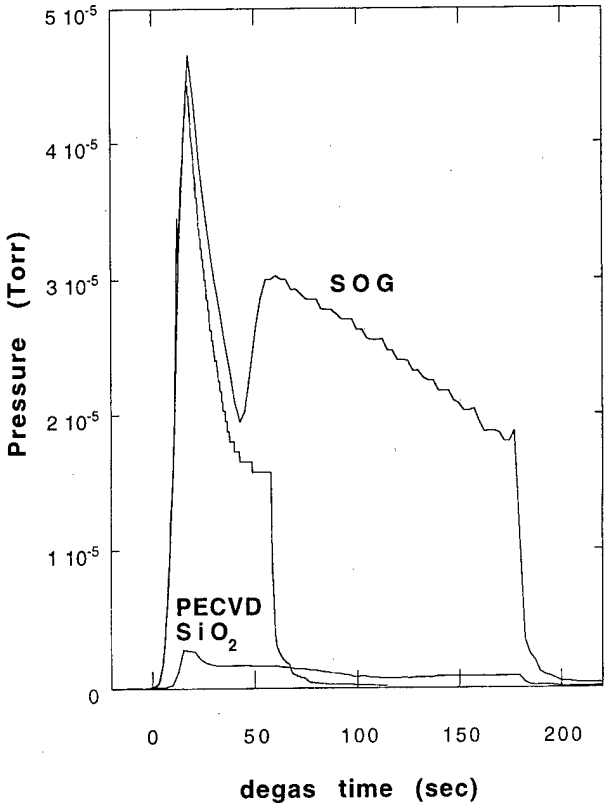


FIG. 2. Total pressure during resp. 180 and 60 s degassing of a device wafer. Included is also the degas characteristic for a blanket wafer coated with 300 nm PECVD SiO_2. The degas temperature is nominally at 380 °C.

RTP chamber for a sufficiently long time to reach the background water level (a few ppb). The total water content in the sampled gas was determined from signals at masses 18 and 19, through calibration with the help of a standard gas generator. The metrological details will be reported elsewhere.

III. RESULTS AND DISCUSSION

A. Desorption

We first focus on the *in situ* degas characteristics in the cluster tool. The degas process was developed as a three-step process, in order to provide the highest possible temperature stability. In the first step the full lamp power is used to raise the wafer temperature to the desired temperature as rapidly as possible. The second step with reduced power is included to limit the overshoot and finally the temperature is maintained constant around 380 °C for the bulk of the process. It should be noted that this process development relied on embedded thermocouple measurements on bare Si wafers. The actual temperature-time characteristic of coated and structured wafers during degassing may still be different (e.g., TiN refractory layers having a higher thermal absorbance compared to Si). Using this specific process the total pressure during 180 and 60 s of degassing for a device wafer (with exposed vias) is shown in Fig. 2. Its most striking characteristic is the appearance of two distinct peaks during degassing for 180 s, whose origin will be discussed later.

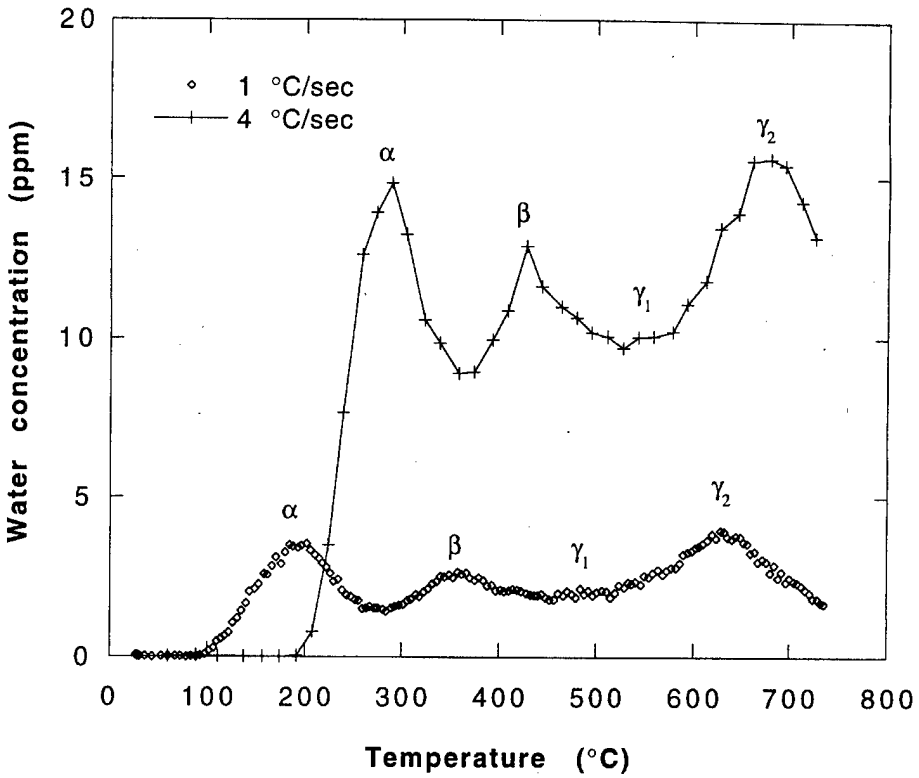


FIG. 3. Thermal desorption measurements on blanket SOG wafers.

Since no further spectroscopic analysis could be done directly on the degas chamber, a more profound TDS study was executed in the RTP system on the blanket SOG films. Upon heating to 700 °C the primary peaks observed are H_2O, with major mass peaks at 18 and 19. Many satellite peaks, originating from water and water-nitrogen complexes are also observed. Other peaks were confirmed at 16 and 29, which were attributed to CH_4 and C_2H_5. Based on the reaction to form the SOG, a preconception might indeed be to see some of the solvent and organics which participate in the reaction, in addition to the expected water molecules.

In the next step the H_2O concentration was measured as a function of temperature, using three different ramping rates. This allows for a further distinction between the different water desorption mechanisms, each occuring at a specific temperature. For each ramping rate a separate blanket SOG wafer was used. Figure 3 shows the obtained TDS spectra for ramping at resp. 1 and 4 °C/s. At least three distinct peaks can be identified. In the case of the lowest ramping rate (1 °C/s), the first peak (α) locates at 192 °C. The second (β) and third (γ_2) are respectively at 355 and 631 °C. A smaller peak (γ_1) is present between β and γ_2. The higher ramping rates shift the peaks towards higher temperatures. Table I summarizes the apparent Arrhenius parameters for the different water desorption mechanisms, as obtained by the kinetic treatment given in the Appendix. The small frequency factors suggest the contribution of (re)adsorption, especially at lower temperature.

Judging from the low peak temperatures and since no preheating was executed before the actual TDS analysis was started, α water can reasonably be attributed to physisorbed water. The observed E_a value (23±2 kJ/mol) is almost half of the enthalpy of water vaporization (40.7 kJ/mol) which corresponds to the strength of one hydrogen bond. This suggests that α water molecules loosely bond to each other and to surface silanol groups in the SOG.

The activation energy for the desorption of β water (55 ±17 kJ/mol) is slightly higher than the heat of water adsorption. This water is generally assigned to tightly, hydrogen-bonded water, where one water molecule is known to be released from two silanol groups:[11]

$$\equiv Si{-}OH \ldots H_2O \ldots OH{-}Si\equiv \rightarrow 2\equiv Si{-}OH + H_2O.$$

The observed E_a value is fairly consistent with the hydroxyl-water bonding energy (>60 kJ/mol).[12] The β desorption reaction may be accompanied by the formation of distorted siloxane, through hydrogen bonding of the silanol groups.[13] Further reduction of these near-neighbor silanols may then give rise to additional water desorption:

$$\equiv Si{-}OH \ldots H_2O \ldots OH{-}Si\equiv \leftrightarrow \equiv Si{-}OH \ldots OH{-}Si\equiv + H_2O$$

$$\equiv Si{-}OH \ldots OH{-}Si\equiv \rightarrow \equiv Si{-}O{-}Si\equiv + H_2O.$$

TABLE I. Apparent Arrhenius parameters of water desorption.

Water	A_0 (s^{-1})[a]	E_a (kJ/mol)	Assignment
α	5.1	23±2	Physisorbed water
β	6.0×10^2	55±17	Tightly hydrogen-bonded water
γ_1	2.5×10^4	89±1	Hydrogen-bonded silanols
γ_2	1.4×10^{10}	202±18	Isolated silanols

[a]If second order, unit is different.

Since the latter is an activated process, it will also occur independently at higher temperatures. Hence γ_1 water can be assigned to this process, where the observed activation energy (89±1 kJ/mol) is fairly in agreement with the one reported by Tompkins and Deal (80 kJ/mol).[14]

Finally γ_2 water is attributed to the direct condensation of H_2O out of isolated silanols:

$$\equiv Si{-}OH + OH{-}Si\equiv \rightarrow \equiv Si{-}O{-}Si\equiv + H_2O.$$

The very high activation energy of 202±18 kJ/mol is consistent with the heat of metal-oxide hydroxylation by water.[12] This process requires still higher temperatures, above the ones normally used in integrated circuit metallization (500 °C). Therefore this desorption reaction will not contribute to a possible poisoned via problem.

Finally we can apply the previously obtained kinetic data to discuss our actual *in situ* degas process in the cluster tool (cf. Fig. 2). Assuming isothermal annealing and solving simple rate equations [(A1), see the Appendix] with the values listed in Table I, we can calculate the residual amount of each type of water as a function of degas time and temperature. Results are shown in Table II. At our actual degas temperature of 380 °C, 60 s is sufficiently long to remove most of the physisorbed (α) water. On the other hand almost 25% of the tightly hydrogen-bonded β water is still retained and γ water cannot be removed substantially. By prolonging the degas time to 180 s, the β water is almost fully released,

TABLE II. Results of kinetic simulations of isothermal degassing.

Temp (°C)	Time (s)	Water	Water retained
300	60	α[a]	10%
		β[b]	70%
		γ_1[b]	99%
380	60	α[a]	1.5%
		β[a]	23%
		γ_1[b]	89%
	180	α[a]	0%
		β[a]	1.3%
		γ_1[b]	70%
		γ_2[b]	100%
500	60	α[a]	0.03%
		β[a]	0.1%
		γ_1[b]	22%
		γ_2[b]	98%
	180	γ_1[b]	19%
		γ_2[b]	95%

[a]First order and
[b]second order are assumed.

while γ_1 desorption starts to become significant. Despite the experimental difference between the APIMS-TDS measurements and the *in situ* degas process in the cluster tool, we believe that concerning the different desorption mechanisms, the same conclusions hold. In fact, the desorption rate estimated from isothermal desorption data (at 400 °C≈10 ppm $=1.7\times10^{-3}$ cm^3/s$=4\times10^{-5}$ Torr for an effective pumping speed of 30 l/s) is fairly comparable to the degassing data obtained in the cluster tool from a device wafer with open via structures (4 to 5×10^{-5} Torr, cf. Fig. 2; note that the contribution of PECVD SiO_2 was shown to be negligibly small). This suggests that almost all the α- and β-absorbed water inside the SOG layer is released through the via holes. Therefore we can get more insight in the origin of the double peaked degas characteristic of Fig. 2 by referring back to our APIMS-TDS analysis. The first peak relates to physisorption (α-water), which requires only moderate heating in vacuum. Although the degas temperature is also sufficiently high to desorb the tightly, hydrogen-bonded water (β water, cf. Table II), care should be taken for explaining the second peak. Desorption kinetics only predict an exponential decay during isothermal annealing. Indeed, when a blanket SOG wafer was isothermally annealed at 400 °C in the TDS system, the initial water concentration showed an exponential decrease, the decay constant (2.5×10^{-2} s^{-1}) being consistent with the value predicted from the previous kinetic data (2.1×10^{-2} s^{-1}). One possible reason for the double peaking is temperature instability during the *in situ* degas process, as a continuous temperature increase indeed produces desorption peaks (cf. Fig. 3). Although the actual wafer temperature was confirmed to be almost constant in the case of bare Si, device wafers can still have very different optical properties, resulting in a slower temperature rise or temperature instability. Moreover the formation of intermediate products from water should be pointed out as well. If primary (α)water could be trapped tightly on, e.g., metal oxide in the bottom of the via hole, the rather long lifetime of these intermediate products can result in a delayed desorption of the β water. In any case, judging from the actual degas temperature, both α and β water are undoubtedly released during 180 s of degassing. More distinct peaks, originating from significant desorption of γ water, were not observed because of the too high temperatures involved. In our future discussions we will denote the desorption of α and β water as resp. physisorption and chemisorption, although the latter may not be fully appropriate to describe the release of tightly bonded water, double hydrogen bonded to two silanols.

B. Electrical evaluation

Based on the results of the desorption study, we identified the total degas time as most critical parameter, necessary for both sufficient physical (α) and chemical (β) desorption. For the electrical evaluation, our standard 180 s degas process was therefore compared directly to a degas process with exactly the same parameters but with a reduced degas time of only 60 s. Figure 2 shows that this mainly releases the physisorpted α water, since no second peak could be observed.

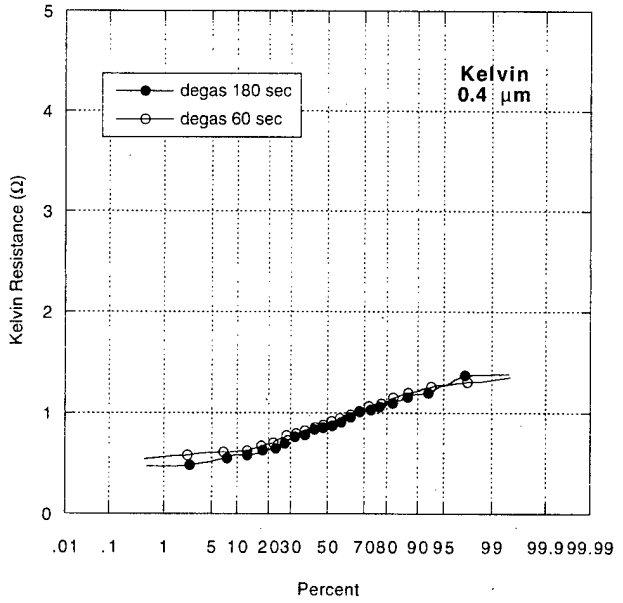


FIG. 4. Statistical distribution of single via resistances, as measured on 0.4 μm isolated Kelvin structures. The dots represent only 20% of the total amount of measurements.

This also agrees with the kinetic simulations of Table II. The electrical measurements were executed on both isolated Kelvin structures and via chains. While the first consist of only one single via, the latter include several arrays of dense vias, with a spacing of no more than 3 μm. Hence these via chains not only improve the statistical significance of yield numbers, but also allow for identifying possible density effects.

The typical layout of our standard via test chain can be described as follows. The total amount of vias included is 50 000 and the total chip area occupied is in the order of 1 mm^2. Intermediate connections for the electrical measurement of chains with 10, 100, 1000, and 10 000 vias are available as well. The main chain direction is horizontally, with chains from 10, 100, and 1000 vias all extending on the lower edge. To include one more horizontal array of vias, a connection is made alternately at the right and left edge. This then results in chains with 10 000 and 50 000 vias. The bottom dimension of all via structures was 0.4 μm. For each of the degas conditions (60 and 180 s) two identical wafers were processed, resulting in a total of 100 individual measurements for each electrical structure.

The statistical distribution of single via resistances, as measured on 0.4 μm isolated Kelvin structures, is represented in Fig. 4. No significant difference could be observed between the 60 and 180 s degas process, neither on the individual via resistance nor on the yield (considered on 100 individual vias). One should notice that "yield" has here its general meaning of the percentage of a population satisfying a predefined resistance specification. For an individual via resistance this is typically set at $\leqslant 3\ \Omega$/via.

For the chain structures, the statistical distribution of the average via resistance, as measured on chains with resp. 100, 1000, 10 000, and 50 000 vias, is shown in Fig. 5. There seems again to be no difference between the two degas processes for the yield measured on the 100 and 1000 via chains. However a large decrease in yield is observed for the 60 s degas process when considering the 10 000 and 50 000 via chains. While a factor 10 increase when going from 100 to 1000 vias did not change the yield number, an additional factor 10 when going from 1000 to 10 000 vias resulted in a detrimental value of less then 50%. This sudden yield loss is believed to be unusually large to be caused only by the increased number of vias in the chain. No such drastic yield decrease was observed for the 180 s degas process. Therefore the yield loss for the 60 s degas process is attributed to an additional density effect, introduced by the typical layout of our standard via chain. It is indeed very striking that this severe yield loss for the 60 s degas process occured specifically for the 10 000 via chains, i.e., the first intermediate chain including "nonedged" vias. For the 100 and 1000 via chains, only relatively isolated vias were included, all located at the edge of the test structure. This density effect in combination with an insufficient degas turns out to affect the hot Al filling process to a large extent. To our knowledge very little has been published before concerning this structure dependence of the poisoned via problem.[15] Instead the problem has often been described as the "random occurrence of high via resistances."[10]

One more additional check for the observed density effect would be to measure a horizontal array of 1000 vias in the bulk of the test structure. Since no suitable bonding pads were available on the standard design, additional pads were fabricated with focused ion beam (FIB), using standard chip repair techniques. As a reference additional connections were made as well on the top and lower edge of the structure. Three chips were treated like this. Of these nine additional "bulk chains" four showed a clear electrical failure through a nonlinear $I-V$ characteristic, while all six reference "edge chains" behaved completely ohmic (constant resistance within the applied voltage range). Although this manual method practically does not allow for a full statistical picture, our additional measurements do provide more evidence for a density effect in the electrical behavior when performing an insufficient degassing step. It will be shown now that this may be correlated with the mean free path of desorbing H_2O molecules in the different process chambers of the cluster tool.

As mentioned above the 60 s degas process is insufficient to fully remove the absorbed β water from the SOG. However after the actual degas process the wafer is still kept inside a vacuum environment and will be barely allowed to cool down before the deposition of the Ti wetting layer. Therefore the remaining H_2O (mostly β water) will still desorb when arriving for Ti deposition (in our setup typically 60 s after the degassing). The wafer is then kept at an additional 30 to 60 s inside the Ti chamber before the actual sputter deposition (at 300 °C), to allow for gas and temperature stabilization and to undo possible nitrification of the target. However, although during degassing the vacuum environment is in the order of 10^{-5} to 10^{-6} Torr (and during

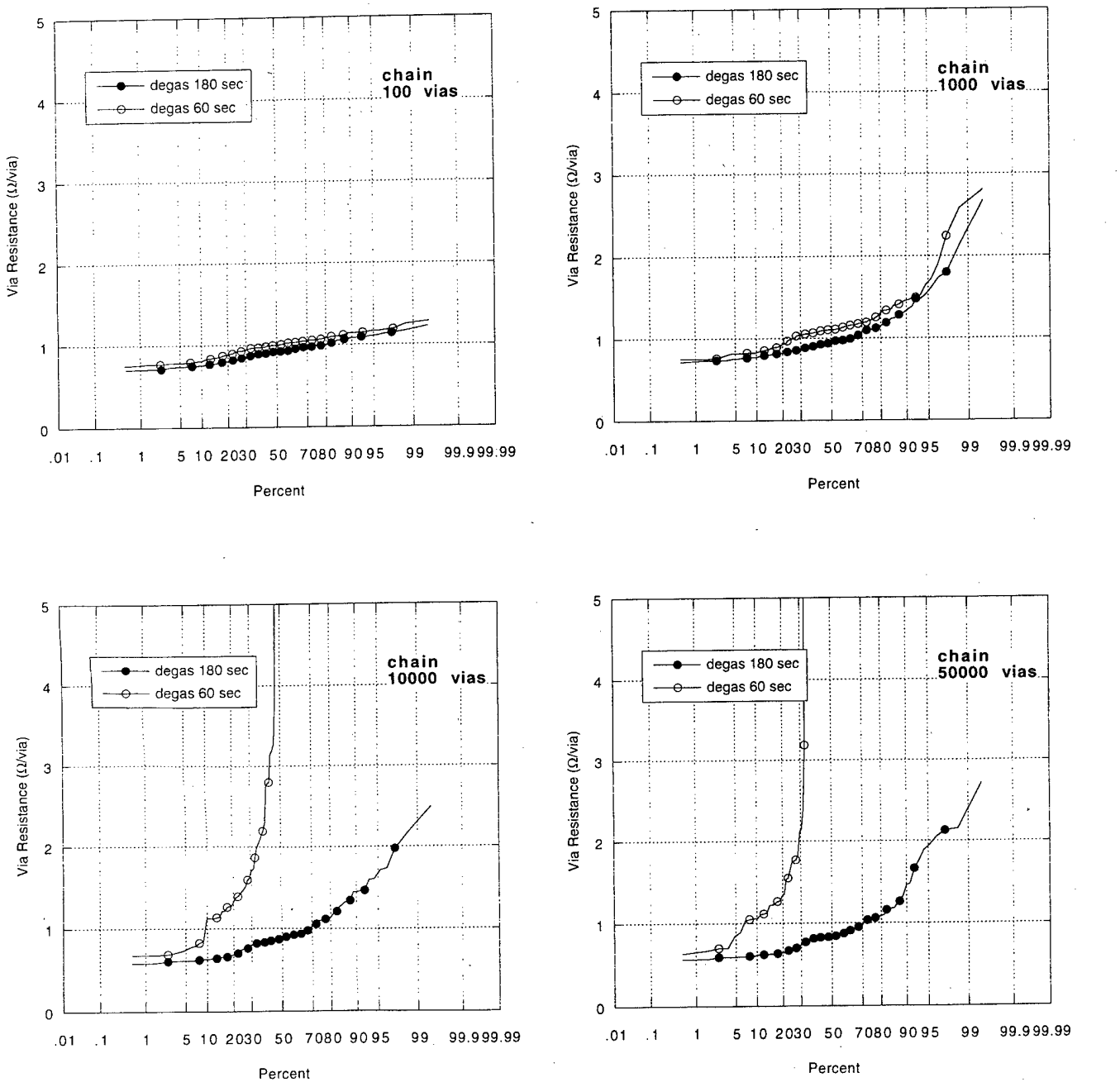


FIG. 5. Statistical distribution of the average via resistance, as measured on via chains with 100, 1000, 10 000, and 50 000 vias. The dots represent only 20% of the total amount of measurements.

the subsequent transport in the order of 10^{-8} Torr), the pressure is significantly increased during Ti deposition to 3 to 4×10^{-3} Torr due to the inlet of both backside and process Ar gas. This accordingly decreases the mean free path of water molecules from the order of meters towards millimeters, so that it actually becomes of the same magnitude as our via-chain test structure. Note that, according to cosine law, the water molecules that are released from the sidewall of the via (where the SOG is exposed) are most likely recaptured at the opposite sidewall. So the mean residence time could very well be longer than that predicted by desorption kinetics. Once the water molecules are released outside the via, they will be scattered and captured on the wafer surface. In view of the reduced mean free path inside the Ti chamber, this process will create a gradient in partial pressure (or surface water concentration) over the via structures upon desorption, with the highest pressure values in the most dense regions. This may locally prevent the β water from fully desorbing and hence may contaminate the deposited Ti in the dense regions, which is known to negatively influence the subsequent hot Al(Cu) filling.[16] At the more isolated edges the lower partial pressure may still allow for a sufficient desorption, resulting in a perfect hot Al(Cu) fill. We speculate that once the Ti deposition is completed no more harm is to be expected from incomplete degassing. This is opposite to the common belief in industry that the Al may be pulled out of

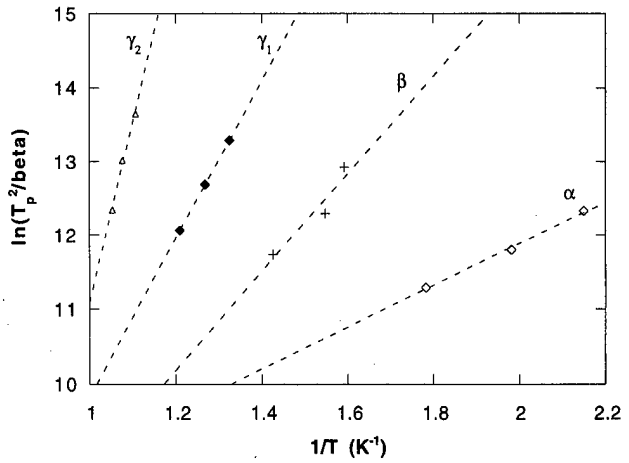


FIG. 6. Relationship between $\ln(T_p^2/\beta)$ and reciprocal absolute temperature for different water desorption mechanisms (cf. the Appendix).

the via by SOG outgassing during the high temperature deposition itself. A more crucial parameter in the case of incomplete degassing may be the delay time between the degas process and the Ti deposition. In order to avoid any irreproducibility neither from hardware limitations nor from density effects, the degas process should therefore be tuned directly to allow for both physical and chemical desorption.

IV. SUMMARY

In this work we combined desorption studies and electrical results to gain a better understanding of the poisoned via problem. It was shown that a 180 s degas process at a nominal temperature of 380 °C was sufficient to release both the physisorbed and chemisorbed water from the SOG. The latter was identified by means of APIMS-TDS as tightly bonded water, double hydrogen bonded to two silonal groups. Slight process modifications, like the use of Ar backgas to improve the heat transfer,[17] may still reduce the overall degas time.

Insufficient degassing, leaving the chemisorbed water unreleased inside the SOG, may then be the result of either a too low temperature or a too short degas time. We showed the latter to have a detrimental impact on the electrical yield of large, dense via chains, while the resistance of isolated structures was unaffected. This density effect was attributed to a continued desorption during the first minutes of further processing in the cluster tool after the actual degas process. Since during the actual deposition of the wetting layer, the vacuum decreased by several orders of magnitude, the large decrease in mean free path for H_2O molecules will generate a pressure gradient, preventing the H_2O from fully desorbing in the most dense regions and hence locally contaminating the Ti wetting layer. Further evidence by means of FIB techniques for this density effect in the SOG degas characteristic also suggests that in general care should be taken in the interpretation of yield results obtained from large chainlike test structures.

ACKNOWLEDGMENTS

The authors would like to acknowledge Z. Nényei and T. Gutt from AST for the technical support on the RTP system, H. Bender for the FIB processing, and L. Papens and B. Scholte Van Mast for their support on the cluster tool. This work was performed with the financial support of the Flemish Institute of Scientific and Technological Research. Karen Maex is a research director of the Fund for Scientific Research Flanders.

APPENDIX

The rate of water release can be written as

$$-\frac{dC}{dt}=k_d\cdot C^n=A_0\cdot C^n\cdot\exp\left(-\frac{E_a}{RT}\right), \tag{A1}$$

where C is the concentration, t the time, k_d the rate constant, A_0 a pre-exponential factor, E_a the activation energy, and RT having its usual meaning. At a fixed carrier gas flow rate, the TDS signal intensity is proportional to the reaction rate. The temperature is changed linearly with time, so $T=T_0+\beta.t$. In the case of a first order reaction ($n=1$), the differential of Eq. (A1) is equal to zero at a peak temperature T_p, which results in

$$\ln\left(\frac{\beta}{T_p^2}\right)=\ln\left(\frac{R.A_0}{E_a}\right)-\left(\frac{E_a}{R}\right)\cdot\frac{1}{T_p}. \tag{A2}$$

In the case of a second order reaction, we obtain

$$\ln\left(\frac{\beta}{T_p^2}\right)=\ln\left(\frac{2.R.A_0.C_p}{E_a}\right)-\left(\frac{E_a}{R}\right)\cdot\frac{1}{T_p}. \tag{A3}$$

Therefore, E_a/R is calculated as the slope of $\ln(\beta/T_p^2)$ vs $1/T_p$. This slope is independent of the reaction order. If free readsorption takes place, adsorption is not an activated process, making E_a identical to the heat of adsorption. Figure 6 shows the plots of $\ln(\beta/T_p^2)$ vs $1/T_p$ for different types of absorbed water. Fairly good linearity is confirmed.

In the case of first order reactions, the pre-exponential term A_0 is obtained by substituting E_a into Eq. (A2). For second order reactions, we need to know the surface concentration at the peak temperature (C_p). When the peak is symmetric and the initial $C=1$, which is indeed assumed for most of high temperature TDS experiments on hydrophilic silica surfaces (like SOG), $C_p=0.5$. Therefore, Eq. (A2) can be used as well to estimate A_0 for $n=2$.

[1]M. Inoue, K. Hashizume, and H. Tsuchikawa, J. Vac. Sci. Technol. A **6**, 1636 (1988).
[2]H. Nishimura, S. Ogawa, and T. Yamada, J. Vac. Sci. Technol. B **13**, 198 (1995).
[3]R. G. Filippi, E. N. Levine, and K. P. Rodbell, J. Vac. Sci. Technol. B **15**, 750 (1997).
[4]S. M. Rossnagel and R. Sward, J. Vac. Sci. Technol. A **13**, 156 (1995).
[5]S. M. Rossnagel, J. Vac. Sci. Technol. B **13**, 125 (1995).
[6]C. H. Ting and H. Y. Lin, Proc. 1987 IEEE VMIC Conference, 1987, p. 61.
[7]Y. Homma, T. Furusawa, H. Morishima, and H. Sato, Solid-State Electron. **41**, 1005 (1997).

[8]C. Chiang, N. V. Lam, J. K. Chu, N. Cox, D. Fraser, J. Bozarth, and B. Mumford, Proc. 1987 IEEE VMIC Conference, 1987, p. 404.
[9]H. G. Tompkins and C. Tracy, J. Electrochem. Soc. **136**, 2331 (1989).
[10]H. G. Tompkins and C. Tracy, J. Vac. Sci. Technol. B **8**, 558 (1990).
[11]W. Stöber, Kolloid-Z. **145**, 17 (1956).
[12]J. Texter, K. Klier, and A. C. Zettlemoyer, Prog. Surf. Memb. Sci. **12**, 327 (1978).
[13]J. J. Fripiat and J. Uytterhoeven, J. Phys. Chem. **66**, 800 (1962).
[14]H. G. Tompkins and P. W. Deal, J. Vac. Sci. Technol. B **11**, 727 (1993).
[15]M. D. Kellam, W. B. Rogers, R. W. Sayer, and R. C. Chapman, Proc. 1991 ECS ULSI Science and Technology, 1991, p. 61.
[16]K.-C. Chen, S.-T. Hsia, J.-N. Kuo, C.-Y. Lee, H. Yen, and B.-C. Chumg, Proc. MRS Adv. Metall. for ULSI Appl., 1994, p. 503.
[17]L.-J. Chen, K.-C. Chen, and S.-T. Hsia, Proc. 1994 IEEE VMIC Conference, 1994, p. 281.

RAPID COMMUNICATIONS

This section is reserved for short submissions which contain important new results and are intended for accelerated publication.

Analysis of scanning force microscope force-distance data beyond the Hookian approximation

Fredy R. Zypman[a)]
Department of Physics, University of Puerto Rico, 100 908th Road, Humacao, Puerto Rico 00791-4300

Steven J. Eppell[b)]
Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106-7207

(Received 12 February 1998; accepted 24 April 1998)

A new method for obtaining force–distance curves using scanning force microscopy is suggested. The theoretical justification for this method is discussed. The method, involving frequency domain measurements of cantilever motion, is predicted to provide improved force data using standard single point measurement of the cantilever deflection. Shortcomings of the standard Hookian model that are alleviated by our model are discussed. Spectroscopic information and knowledge of the applied force during intermittent contact mode imaging are shown to be more appropriately determined using the new method. © *1998 American Vacuum Society.* [S0734-211X(98)03304-6]

Recent experimental advances in scanning force microscopes (SFMs)[1–5] have prompted us to reassess the basic assumptions underlying current theoretical models used to convert experimental SFM data into useful microscopic physical quantities.[6–11] These models are built upon the assumption that the cantilever–tip assembly behaves as a mechanical simple harmonic oscillator (SHO). A system properly described by these models cannot naturally vibrate at more than one frequency.

The SFM community is aware of the fact that multiple vibrational frequencies are excited during standard SFM operation. As a case in point, Fig. 1 shows experimental data of the free oscillation of a commercially obtained SFM cantilever. Clearly, more than one frequency is present in the system precluding generalized application of the SHO theory. Given this situation, it is not appropriate to obtain forces from measured displacements of a SFM cantilever using Hooke's Law ($F=-k\Delta z$) unless it is known *a priori* that the system contains no vibrations above the first resonance.

Previous authors have dealt with the motion of the cantilever at high frequencies. For example, Chen *et al.*[8] studied the natural oscillations of a cantilever with no force of interaction between the tip and surface. Hirsekorn *et al.*[4] studied the motion of a cantilever with the tip attached to an ultrasonically excited surface. In both cases, good agreement between theory and experiment was found. The motivation for these studies lies in the general concern that high frequency modes present in the oscillation of the cantilever nullify the utility of the SHO theory in the analysis of SFM data.

A natural consequence of this state of affairs is that the true force–distance curve cannot be obtained by simply multiplying measured cantilever displacements by some previously determined spring constant. In order to solve this problem, we propose an alternative method utilizing measurement of a resonant frequency of a cantilever loaded by a tip–sample force. As the tip–sample distance varies, this frequency changes.[12] This phenomenon is used in a qualitative way when imaging in noncontact mode and has been shown to be useful quantitatively in obtaining force–distance curves outside the snap-to-contact region.[13] We will show, theoretically, how to extend the use of this frequency shift to obtain a better and more complete force–distance curve than that provided by the SHO theory.

Figure 2 shows a diagram of the geometry relevant to the following discussion. Typical SFMs measure the shape of the cantilever, $u(x,t)$, at one x position near the free end of the cantilever. We assume the cantilever is rigidly attached at one end so that at $x=0$ no translation or rotation occurs. At the free end, $x=L$, the cantilever is loaded with a vertical force, F. The function, $u(x,t)$ is obtained from

$$\frac{EI}{\rho A}\frac{\partial^4 u(x,t)}{\partial x^4}+\frac{\partial^2 u(x,t)}{\partial t^2}=0 \tag{1}$$

together with the above boundary conditions. In this equation, E is the Young's modulus, A is the cross-sectional area, I is the geometric moment of inertia of this area, and ρ is the linear density. Note that we have accounted for the tip–sample force in the boundary conditions.

Since the vibrations in typical SFM work are small, i.e., $u(x,t)/L\ll 1$, the tip probes a small region of the force–distance curve.[14] So it is reasonable to write the tip–sample force as

$$F(d,t)=F_s(d)+G(d)\Delta u(L,t), \tag{2}$$

a)Electronic mail: f_zypman@cuhac.upr.clu.edu
b)Electronic mail: sje@po.cwru.edu

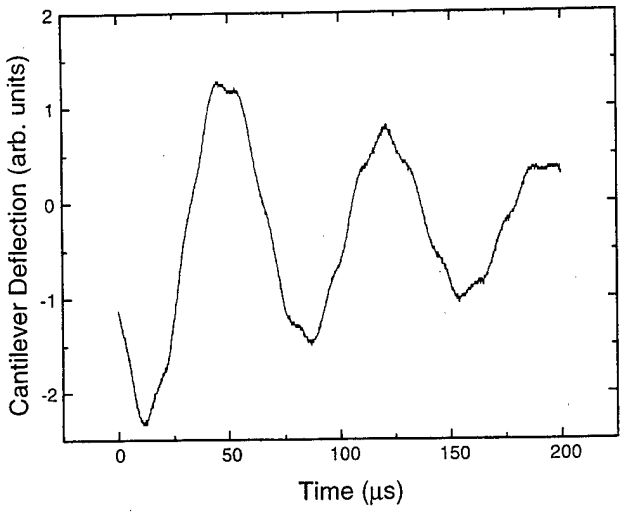


FIG. 1. Free oscillation of a 200 μm thick legged Nanoprobe® Si_3N_4 microlever recorded using a Park Scientific Instruments Autoprobe CP. The oscillation was recorded with a digital oscilloscope just after the tip snapped off the surface (a polished piece of brass). The presence of modes higher than the fundamental is indicated by structure and asymmetry in the decaying sinusoid.

$$d \equiv \frac{1}{T}\int_0^T u(L,t)dt, \quad T \equiv \text{period of oscillation},$$

where d is the average separation between tip and sample, F_s is the average force, G is the slope of the force–distance curve at separation d, and $\Delta u = u(L,t) - u_s(L)$. Note $u_s(x)$ is the curve which describes the static deflection of the cantilever solicited by F_s. The vibration of the cantilever at the free end, $u(L,t)$, is referenced to this static curve yielding $\Delta u(L,t)$.

The form for F given above is reasonable only if the average velocity at which the sample approaches the tip, $(d/dt)d$, is small compared with the instantaneous velocity due to oscillations present in the cantilever, $(d/dt)\Delta u$. To make this point clear, consider a scenario in which the cantilever is excited in its lowest resonant mode. For a flimsy cantilever this corresponds to approximately 10 kHz. Typically one data point is collected each millisecond. The cantilever will oscillate ten times during that interval. So one can study the motion of the cantilever at a particular tip–sample average separation as if that separation were constant. Therefore both F_s and G are slowly varying implicit functions of time as compared with $u(x,t)$. This means that the path traced out by the tip of the cantilever as it approaches the sample surface is the superposition of the conventional force–distance curve and a set of high frequency vibrations, $u(x,t)$ due to resonant oscillations of the cantilever. Since $(d/dt)d < (d/dt)\Delta u$, we are able to make the substitution $\Delta u = \Delta d$ in Eq. (2). Implied in this substitution is a time average of the tip–sample spacing as is evident from the definition of d given in Eq. (2). To obtain the commonly reported force–distance curve, it is then necessary to integrate Eq. (2), as is common practice in SFM resonance methods.[15]

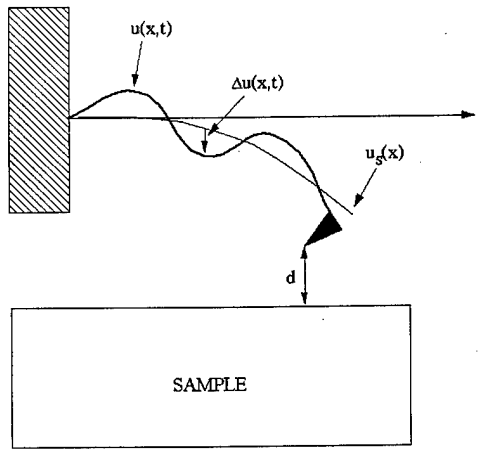


FIG. 2. Geometry of a cantilever oscillation, $u(x,t)$, about its equilibrium deflection shape, $u_s(x)$. The cantilever is a rectangular bar. The tip–sample separation distance, d, is taken as a time average over several periods of the cantilever.

After some substantial calculation, it is found that the resonance frequencies for the cantilever under consideration are given by

$$f_n = \left(\frac{EI}{4\pi\rho A L^4}\right)^{1/2} \xi_n^2, \quad n = 0,1,2..., \tag{3}$$

where ξ_n are the solutions to

$$\frac{1+\cos\,\xi\,\cosh\,\xi}{\sin\,\xi\,\cosh\,\xi - \cos\,\xi\,\sinh\,\xi}\,\xi^3 = \frac{GL^3}{EI}. \tag{4}$$

Equations (3) and (4) provide a means to obtain G from experimentally measured quantities. One would substitute a measured resonant frequency, f_n, into the left side of Eq. (3), solve for ξ_n, substitute this solution into the left side of Eq. (4), and solve for G.

Information regarding the magnitude of the force between tip and sample resides in the shape of the cantilever. Our model contains an explicit and exact connection between cantilever shape and frequency of oscillation. This allows us to use standard existing atomic force microscopy technology which measures cantilever deflection at only one point, and still obtain information contained in any mode of oscillation. In the SHO model, the problem of relating cantilever shape and frequency is not well defined. One would have to add to the model *ad hoc* cantilever shapes in order to obtain tip–sample forces from oscillation frequency measurements. An example of the power of our method is found in the analysis of the snap-to-contact portion of the force–distance curve. As the tip crosses the snap-to-contact point, the lowest vibrational mode experiences a transition which sharply increases its frequency. Also in this region, a static analysis of the cantilever is not possible. If one desired to use a Hookian approach to this problem, it would be necessary to assume more than one mode of oscillation and to have previously measured a spring constant for each significant mode. In addition, the effective spring constant of the system changes as the tip–sample force gradient changes. This means that it is

not possible to write a linear expression of the force balance equation as has been previously suggested.[16] Our method alleviates this problem and places less constraint on the experiment. We are able to use current SFM instruments which provide measurement of only one point on the cantilever. Instead of finding many spring constants, we only need input a material constant for the cantilever, E, its length, and its area moment of inertia, I. The moment of inertia and Young's modulus can be obtained by fitting the data at any frequency since, unlike the spring constant, E and I do not change with frequency or the tip–sample force gradient. Previous publications have shown how to use the formalism presented above to deal with both viscous damping and oscillations after the tip has come into direct contact with the sample.[4,5,17–19] Our method augments these studies by showing how to deal with the forces an oscillating tip/cantilever experiences due to the longer range forces that couple the probe tip to the sample surface prior to direct contact.

Currently, a typical SFM force–distance experiment involves measurement of the position of the tip, the position of the sample, and a single value, k, is used to convert this distance information to a force. This method has been shown to provide good agreement between experiment and the SHO theory outside of the snap-to-contact region and when operating cantilevers at or below their first resonant peak. We are now suggesting that if one wishes to explore the snap-to-contact region or to operate at frequencies above the cantilever's first resonance, or simply to improve accuracy, it is necessary to measure a resonant frequency at each data point. The mode chosen must be higher than the first resonance if one desires to obtain information during snap-to-contact. The shift in the frequency contains the information necessary to reconstruct the force–distance curve. In particular, Eqs. (3) and (4) provide the slope of the force–distance curve at each tip–sample separation. Knowledge of the frequency of oscillation is also used to filter out vibrational cantilever motion yielding the static deflection.

[1]Q. Zhong, D. Inniss, K. Kjoller, and V. B. Elings, Surf. Sci. **290**, L688 (1993).
[2]W. H. Han, S. M. Lindsay, and T. Jing, Appl. Phys. Lett. **69**, 4111 (1996).
[3]G. U. Lee, D. A. Kidwell, and R. J. Colton, Langmuir **10**, 354 (1994).
[4]S. Hirsekorn, U. Rabe, and W. Arnold, Nanotechnology **8**, 57 (1997).
[5]U. Rabe, K. Janser, and W. Arnold, Rev. Sci. Instrum. **67**, 3281 (1996).
[6]N. A. Burnham, J. Vac. Sci. Technol. B **12**, 2219 (1994).
[7]J. Tamayo and R. Garcia, Langmuir **12**, 4430 (1996).
[8]G. Y. Chen, R. J. Warmack, T. Thundat, and D. P. Allison, Rev. Sci. Instrum. **65**, 2532 (1994).
[9]T. J. Senden, C. J. Drummond, and P. Kekicheff, Langmuir **10**, 358 (1994).
[10]H. J. Butt, M. Jaschke, and W. Ducker, Bioelectrochem. Bioenerg. **38**, 191 (1995).
[11]J. P. Cleveland, S. Manne, D. Bocek, and P. K. Hansma, Rev. Sci. Instrum. **64**, 403 (1993).
[12]R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge University Press, Cambridge, 1994).
[13]Y. Martin, H. K. Wickramasinghe, and C. C. Williams, J. Appl. Phys. **61**, 4723 (1987).
[14]This is the same small amplitude approximation used in typical SFM work to equate the z displacement of the cantilever with the motion of the laser spot on the photodetector. Thus the signal amplitude can be made large with respect to thermal noise without violating the small amplitude assumption.
[15]P. Fontaine, P. Guenoun, and J. Daillant, Rev. Sci. Instrum. **68**, 4145 (1997).
[16]F. R. Zypman and S. J. Eppell, J. Vac. Sci. Technol. B **15**, 1853 (1997).
[17]S. Weigert, M. Dreier, and M. Hegner, Appl. Phys. Lett. **69**, 2834 (1996).
[18]F. J. Elmer and M. Dreier, J. Appl. Phys. **81**, 7709 (1997).
[19]O. B. Wright and N. Nishiguchi, Appl. Phys. Lett. **71**, 626 (1997).

Microtrenching resulting from specular reflection during chlorine etching of silicon

Robert J. Hoekstra[a] and Mark J. Kushner[b]
Department of Electrical and Computer Engineering, University of Illinois, 1406 W. Green St., Urbana, Illinois 61801

Valeriy Sukharev and Phillipe Schoenborn
LSI Logic Corporation, 3115 Alfred St., Santa Clara, California 95054

(Received 9 February 1998; accepted 10 April 1998)

In an effort to increase throughput, the microelectronics fabrication industry has transitioned to high plasma density etching reactors using large source (>800 W) and moderate substrate bias (>100 W) powers in which the ion to neutral radical flux is large compared to reactive-ion-etching systems. These conditions can lead to microtrenching where etch rates are largest at the base of the sidewalls. Microtrenching has been attributed to specular reflection of high energy particles, usually ions, at grazing angles on the sidewalls of the mask and trench. These reflections produce a "focusing" of flux to the corners of the trench which results in locally enhanced etching. In this letter, integrated plasma equipment and Monte Carlo feature profile models have been used to examine the processes and conditions which produce focused fluxes and microtrenching, including the degree of specular reflection and sidewall slope of the mask. Quantitative comparisons are made to experimental measurements of etch profiles. © *1998 American Vacuum Society.*
[S0734-211X(98)00804-X]

As the feature sizes of microelectronics devices decrease to sub-0.5 μm dimensions, the semiconductor fabrication industry is increasingly employing dry etching techniques using high plasma density reactors such as inductively coupled plasma (ICP) systems.[1] These tools differ from conventional reactive-ion-etching (RIE) reactors in that the ratio of the ion flux to the reactive neutral flux to the substrate is larger. Microtrenching is one possible consequence of these conditions. Microtrenching refers to profiles for which the etch rate is larger near the corners of a trench compared to the center of the trench. The etch profile across the floor of the trench is therefore either convex or has vertical slots at the base of the sidewalls. Microtrenching is believed to be produced by the impact of high energy particles (mostly ions) at grazing angles (>80°) on the sidewalls followed by specular reflection where the particles retain a large fraction of their energy and directionality. These conditions lead to "focusing" of the high energy particles at the base of the sidewalls of the feature, resulting in higher etch rates at those locations. Microtrenching can lead to large differences in etch depth across the bottom of features and the possibility of "punchthrough" on etch stops or other thin layers such as gate oxide. Microtrenching was first discussed by Nguyen *et al.*,[2] and the proposal that specular reflection is a major contributing cause was first made by Dalton *et al.*[3]

As plasma equipment models and profile simulators have matured, the ability to self-consistently predict the consequences of plasma-surface interactions on etch profiles has also significantly improved.[4–7] In this regard, in this letter we present results from integrated plasma equipment and profile models to investigate the consequences of specular reflection by high energy grazing-angle ions and neutrals on profile evolution during chlorine plasma etching of Si. The plasma simulator we used in this study is the hybrid plasma equipment model (HPEM) and the profile simulator is the Monte Carlo feature profile model (MCFPM). The models, and their method of integration, are described in Refs. 4 and 8. In the MCFPM, the trench is resolved in two dimensions using a rectilinear numerical mesh having 500×500 cells for a 1 μm×1 μm region. Computational particles are directed towards the surface representing the energy and angularly resolved ion and neutral fluxes produced by the HPEM. Monte Carlo techniques are applied to changed the identity of a mesh cell to represent, for example, adsorption, passivation, and etching processes. The reaction mechanism we used is based on successive chlorination of Si by neutral Cl atoms followed by ion stimulated desorption of $SiCl_n$.[4]

The MCFPM differs from that previously described in the method of treating ion (or high energy neutral) reflections from surfaces. Due to the statistically rough surface inherent to Monte Carlo based simulators, it is necessary to locally smooth the surface at the site of impact to eliminate unrealistic high angle scattering from sharp boundaries of the numerical mesh between the surface and plasma. This smoothing was accomplished by sampling the actual plasma-surface boundary 10–20 cells on either side of the point of impact and making a least-squares fit to the surface. The incident particle then collides and reflects from the smoothed surface.

All experimental and model results discussed in this article are for an inductively coupled LAM 9400SC plasma etching reactor.[9] The reactor operating parameters are: 600 W inductively coupled power, 100 W substrate bias at 13.56 MHz, and 10 mTorr of Cl_2 at a flow rate of 60 sccm. The

[a]Electronic mail: stretch@uigela.ece.uiuc.edu
[b]Electronic mail: mjk@uiuc.edu

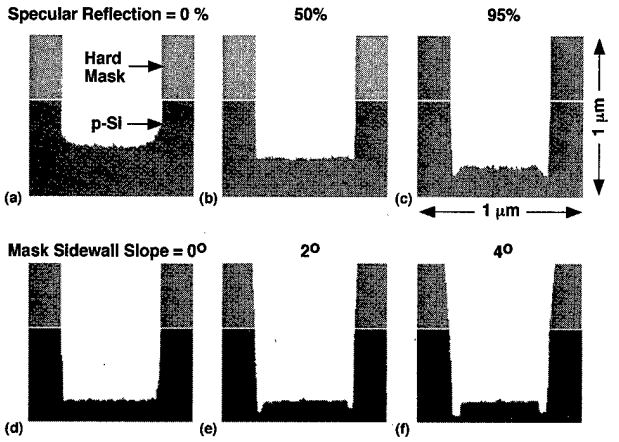


FIG. 1. Etch profiles for chlorine etching of silicon for (a) 0%, (b) 50%, and (c) 95% specular reflection and for outward slope of the mask sidewalls of (d) 0°, (e) 2°, and (f) 4°.

system we are addressing is etching of crystalline Si using a hard mask of SiO_2. The angular dependence of the etch yield of chlorine atomic ions on silicon we used in the MCFPM is that measured by Chang and Sawin.[10] We specified that ions retain as much as 99% of their energy when reflecting at grazing angles ($>80°$) based on results from molecular dynamics simulations by Helmer and Graves.[11] We ignored the effects of surface charging on ion trajectories in the trench due to the moderately high conductivity of the substrate, though charging in the trench is an important process in producing notching.[12]

To demonstrate the dependence of microtrenching on the degree of specular reflection (SR), the fraction of grazing ions allowed to retain their energy was varied from 0% to 95%. The resulting profiles, shown in Figs. 1(a)–1(c) for 0.6 μm wide trenches, reveal significant changes in the morphology of the bottom of the trench. As the fraction of SR is increased, the corners of the trench evolve from being rounded at 0% SR, to being sharp and square at 50%, and finally to having microtrenches at the base of the sidewalls at 95% SR. The onset of microtrenching is a direct result of ion reflection from the sidewalls leading to "focusing" or enhancement of the particle flux at the base of the sidewall. The increased particle flux produces a higher etch rate which, if sufficiently focused, generates microtrenching, generally at $SR>90\%$.

The slope of the sidewall of the mask can also play an important role in the initial development of microtrenching. A finite slope of the mask increases the solid angle of the ion flux from the plasma that can reflect from its sidewalls. Since the angular spreads of the ion flux for the conditions of interest are typically $<4°-5°$, a small variation in the slope of the mask sidewall accesses a significantly larger fraction of the ion flux. For example, the etch profiles shown in Figs. 1(d)–1(f) were obtained by varying the slope of the mask sidewalls from 0° to 4°. As the slope is increased the microtrenching becomes more pronounced and broader. Note that as the depth of the trench increases the relative area of

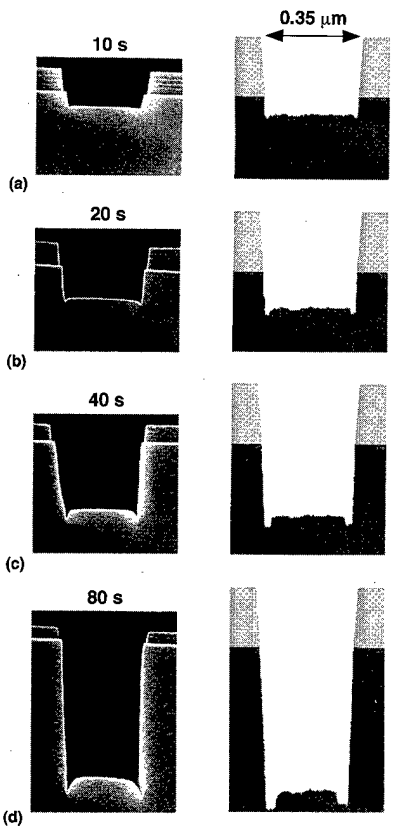


FIG. 2. Comparison of experimental etch profiles (left-hand side) and results from the model (right-hand side) for etch times of (a) 10, (b) 20, (c) 40, and (d) 80 s.

the mask sidewall decreases compared to the exposed sidewall of the trench. Therefore the influence of the slope of the mask on microtrenching is less pronounced for deeper trenches. However, since the solid angle of the ion flux accessible by the sidewalls of the mask is always larger than that for the trench, its slope is always an important consideration.

The evolution of microtrenching is shown in Fig. 2 where predictions from the simulation are compared to experimental results for 0.35 μm wide trenches at different times during the etch. A 0.2 μm thick hard mask was used in the model, while a 0.1 μm hard mask was used experimentally which etched at a rate approximately 5% that of the silicon. The predicted and experimental profiles show the same trends. As the trench deepens, the depth and width of the microtrenches increases as more sidewall area is available to reflect and focus ions. The model shows a more severe broadening of the microtrenches than found in the experiments which may be a consequence of the discreteness of the numerical mesh.

The angular dependence of sputtering yield is typically depressed at normal incidence with a maximum near 60° as recently demonstrated by molecular dynamics simulations by Hanson *et al.*[13] for the sputter yield of Cl^+ on Si. Experiments by Chang and Sawin,[10] however, show a broad maximum in the etch yield of Cl^+ on Si from 0° to 40° with a gradual decrease approaching 90°. These results indicate that

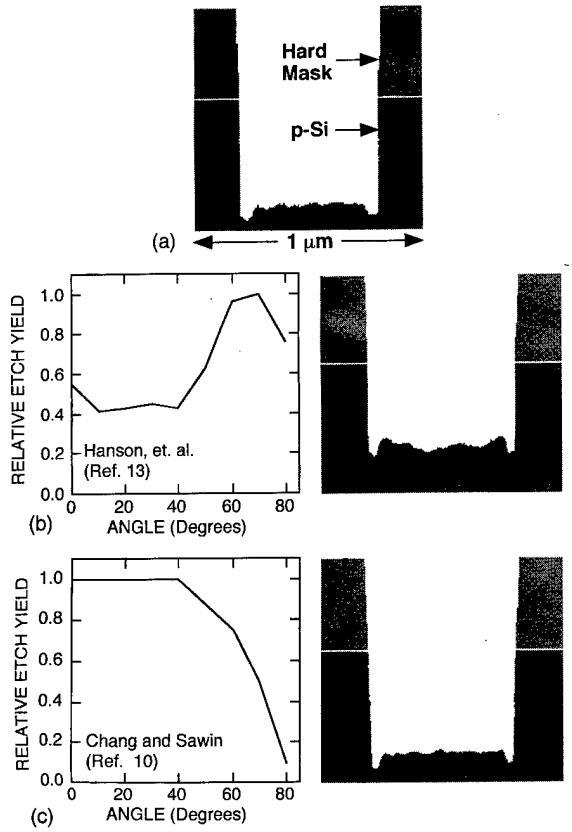


FIG. 3. Etch profiles for different models of the angular dependence of etch yield. (a) Uniform angular dependence, (b) angular dependence as given by Hanson *et al.* (Ref. 13), shown at left for 100 eV ions, (c) angular dependence as given by Chang and Sawin (Ref. 10), shown at left for 35 eV ions.

the mechanism for ion enhanced etching may differ from simple sputtering. The precise forms of these etch yields do have an influence on microtrenching. These trends are shown in Fig. 3 where profiles are plotted for the same etch times while using different angular dependencies for the etch yield. The slope of the mask sidewall is 2°. The profile in Fig. 3(a), the base case, was obtained with no angular dependence of the etch yield. The sidewalls of the trench have a shallow angle and the microtrench is broad due to there being a broad angular flux of reflected ions. The profile in Fig. 3(b) was obtained using the angular dependence of Hanson *et al.* The etch rate is smaller than for the base case since the relative etch yield at near normal incidence is lower. The sidewalls are nearly vertical due to the larger etch yield at grazing angles. There is also structure on the floor of the trench resulting from the extrema in the etch yield and energy loss of ions reflecting from the walls. The profile generated using the etch yields from Chang and Sawin [Fig. 3(c)], has narrower microtrenches than the base case, due to there being fewer losses of grazing incidence ions on the sidewalls. The sidewall slope is larger than for the yields of Hanson *et al.* due to the decrease in etch yield at grazing angles. The total etch rate is commensurate with the base case due to the larger yield at normal angles.

In conclusion, a Monte Carlo feature profile model has been used to investigate the effects of specular reflection of grazing-angle ions on the profile evolution and microtrenching in chlorine plasma etching of silicon. We found that the SR of ions from the sidewalls must exceed 90% at grazing incidence ($>80°$) to reproduce experimentally observed microtrenching. The slope of the sidewall of the mask also has an important influence on microtrenching. Sidewall slopes of 2°–4°, commensurate with the angular spread of the incoming ion flux, increases microtrenching by accessing a larger solid angle of the ions. The angular dependence of the etch yield influences microtrenching as well. Low yields at high angles of incidence allow more ions to retain a larger fraction of their energy after reflection from the sidewalls, thereby producing a narrower microtrench. Microtrenching increases with increasing etch depth due to the larger sidewall area available for ion reflection.

This work was supported by the Semiconductor Research Corporation, National Science Foundation (ECS 94-04133, CTS 94-12565), and the University of Wisconsin ERC for Plasma Aided Manufacturing.

[1]J. H. Keller, Plasma Sources Sci. Technol. **5**, 166 (1996).
[2]S. V. Nguyen, D. Dobuzinsky, S. Stiffler, and G. Chrisman, J. Electrochem. Soc. **138**, 1112 (1991).
[3]T. J. Dalton, J. C. Arnold, H. H. Sawin, S. Swan, and D. Corliss, J. Electrochem. Soc. **140**, 2395 (1993).
[4]R. J. Hoekstra, M. J. Grapperhaus, and M. J. Kushner, J. Vac. Sci. Technol. A **15**, 1913 (1997).
[5]V. Singh, E. S. G. Shaqfeh, and J. P. McVittie, J. Vac. Sci. Technol. B **12**, 2952 (1994).
[6]B. Abraham-Shrauner and W. Chen, J. Vac. Sci. Technol. B **14**, 3492 (1997).
[7]N. Hamaguchi and S. M. Rossnagel, J. Vac. Sci. Technol. B **13**, 183 (1995).
[8]M. J. Grapperhaus and M. J. Kushner, J. Appl. Phys. **81**, 569 (1997).
[9]LAM Research Corp., Fremont, CA; URL http://www.lamrc.com
[10]J. Chang and H. H. Sawin, J. Vac. Sci. Technol. A **15**, 610 (1997).
[11]B. A. Helmer and D. B. Graves, presented at the 44th Annual Symposium of the American Vacuum Society, San Jose, CA, October 1997, paper PS-ThA10.
[12]G. S. Hwang and K. P. Giapis, J. Vac. Sci. Technol. B **15**, 70 (1997).
[13]D. E. Hanson, A. F. Voter, and J. D. Kress, J. Appl. Phys. **82**, 3552 (1997).

Aligned pipe arrays formation by silicon anodic etching

M. C. dos Santos and O. Teschke[a)]
Nano-structure Laboratory, Instituto de Física, UNICAMP, Caixa Postal 6165, 13081-970, Campinas, SP, Brazil

(Received 14 August 1997; accepted 15 May 1998)

A simple technique is described here that produces aligned curved or straight arrays of pipes. This opens the way to produce complete systems (electronic, photonic and fiber optic connectors) from silicon with large scale fabrication techniques. Our strategy in creating parallel pipes is to use high current densities (~100 mA/cm^2) silicon anodization in HF solutions, in a cell where the current direction and the preferential etching directions are not coincident in (111) silicon substrates. The etched structure in a nanoscale range was observed using atomic force microscopy, which shows steps. In a macroscopic scale steps were observed in the walls of arrays of tenths of millimeters diameter pipes. Pores arrays were obtained forming a ~90° angle with the current direction. Molecular mechanics simulations of the pipe wall structure show that a preferential etching along the ⟨100⟩ direction and passivation of the (111) planes are the mechanisms responsible for the formation of pore arrays structure. © *1998 American Vacuum Society.* [S0734-211X(98)07404-6]

I. INTRODUCTION

Silicon etched structures are becoming increasingly important due to the prospect of producing fully integrated electronic devices.[1] The study of the properties of these structures is essential to control material fabrication processes. In this article we will show that, during porous silicon formation by anodization of silicon in HF solutions, arrays of microchannels may be formed by increasing the silicon etching velocity. This opens the way to produce complete systems (electronic, photonic and fiber optic connectors) from silicon with large scale fabrication techniques.

Pore formation in silicon is a result of electrochemical anodization in hydrofluoric acid solutions. The first reports of silicon etching in HF solutions were published by Uhlir[2] in 1956 and Turner[3] in 1958 who observed color changes of the silicon surface as a result of the etching process. Subsequently, Theunissen[4] showed that the interference colors were due to the formation of a porous structure within the silicon substrate. At current densities approaching the electropolishing regime, the pores are elongated, forming pipes. Pipe formation in *n*-type silicon anodically biased in the dark has been described by Theunissen *et al.*[4–6] The resulting channels are needle-shaped and strongly branched. Lehmann and Foll [7] have detailed the different types of "pipe" geometries which occur under different anodization conditions, and have explained "piping" in terms of the diffusion of holes to the silicon surface.

Since anodic biasing of *n*-type silicon always creates a space-charge region (depletion),[8] trench formation occurs spontaneously even at polished surfaces.[4–6] The situation is diverse for the *p*-type material. Since the silicon is positively polarized, there is no depletion layer but an accumulation layer. Consequently, a different mechanism has to be responsible for pipe formation.

In this article, we report atomic force microscopy (AFM) and optical microscopy images which reveal the formation of arrays of curved pipes in etched silicon, as well as straight arrays, which were fabricated at 45° and 90° with the silicon slab surface orientation. We will show that the hydrogen generated by the silicon dissolution reaction as a by-product prevents the pore wall dissolution by fluorine atoms. This is demonstrated by simulating the pore formation mechanism using molecular mechanics calculations

II. EXPERIMENT

The electrochemical cell was made of PTFE, with a rectangular cross section (32×65 mm^2). The counter electrode was a circular platinum mesh with a diameter of 35 mm. Three types of samples were prepared such that the surfaces exposed to anodization were different, according to the following procedures: (i) Fig. 1 shows a (111) silicon 5 mm disk from where ~1 mm thick slices were cut, as indicated, to produce samples A ({010} surfaces) and B ({211} surfaces); and (ii) 2 mm thick with 5 cm diameter (111) oriented silicon wafers were usually cut into 4×20 mm rectangular slabs (samples C), although other sizes were also used. Each silicon slab was mounted perpendicularly to the electrolyte level and contact was made near the edge of the wafer. Only the lower part of the wafer was immersed in the solution. A Teflon coated magnetic stirrer was used to keep the solution well mixed, and also to assist in removing bubbles which accumulate on the wafer surface. A detailed account of the sample preparation technique has been previously described.[9,10] The anodization was carried out in 25% HF solutions using a EGG/PAR 273A galvanostat/potentiostat. Optical microscopic observations were made using a Zeiss Axioplan optical microscope and AFM images were taken in a TopoMetrix TMX-2000 microscope.

III. RESULTS AND DISCUSSION

Porous silicon is obtained at low current densities when high resistivity samples (~1–10 Ω cm) are etched.[11] The

a)Electronic mail: oteschke@ifi.unicamp.br

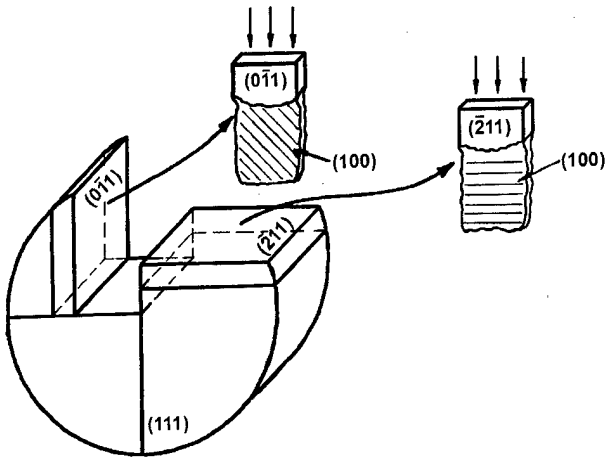


FIG. 1. Schematic diagram showing the crystal orientation and how samples were cut from the silicon wafer. Sample planes exposed to corrosion are indicated.

morphology of a structure formed by applying a low anodic current density to these samples is an interconnected skeleton of nanowires having diameters as small as ~2 nm.[10] It is also generally known that at low current densities the pores are randomly directed in the porous silicon layer. For 0.006 Ω cm silicon, there is neither a microporous layer nor pipe formation when the samples are anodized at low current densities; instead, terraces are formed. Figure 2 shows an AFM image obtained from sample C (0.006 Ω cm), where the exposed surface is (111) oriented, anodized for 5 min at a current density of 5 mA/cm^2. A structure of terraces with ~20 Å steps is observed. Since the AFM tip radius of curvature is ~35 Å, the observed image is the convolution of the etched structure and the tip, which precluded a determination of the angle between the terraces.

A macroscopic observation of the etched structure is then performed. Pore growth was induced by a pattern of pits produced by cutting the sample surface with a diamond-edged cutter. After pipe array formation the slab is sliced by a razor blade and its internal surface is observed. Figure 3(a) shows a region where there are pipes in a typical p-type 0.006 Ω cm silicon (sample A) formed at high current densities (100 mA/cm^2). The current direction is in the plane of the photograph and it is not coincident with a ⟨100⟩ preferential etching direction. Arrays of parallel pipes at an angle of 45° with slab axis direction is observed.

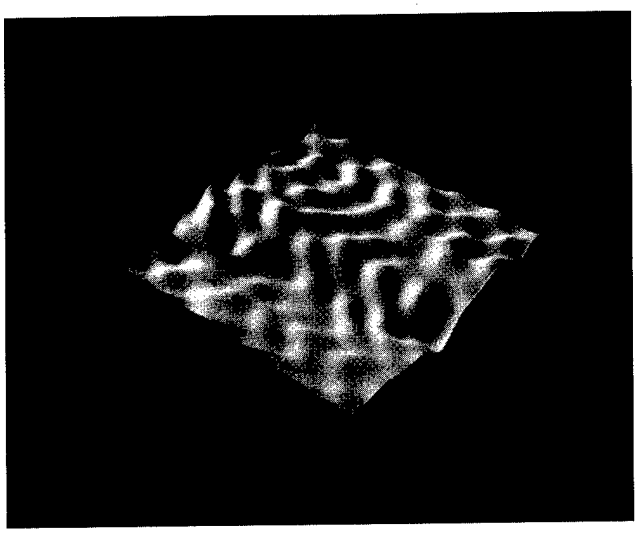


FIG. 2. AFM image of (111)-oriented silicon sample anodized 5 min at 5 mA/cm^2. A 138 Å length is displayed in the x and y directions and 90 Å in the z direction.

Similar structures were also obtained tilted at 90° angle with the slab axis direction by simply cutting the silicon slice as in sample B, where the (211) surface is exposed. Figure 3(b) is a 250× optical view of a sliced slab.

Figure 3(c) shows an array of curved pipes obtained for a sample cut at a nonorthogonal direction with the ⟨111⟩ direction (sample C). It shows pipes with different radius of curvature and various diameters. Figure 3(d) shows an array of pipes where the walls have steps that form an angle of ~70° between them. From these results it is concluded that the diameter distribution of straight etched pipes is almost uniform. The diameter depends on the initial position of the pit in the random pit pattern generated by edge cutting the wafer into a rectangular shape. After 10 min anodization time of the original pit distribution only a few percent survived.

In order to understand the pipe formation it is necessary to deal with the mechanism of the anodic surface dissolution. Although the detailed mechanism is still not clear, it probably occurs as follows: At the onset of the anodic current the silicon surface is all H-covered.[12,13] A field built up across the space-charge layer, moves holes towards the surface at kinks, defects or tensioned regions. This induces a nucleophilic attack on the Si–H bonds by F^- [or HF_2^- (Ref. 3)] ions, forming SiF_2 groups at these sites and H^+ ions in the solution. The Si–Si back bonds of the SiF_2 groups are stretched due to the fluorine electronegativity allowing the insertion of F^- (or HF_2^-) ions. The reaction detaches the SiF_2 group from the surface into the solution, and forms two new Si–F bonds which react again as

$$[Si_N]SiF_2 + 2h^+_{sol} + 2F^-_{sol} \rightarrow [Si_{N-1}]SiF_2 + (SiF_2)_{sol}. \quad (1)$$

This changes the surface geometry and the electric field distribution so that the next hole transference will preferentially occur at this location thus enlarging the "pore" (see Fig. 4). Equation (1) describes the silicon dissolution reaction (SDR). The intermediary product SiF_2 reacts rapidly with HF, forming H_2 and SiF_4 (Ref. 14):

$$(SiF_2)_{sol} + 2(HF)_{sol} \rightarrow (H_2)_{sol} + (SiF_4)_{sol}. \quad (2)$$

The H_2 oxidation reaction, which occurs at the Si–H surface bonds covering the electrode surface, is given by[13]

$$(Si{-}H)_{solid} + (H_2)_{sol} \rightarrow (Si{-}H)_{solid} + 2e^- + 2H^+_{sol}. \quad (3)$$

This oxidation of H_2 fixes the electrode potential at a value which prevents its dissolution by fluoride ions.[14] This

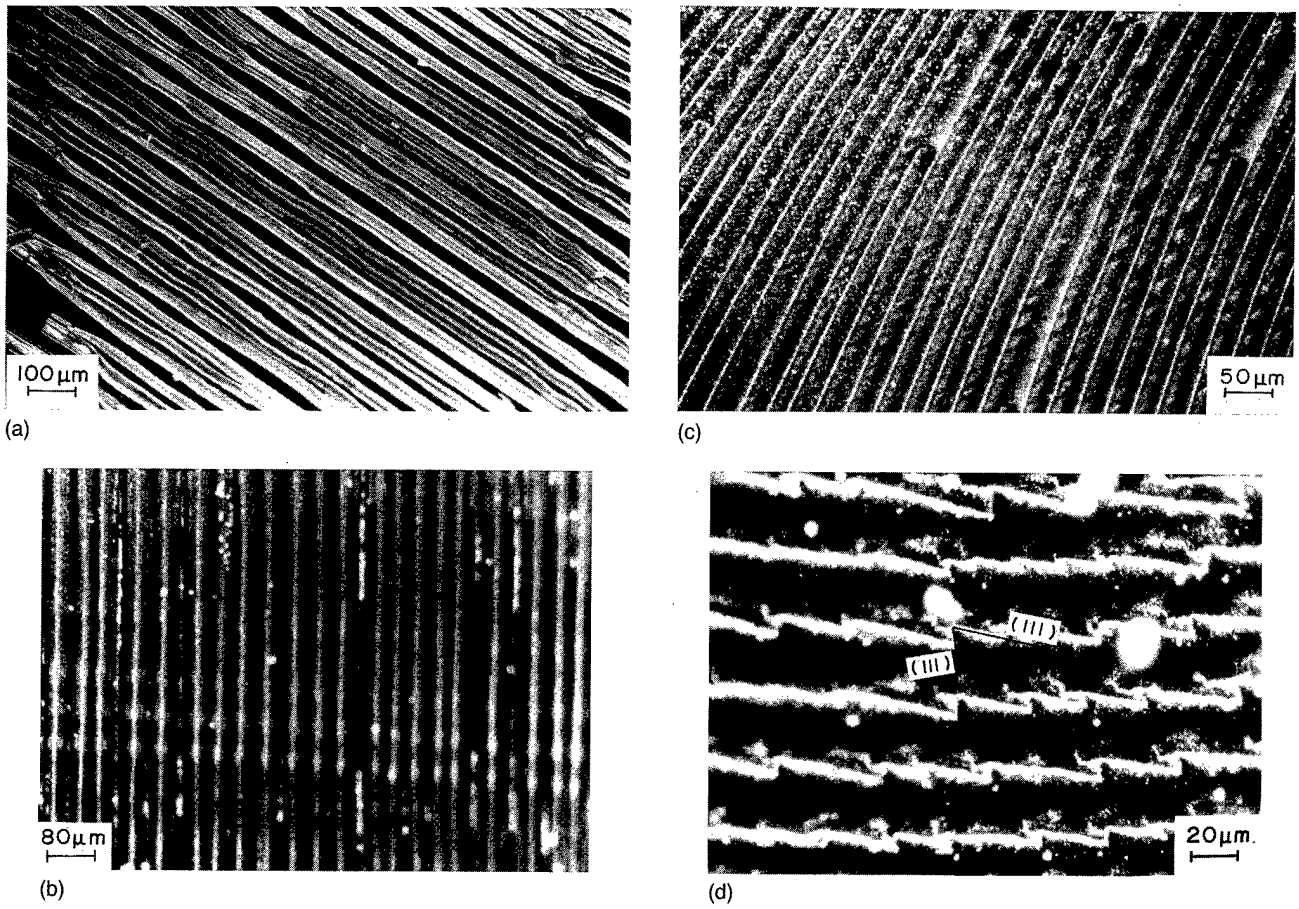


FIG. 3. Array of parallel pipes etched at a current density of ~200 mA/cm^2. (a) Array of pipes at an angle of 45° with the (111) edges (sample A); (b) array of pipes at an angle of 90° with the (011) edges (sample B); (c) array of curved pipes (sample C); (d) array of pipes showing stepped walls formed by the intersection of (111) silicon planes.

is illustrated in Fig. 4. It also displays the schematic distribution of the electric field lines which leads to pipe formation.

Previous work on pipe formation in *n*-type silicon concluded that the space-charge region itself protects the pipe walls from being etched as follows:[7] If all minority carriers are collected by the pipe tips there are none left to permit etching of the wall. This means that a depletion effect is the reason for pipe formation, not an inhibition layer of a different chemical composition at the pipe walls. Here we report on pipe formation in *p*-type material; since the silicon is positively polarized there is no depletion layer and consequently we propose a different mechanism for pipe formation. The hydrogen covered silicon surface (Si–H bonds) which forms the lateral walls of the pipes is passivated by the hydrogen oxidation reaction (H_2 released at the reaction sites) which injects charges into the H-covered sites, preventing wall dissolution by fluorine atoms and increasing the dissolution rate at the bottom of the pipe to where the injected charges migrate. The observed doping dependence of the pipe formation is due to a complex interplay of the silicon dissolution reaction [Eq. (1)] on the dopant concentration, and its by-product reaction [Eq. (2)]. The molecular hydrogen concentration on the H-covered sites then play an important role in the rate of the dissolution reaction at the pore bottoms.

A second effect to be considered is the ability of hydrogen to interact with structural defects and impurities present in the silicon lattice. In particular it has been found that, in crystalline silicon, all the acceptors were passivated by hydrogen, causing a dramatic increase in the resistivity of *p*-type material.[15] This effect adds to the passivation effect produced by hydrogen at the pore walls.

Guided by these results, we performed a large cluster calculation including typically 1000–2000 silicon atoms at the molecular mechanics level. These calculations were carried out within Cerius,[2] a program developed by Molecular Simulations Incorporated, and the universal force field parametrization[16] has been adopted. Conformations obtained were in agreement with our previous AM1 geometry optimization.[13]

Enthalpies of reaction of the cluster with HF were found to be very much dependent on the local symmetry around the silicon atom involved in the reaction. The main conclusions could be summarized as follows: Electrochemical dissolution of silicon in concentrated hydrofluoric acid solutions involves the reaction of Si with HF. Once $>SiF_2$ species are formed at the surface, Si–Si back bonds weaken and SiF_2

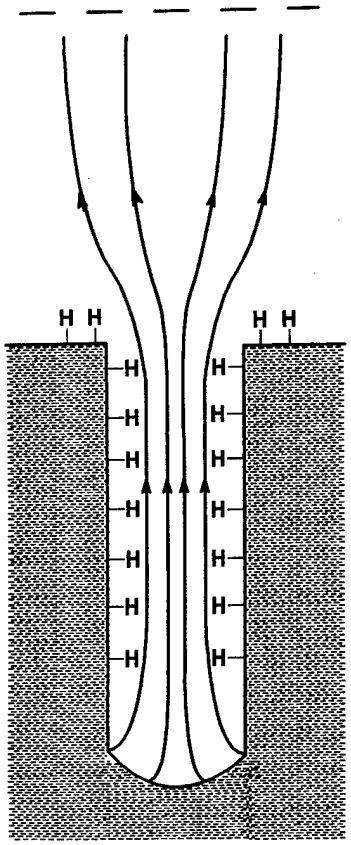
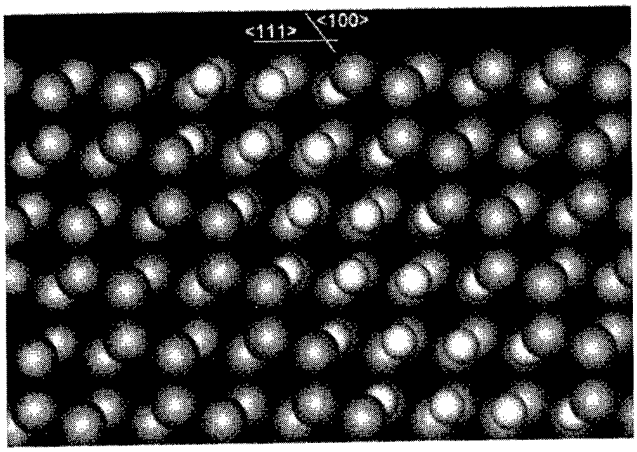
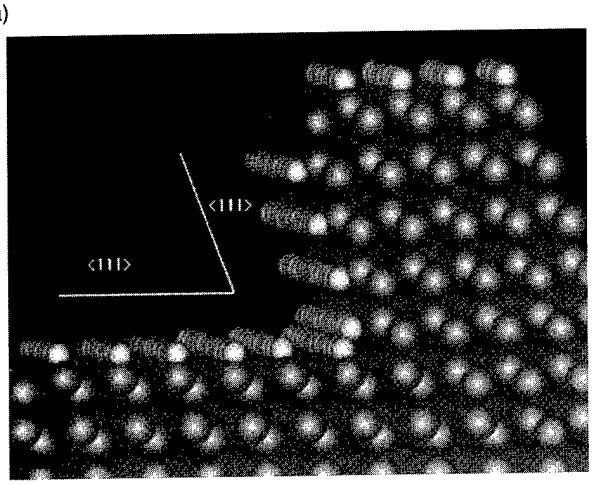


FIG. 4. Schematic diagram of pipe formation mechanism showing the distribution of the electric field lines and the H-covered wall passivated by the hydrogen oxidation reaction.



(a)



(b)

FIG. 5. Molecular mechanics simulations of the pore walls: (a) transversal section of a small straight pipe forming an angle of 54.7° with the (111) silicon surface and (b) lateral view of a stepped wall that gives rise to curved pipes. The step is formed by intersecting (111) planes, at an angle of 70.5°. Dark and light gray spheres represent hydrogen and silicon atoms, respectively.

molecules separate from the substrate to the solution to further react with HF, producing SiF_4 and H_2.[15] Reaction enthalpies calculated on model clusters at the AM1 level showed that silicon atoms at the (111) surface are much more stable than those at the (100) surface. The energy released in the formation of a SiF_2 fragment on a (111) surface is 14.0 kcal/mol and this reaction requires a high activation energy step which is the breaking of a Si–Si bond. On the other hand, no Si–Si bond cleavage is needed to create such a fragment on a (100) surface, and the energy release obtained in this case is 37.0 kcal/mol. This means that the dissolution reaction tends to follow a path that is determined by the easy direction of etching (atoms along the ⟨100⟩ direction), and that the exposure of distinct silicon surfaces should determine different patterns of pores at the electrode surface.

Slices oriented as samples A and B have been graphically obtained by faceting a silicon crystal through the planes (111), (211) and (011). A slice having (111) and (211) plane edges, as in sample A, has the ⟨100⟩ direction forming an angle of 45° with the edges while in a slice having (111) and (011) plane edges, as in sample B, the ⟨100⟩ direction is perpendicular to the (011) plane. These simple geometrical arguments suggest that samples prepared in this way, as shown in Figs. 3(a) and 3(b), would present pipes at 45° and 90° with the edges if one assumes that pipe growing follows the ⟨100⟩ direction.

Figure 5(a) shows the optimized structure of the transversal section of the smallest pipe that could be built starting from a (111) surface and following the path given by the ⟨100⟩ direction. This procedure leaves singly hydrogenated silicon atoms at the pipe walls. This tube is cylindrical in shape, with the axis forming an angle of 54.7° with the (111) surface. Figure 5(b) shows another pipe wall involving a step formed by different (111) planes, which resulted in an angle of 70.5° between them. These structures are in agreement with the ones observed in Figs. 2 and 3. It is also interesting to notice that the intersection of (111) surfaces leaves silicon atoms at the corners that are connected to two other silicon atoms in the wall, that is, in a symmetry similar to the one at the (100) surface. Further fluorine attack on the walls should proceed through these atoms until another step is found.

The high current density (~100 mA/cm^2) anodization process described in this article results in an array of straight parallel pipes with an almost uniform distribution of diameters. Since the silicon slabs used were (111) oriented, their lateral walls have a different orientation and consequently

the attack takes place preferentially at the plate lateral walls. The pore pattern was produced by cutting the wafers with a diamond-edge cutter. When the current direction is not aligned with the ⟨100⟩ direction the pipes are not aligned with the current, and the etching will proceed in the ⟨100⟩ direction until the etch front hits the {111} planes. The pore walls are consequently {111} planes. In Fig. 3(b), pores follow preferentially the direction normal to the current direction in the plane of the photograph, while Fig. 3(a) has pores preferentially aligned at 45° with the current direction.

Regions of small radius of curvature of the aligned pipes results in the decomposition of the curved pipe walls into the two (111) planes forming an angle of 70.5° between then. This is shown in Fig. 3(d). These pore stepped walls are formed by etching various (100) planes and the remaining structure is formed by the more slowly etched (111) planes.

Figures 3(a) and 3(b) also show that the symmetry of the exposed silicon surface may lead to pipes in any direction, i.e., pipe formation with any orientation or radius of curvature, down to crystallographic steps, may be possible by step etching along the ⟨100⟩ etching directions forming an array of pores with (111) walls.

IV. CONCLUSIONS

In this article we present experimental results showing that silicon slab anodization in HF solution forms arrays of straight or curved parallel pipes. The proposed mechanisms of pipe formation are a preferential etching along the ⟨100⟩ directions and a preferential hydrogen passivation of the (111) planes, in agreement with the observation that during the etching process we may obtain walls formed by (100) and (111) planes. Molecular mechanics simulations of the pipe wall structure give further support for these mechanisms.

ACKNOWLEDGMENTS

The authors are grateful to J. R. Castro and L. O. Bonugli for technical assistance. This work was supported in part by FAPESP 96/4431-9, FINEP and CNPq Grant No. 523.268/95-5.

[1]K. D. Hirschman, L. Tsybeskov, S. P. Duttagupta, and P. M. Fauchet, Nature (London) **384**, 338 (1996).
[2]A. Uhlir, Bell Syst. Tech. J. **35**, 333 (1956).
[3]D. R. Turner, J. Electrochem. Soc. **105**, 402 (1958).
[4]M. J. J. Theunissen, J. Electrochem. Soc. **119**, 351 (1972).
[5]M. J. J. Theunissen, J. A. Appels, and W. H. C. G. Verkuylen, J. Electrochem. Soc. **117**, 959 (1970).
[6]M. J. Hill, J. Electrochem. Soc. **120**, 1425 (1973).
[7]V. Lehmann and H. Foll, J. Electrochem. Soc. **137**, 653 (1990).
[8]S. M. Sze, in *Physics of Semiconductor Devices* (Wiley, New York, 1969).
[9]O. Teschke, M. C. Gonçalves, and F. Galembeck, Appl. Phys. Lett. **63**, 1348 (1993).
[10]O. Teschke, Appl. Phys. Lett. **68**, 2129 (1996).
[11]O. Teschke, F. Alvarez, L. R. Tessler, and M. U. Kleinke, Appl. Phys. Lett. **63**, 1927 (1993).
[12]G. W. Trucks, K. Raghavachari, G. S. Higashi, and Y. J. Chabal, Phys. Rev. Lett. **65**, 504 (1990).
[13]D. M. Soares, M. C. dos Santos, and O. Teschke, Chem. Phys. Lett. **242**, 202 (1995).
[14]D. M. Soares, O. Teschke, and M. C. Santos, Langmuir **12**, 2875 (1996).
[15]S. J. Pearton, W. L. Hansen, E. E. Haller, and J. M. Kahn, J. Appl. Phys. **55**, 1221 (1984).
[16]A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024 (1992).

SHOP NOTES

These are "how to do it" papers. They should be written and illustrated so that the reader may easily follow whatever instruction or advice is being given.

Side-by-side wafer bonding of InP for use with stepper-based lithography

R. W. Ryan, R. F. Kopf,[a] A. Tate, J. Burm, and R. A. Hamm
Lucent Technologies, Bell Laboratories Division, 700 Mountain Avenue, Murray Hill, New Jersey 07974

(Received 27 March 1998; accepted 8 May 1998)

We have developed a simple technique for wafer bonding for use with stepper-based lithography. The technique involves side-by-side bonding of two or more pieces of a wafer together with epoxy. A bonded wafer with the same dimensions, such as thickness, flatness, and position of the large flat, as a whole wafer can be run on a stepper. Since a stepper can align each field independently, all of the area without the actual bond may be patterned. This technique is quite useful for salvaging a broken wafer in the middle of processing. In addition, it may be used for bonding several substrates of different materials together for monolithic integration. © *1998 American Vacuum Society.* [S0734-211X(98)04804-5]

I. INTRODUCTION

There is considerable loss from breakage during processing of InP and GaAs wafers, since they are very fragile. Unfortunately, wafer breakage almost always occurs near the end of processing, rendering the wafer useless when using a stepper, and difficult to handle when using contact lithography. Breakage can be quite costly if it occurs near the end of the process, since much effort and money have already been investigated into it by this time.

We have developed a simple technique for wafer bonding. The technique involves side-by-side bonding of two or more pieces of a wafer together with epoxy. The bonded wafer has the same dimensions, such as thickness, flatness, and position of the large flat, as a whole wafer. Stepper, or projection, lithography involves a step and repeat system which exposes the same pattern many times across a wafer. When the stepper loads the wafer, a rough alignment to the large flat is performed, then a fine alignment to several individual fields is performed. Since the stepper can fine-align each field independently, the fields in the area without the actual bond can be patterned. This technique is quite useful for salvaging a broken wafer, where only the fields which contain the crack need to be eliminated from the alignment process. In addition, this method may also be used for bonding several substrates of different materials together for monolithic integration.[1,2] It is also useful for structures which would otherwise have to be grown using selective regrowth, which require elaborate growth techniques to get proper step coverage and good junctions across the etched mesas.

[a]Electronic mail: rek@lucent.com

II. EXPERIMENT

Metalorganic molecular beam epitaxially grown InGaAs/InP heterojunction bipolar transistor (HBT) wafers were used for this study. The structure has been described previously.[3] This insured that built-in stresses from any lattice mismatch, as well as thermal mismatch, between the epi layers would be taken into account when bonding the wafer back together, and also during subsequent process steps. The substrates were 615 μm thick, and were oriented 2° off from the (100), toward the $\langle 11\text{–}1 \rangle$ direction. These particular substrates had the large flat oriented 45° to the small flat. The emitter level metal was initially patterned on the wafer. The wafer was then cleaved in half and epoxied back together. The wafer was then subjected to various processing steps. Since we were interested in this technique for salvaging a broken wafer, we exposed the wafer to all of the processing steps which occur after the collector mesa step. This is about 2/3 of the total process. This determined if the epoxy bond was both strong enough to withstand normal wafer handling, as well as resilient to the chemicals that the wafers are normally exposed to during processing. We checked alignment of the self-aligned base metal mask, and/or the base mesa mask, in between each processing step using in a JOEL 840A scanning electron microscope (SEM) at a bias of 10 keV and a probe current of 6×10^{-10} A. A Nikon enhanced *G*-line series six-body stepper for use with 2″ substrates was employed for this study.

The bonding process for the wafer was as follows. AZ4210 resist was spun-on each piece at 2 K RPM to obtain a 3.2 μm thick layer, and baked at 110 °C for 1 h. Any resist on the cleaved edge was gently cleaned off with a Q-tip and acetone. The wafer was bonded back together using Miller-

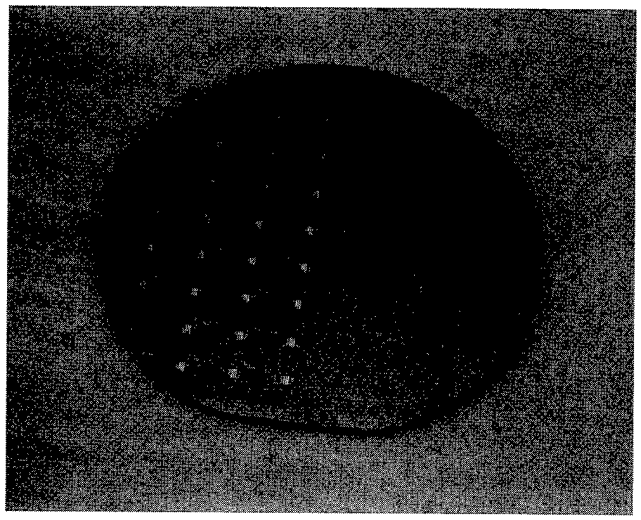


FIG. 1. Wafer after being patterned with emitter metal, cleaved, and epoxied back together.

Stephenson's Epoxy 907, with equal parts of A and B mixed just prior to application. A small bead of epoxy was applied to one edge of the cleaved wafer. The wafer was placed in a Teflon vacuum chuck next to the other half of the wafer. The two halves were gently pushed together until a small bead of epoxy appeared along the top surface to insure that there were no voids in between the pieces. A bead of epoxy formed on the bottom of the wafer as well; conveniently, it does not stick to Teflon. At this point, the vacuum chuck was placed under a microscope, and the wafer was roughly aligned to the original pattern. Then vacuum was applied to the chuck, which was placed on a hot plate at 100 °C, and allowed to cure for 12 h. After curing, the wafer was slid off of the chuck gently, and the epoxy was removed from the back side with a razor blade. The wafer was placed back on the chuck, and acetone spray was used to lift off the epoxy from the front surface.

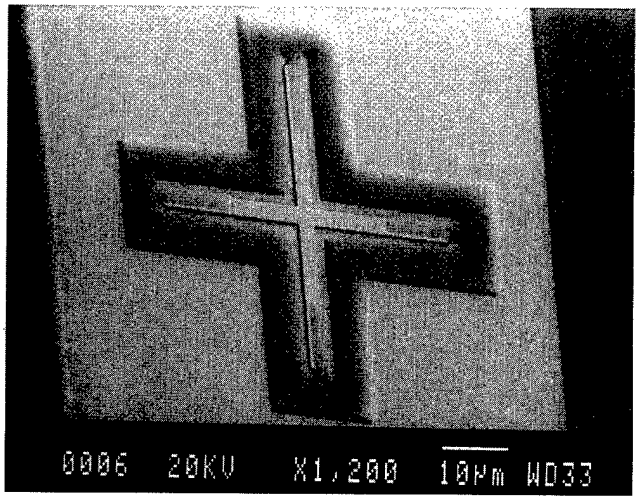


FIG. 2. Resist pattern for the self-aligned base metal in an intact field, which was aligned.

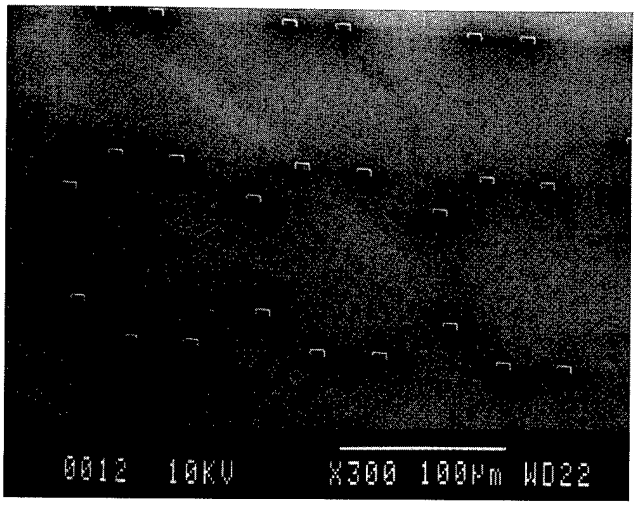


FIG. 3. Resist pattern for the base mesa in an intact circuit field, which was aligned.

III. RESULTS AND DISCUSSION

Figure 1 shows a photograph of a wafer, which has been patterned with emitter metal, cleaved, and epoxied back together. Two wafers were actually tested to insure the process was reproducible. There were two different patterns on the wafer, with 31 fields of each. Estimating from the way that the wafer was cleaved, about 12 of the fields had to be discarded, since the crack went through them. This is still an 80% yield of the original pattern on the wafer. Note that we also attempted to glue a cleaved wafer onto another substrate with the correct large flat orientation. However, the stepper does not have the depth of focus necessary to pattern such a structure, which has a total thickness of over 1000 μm.

Figure 2 shows the resist pattern for the self-aligned base metal in an intact field. The metal cross on the inside is the emitter metal. The outer cross of resist, with the tabs that intersect the ends of the inner cross, is from the base metal pattern. Alignment of these intact fields was within the step-

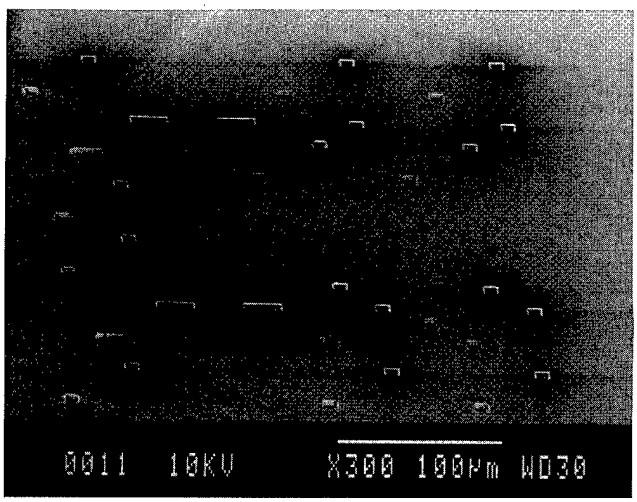


FIG. 4. Resist pattern for the base mesa in a circuit field which contained the crack, and therefore could not be aligned.

per tolerances of ± 0.125 μm registration. We also checked the alignment of the base mesa layout. Since it covers the emitter metal completely, it was easy to determine a large misalignment. Figure 3 shows the alignment of the base mesa pattern in a circuit field, where the resist is completely covering the emitter metal. Figure 4 shows the base mesa pattern in a circuit field which contained the crack, and therefore could not be aligned. Here, the resist pattern is misaligned from the emitter metal by about 20 to 30 μm.

To evaluate the durability of the epoxy, the wafer was subjected to many different processing steps to simulate normal wafer handling. In between each step, the alignment was checked by patterning the wafer with the stepper, then the resist was stripped with acetone spray. The steps evaluated were: (1) Soak in 1:10 ($HCl:H_2O$) for 5 min, rinse in H_2O; (2) soak in acetone overnight, rinse in methanol, and then isopropanol; (3) deposit two coats of spin-on glass, cure at 300 °C/10 min for each coat. Then deposit plasma enhanced chemical vapor deposition SiO_2 at 150 °C; (4) soak in buffered HF (7:1) for 1 h to remove glass and SiO_2; and (5) deposit polyimide and bake at 90, 150, and 230 °C for 1 h each. These steps simulate the process for oxide removal prior to metal deposition, metal liftoff, device encapsulation and passivation, and crossovers. These steps also show the durability of the epoxy to all of the chemicals used in our process after the collector mesa etch, as well as strain due to thermal mismatch during all of the curing cycles.

In conclusion, we have developed a technique for side-by-side bonding of two or more pieces of a wafer together with epoxy. The bonded wafer has the same dimensions, such as thickness, flatness, and position of the large flat, as a whole wafer. Since the stepper can align each field independently, the fields in the area without the actual bond can be patterned. This technique is quite useful for salvaging a broken wafer, where only the fields which contain the crack needs to be eliminated from the alignment process. The epoxied wafer endured exposure to all of the chemicals used in the last 2/3 of our process, as well as the strain due to thermal mismatch during all of the curing cycles.

ACKNOWLEDGMENTS

The authors would like to thank Y.-K. Chen and C. A. Murray for their continuing support.

[1]T. Takamori, H. Wada, and T. Kamijoh, Appl. Surf. Sci. **117-118**, 798 (1997).
[2]Y. Okuno and M. Tamura, Jpn. J. Appl. Phys., Part 2 **35**, L1652 (1996).
[3]R. F. Kopf, R. A. Hamm, R. J. Malik, R. W. Ryan, M. Geva, J. Burm, and A. Tate, J. Electron. Mater. **27**, 69 (1998).

Simple technique for measuring grating periods made using e-beam lithography

K. S. Feder[a] and D. M. Tennant
Lucent Technologies, Bell Labs, Holmdel, New Jersey 07733

(Received 30 September 1997; accepted 29 May 1998)

A simple method for measuring the period of e-beam written gratings is described. The technique allows for measurement of grating periods using a light microscope. Grating periods can be determined with a resolution of about 0.1 Å. Twenty near field holographic photoplates have been made using this method with excellent results. The technique is described for use on a JBX 6000FS electron beam lithography system with fine pitch control software, but may be applicable to other systems. © *1998 American Vacuum Society.* [S0734-211X(98)12904-9]

I. INTRODUCTION

Selected area gratings with periods ranging from 2300 to 2500 Å are needed for use in distributed-feedback and distributed Bragg reflector semiconductor laser arrays designed for emitting wavelengths near 1.55 μm. In our work, prototype laser arrays are being designed for various applications with from 4 to 16 different emitting wavelengths and therefore as many grating periods. The periods of the required gratings decrease nearly linearly in consecutive laser channels, resulting in a spatial period change of a few angstroms.

While significant errors in grating period are rare, the time and expense of laser fabrication warrants verification of grating periods. The standard technique for checking grating periods involves measuring first order diffraction angles.[1] The sample is placed on a rotation stage and illuminated with a laser. With knowledge of the laser wavelength and the angles at which the laser is diffracted the grating periods can be calculated. The method is limited by the resolution of the rotation stage and can be tedious because the small grating area (e.g., 450×40 μm) does not produce a highly visible diffracted beam. In addition the close spacing of gratings with slightly different periods can cause an overlapping of diffraction patterns. Increased laser power and lenses improve accuracy but increase measurement time. Measurements made using diffraction angles take between 30 min and 4 h depending on the number, size and quality of the gratings to be measured. This article describes a simple procedure for e-beam writing a ruler which indicates the period of gratings and can be read under a light microscope.

II. RULER FABRICATION

While our JEOL vector scan electron beam lithography system can be configured in a variety of modes, for this work it was operated at 50 kV with a minimum address step of 250 Å and a corresponding maximum scanner deflection field of 800 μm. Patterns larger than a single writing field can be stitched together with precise control. Both the pattern size accuracy and field stitching accuracy are specified to be 0.06 μm (2σ) and have been measured to be about a factor of 2 better in our system. However, the small area gratings needed for the laser arrays were individually smaller than the 800 μm scanning field and required no stitching. The laser interferometer is the metrology standard for both positioning the sample stage and for calibrating the deflection system. The He/Ne laser-based interferometer has a resolution of 50 Å.

To control the average grating period with sub 0.1 Å resolution we make use of vendor-supplied fine gain control software written for our system.[2] This program allows a fine adjustment override in the gain of the deflector amplifier. For example, gratings programmed at 2500.0 Å may be written with a deflector gain adjustment of −4.000% yielding a period of 2400.0 Å. This then allows much finer adjustments than allowed by coding patterns on a 25 nm grid. This size adjustment can be applied not only to individual fields but with equal precision to the specimen stage. This suggests a novel ruler type measurement technique described here. A reference scale or "ruler" is written on the mask or wafer outside the grating regions after the normal deflection field calibration procedure which typically yields an absolute pitch accuracy of about 0.01%. The ruler is written over an extremely large distance, typically 50 000 μm. After the gain adjustment is performed to produce the gratings, an indicator line is then written which is programmed at the same length as the ruler. The indicator is written a few microns above the scale, but with the same gain adjustment as the gratings. Due to its large size, the indicator line is shrunk a measurable distance by a small change in gain. This is illustrated in Fig. 1.

An indicator line programmed at 50 000 μm and exposed with a gain of −4% will be written 48 000 μm long. Each end of the indicator will be pulled back 1000 μm from the end of the ruler. At this scale, every 10 μm the indicator is reduced corresponds to a 1 Å reduction in the period of gratings programmed at 2500 Å. The gain adjustment acts on both ends of the indicator. If the scale is calibrated correctly, the grating period can be read at either end of the indicator as one would measure with a ruler.

Substrates smaller than 50 000 μm can also be used as only one end of the scale is needed for measurements, allowing that only the end area be written on the specimen. Figure

[a]Electronic mail: Feder@lucent.com

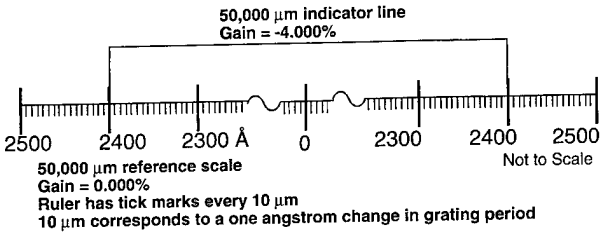


FIG. 1. Schematic of ruler measurement system. Gratings programmed at 2500 Å and exposed with a gain of −4% will have a period of 2400 Å.

2 shows a typical ruler as viewed in a light microscope. The ruler indicates grating periods of 2400.0, 2402.5, 2405.0, and 2407.5 Å. We have previously verified the method using similarly produced gratings with measured periods to within 1 Å (measurement limit) by using the diffracted angle method described above.

III. DISCUSSION AND CONCLUSION

A simple ruler method of measuring grating periods has been described. Periodically grating array measurements using the ruler technique are compared to those obtained using the first order diffraction method. To date all grating measurements using this technique have been within the experimental error of the first order diffraction method. While it relies on the inherent accuracy of the electron beam system, we believe the simplicity of this method makes it very attractive over the more fundamental diffraction method. Measurement error due to optical microscopy readings of the ruler is estimated at about ±0.1 Å. Error in marking the endpoint of the marker could be reduced by increasing the length of the reference scale, but the current level is judged suitable for this work. Despite this success, there is a possibility of a systematic error. The ruler measurement and grating periods are both calibrated by the interferometer built in the electron beam machine. The accuracy of the ruler method can be limited by the errors of the individual electron beam writing tool. In our case these are quite small, however. The two dominant sources of e-beam system errors which contribute to the ruler accuracy are the stage positioning error and the field stitching error. The first of these, the stage errors, have been measured to be less than 1 ppm corresponding to about 0.05 μm at the extrema of the ruler. The second, the field stitching errors, are specified to be 0.06 μm (2 σ) and have been measured to be less than 0.04 μm (2 σ). Since these are independent errors the total is well below 0.1 μm. Since our resolution goal is met with length changes of 1 μm, the writing errors under normal circumstances are not significant. A failure by the interferometer could cause incorrect readings. Simple fixes include checking the calibration of the e-beam machine or measuring the length of the reference scale against a known standard. Random errors due to beam drift can also occur during e-beam exposures. Such errors might be impossible to discern under an optical microscope. Thus far, the simplicity of this method outweighs these concerns.

FIG. 2. Typical ruler measurement. Grating periods measured: 2400.0, 2402.5, 2405.0, and 2407.5 Å.

[1]D. M. Tennant, T. L. Koch, P. P. Mulgrew, R. P. Gnall, F. Ostermeyer, and J.-M. Verdiell, J. Vac. Sci. Technol. B **10**, 2530 (1992).
[2]D. M. Tennant, T. L. Koch, J.-M. Verdiell, K. Feder, R. P. Gnall, U. Koren, M. G. Young, B. I. Miller, M. A. NewKirk, and B. Tell, J. Vac. Sci. Technol. B **11**, 2509 (1993).

BRIEF REPORTS AND COMMENTS

This section is intended for the publication of (1) brief reports which do not require the formal structure of regular journal articles, and (2) comments on items previously published in the journal.

Electrical characteristics of Ta_2O_5 films on Si prepared by dc magnetron reactive sputtering and annealed rapidly in N_2O

G. Eftekhari
Electrical Engineering Department, State University of New York at New Paltz, New Paltz, New York 12561-2443

(Received 23 January 1998; accepted 17 April 1998)

The electrical properties of reactive sputtered tantalum oxide annealed rapidly in N_2O ambient improves the leakage current, increases dielectric constant and electric field breakdown. These improvements are the result of incorporation of oxygen into the structure of the oxide. This is simpler than the other methods of supplying oxygen into the oxide such as annealing in oxygen plasma. © *1998 American Vacuum Society.* [S0734-211X(98)03004-2]

I. INTRODUCTION

In recent years, due to its high dielectric constant, high refractive index, and relatively small leakage current, tantalum pentoxide (Ta_2O_5) thin films have been extensively studied for applications in high density dynamic random access memories, decoupling capacitors, and antireflection coating. Tantalum oxide films can be deposited using various techniques such as anodizaiton, thermal oxidation, chemical vapor deposition (CVD), and reactive sputtering. Among these, reactive sputtering is attractive, because (1) it is a low temperature process and (2) it is compatible with the rest of the microelectronic processings. Anodization is also a room temperature process, however, it is hard to integrate this process with the other microelectronic fabrication processes.

There are numerous studies available on the properties of Ta_2O_5/Si structures regarding its chemical structure, leakage current, and dielectric constant (e.g., Refs. 1–7). In recent years, annealing in N_2O ambient has been used to improve properties of thermally grown oxide on Si (e.g., Ref. 8). The improvements are enhancements in the resistance of the oxide to radiation, hot carriers, and impurity penetration (especially horon). This has been applied to Ta_2O_5 by Sun and Chen[9,10] where oxide films are deposited using the CVD method. The purpose of this work is to study the effects of annealing of Ta_2O_5 films on Si in N_2O where the oxide films were deposited using dc reactive sputtering. Chen *et al.*[1] have studied the electrical characteristics of reactive sputtered Ta_2O_5 films. However, their study does not include the effects of annealing, especially the much preferred rapid thermal annealing. The reactive sputtering is attractive because (1) it is a low temperature process and (2) it is an easy process to perform.

II. EXPERIMENT

The (100) Si used in this experiment was p type with resistivity of about 2 Ω cm. Before loading Si samples in the vacuum system, they were cleaned and etched. The base pressure in the chamber was about 8×10^{-7} Torr. At this pressure, the argon and oxygen gases were introduced into the chamber. The pressure inside chamber during sputtering was close to 8 mTorr. The flow rates of argon and oxygen were separately controlled. The flow rate of argon over oxygen was adjusted to ~2.5. The reason for selecting this ratio is the fact that results of reactive sputtering of tantalum oxide reported by Chen *et al.*[1] indicates that good quality films are obtained at this ratio. The thickness of deposited oxide films was 15 nm. The substrate temperature during deposition was set to 150 °C. Some deposited films were subjected to rapid thermal annealing for 60 s at temperatures in the range 700–900 °C. The gate contact was aluminum evaporated on the oxide film after annealing process through a metal mask in contact with the substrate. The contacts were circular with area of 7×10^{-3} cm^2. Samples were characterized by recording current–voltage ($I-V$) and capacitance–voltage ($C-V$) measurements at room temperature. For the purpose of comparison and verifying the effectiveness of annealing in N_2O, some samples were annealed in N_2 ambient as well. The results shown in this report represent the average measurements on many samples processed at the same time. The details of current–voltage and capacitance–voltage measurements and variation of parameters among samples and processing at other Ar/O_2 flow rates will be reported in an expanded article.

III. RESULTS AND DISCUSSION

At the range of electric field (E) shown in Fig. 1, the variations of I versus $E^{1/2}$ with the current plotted on a log scale for as-deposited and annealed structures were found to be linear. The current at higher electric field was dominated by space charge limited conduction process where $I\alpha V^n$ with $n=2-3$. At lower electric field it was ohmic. The linearity

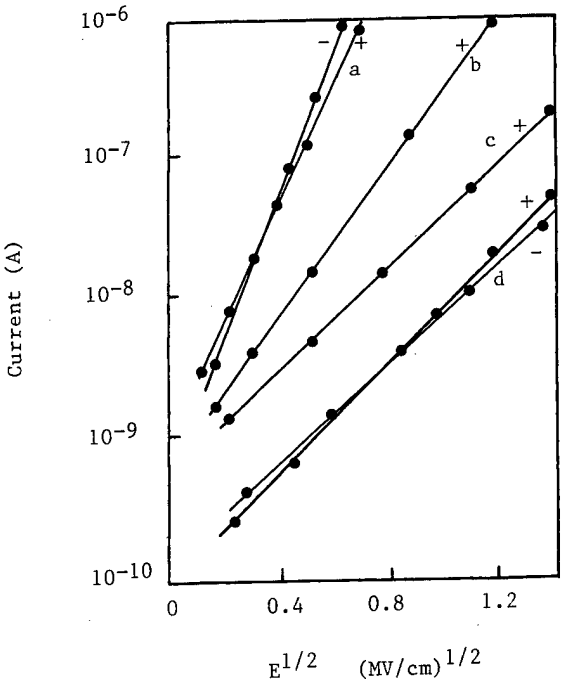


FIG. 1. Plots of I vs $E^{1/2}$ with the current plotted on a log scale for (a) as-deposited and annealed oxides at (b) 600, (c) 700, and (d) 800°C. The + and − signs show the polarity of the gate voltage. Annealing time was 60 s. For clarity in cases (b) and (c) the currents at positive polarity of gate voltage are not shown.

of this graph in the range of electric field shown in Fig. 1 indicates that current in the oxide is either dominated by Schottky emission or Poole–Frenkel (PF) emission. A clear distinction between these two current components can be made by checking the sensitivity of current to the polarity of gate voltage. In Schottky emission, the current is expected to depend on the gate voltage polarity while in PF emission it is not. In this case, it was found that current is less sensitive to the polarity of gate voltage suggesting that current is dominated by PF emission. This is shown in Fig. 1 for as-deposited samples and samples annealed at 800 °C. The transformation of ohmic conduction at low electric field to Poole–Frankel conduction at high electric field is the result of lowering the potential barrier for trapped electrons due to combination of electric field and image force.[11] The slope of PF plots is given by $\beta_{PF}=(q/\pi\epsilon_0\epsilon_r)^{1/2}q/kT$, where q is the electron charge, T the temperature, k the Boltzmann constant, ϵ_0 the permittivity of free space, and ϵ_r the dielectric constant of the tantalum oxide. The calculation of dielectric constant from the slope of PF plots gives us the values of nine for as-deposited samples and 23, 52, and 46 for samples annealed at 600, 700, and 800 °C, respectively. These values are much higher than the values obtained from capacitance measurements at accumulation (at 1 MHz) as shown in Fig. 2. The determination of dielectric constant from PF plots has been shown to be unreliable.[11] If we consider the modified form of PF emission (Ref. 11) which treats the emission process in three dimensions, then values given above should

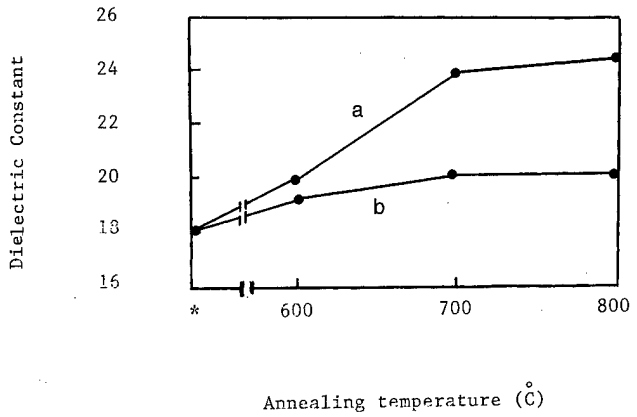


FIG. 2. Variation of dielectric constant of the oxide with annealing temperature. The symbol* indicates as-deposited oxides. Annealing time was 60 s. Curves a and b represent samples annealed in N_2O and N_2 ambients, respectively.

be divided by 4. In this case, the values for dielectric constants are considerably smaller than the values obtained from capacitance measurements.

According to Fig. 1, annealing in N_2O resulted in considerable reduction in leakage current. Since current is dominated by PF emission and in this emission, current is due to trapping and detrapping of electrons from impurity sites, therefore, one could draw the conclusion that the concentration of trap sites is reduced. Aoyama *et al.*[7] have studied the annealing behavior of CVD grown Ta_2O_5 films in oxygen plasma. Their study shows that concentrations of hydrogen and carbon in the films decrease after oxygen treatment, concluding that during oxygen plasma annealing, carbon and hydrogen impurities are oxidized and consequently the concentration of impurity or trap sites for current conduction is decreased. These impurities are unintentionally introduced into the oxide during processing and annealing. In our case, oxygen needed for this process is created by dissociation of N_2O according to $N_2O \rightarrow N_2+O$. Furthermore, any possible oxygen deficiency in the oxide could be improved by introducing oxygen into the structure of the oxide. As a result of

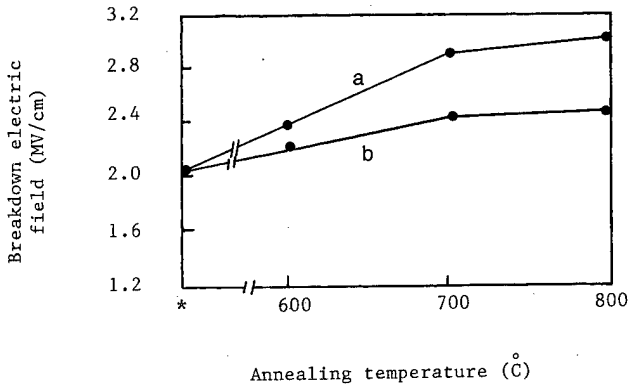


FIG. 3. Variation of breakdown electric field with annealing temperature. The symbol* indicates as-deposited oxides. Annealing time was 60 s. Curves a and b represent samples annealed in N_2O and N_2 ambients, respectively.

TABLE I. Properties of oxide before (upper values) and after (lower values) stressing.

	J(A/cm^2)	ϵ_r	E_b(MV/cm)
as-deposited $E=0.36$ MV/cm	$4\times10^{-7}/4\times10^{-6}$	18/14	2.1/1.1
annealed at 800 °C $E=1$ MV/cm	$8\times10^{-9}/5\times10^{-9}$	24/23	3/2.8

these, breakdown electric field (E_b) of the oxide and stability of the current in the oxide were also improved. The E_b of as-deposited oxide was 2.1 MV/cm and it increased to 3 MV/cm after annealing at 800 °C as shown in Fig. 3. Breakdown electric field here is defined as the electric field where the original status of the $I-V$ characteristic were not recovered as a result of structural damage to the oxide. For the purpose of comparison and demonstration the effectiveness of annealing in N_2O, some samples where annealed in N_2 ambient. Figures 2 and 3 show the results of such annealing on dielectric constant and breakdown electric field, respectively. As seen, the results are inferior when compared to the samples annealed in N_2O ambient. Matsui *et al.*[12] have used annealing in oxygen radicals to improve the properties of CVD deposited Ta_2O_5. Rapid annealing in N_2O seems interesting because of (1) simplicity of process and (2) since annealing time is short, therefore, the possibility of growing further oxide on the existing oxide is very small.

In another experiment, the effectiveness of N_2O annealing in oxide stress was examined. The as-deposited and oxides annealed at 800 °C were stressed by subjecting them to an electric field (0.36 MV/cm for as-deposited) and (1 MV/cm for oxides annealed at 800 °C) for 3 h. The results of current density, dielectric constant and E_b before and after stressing are shown in Table I. As this table shows, oxides annealed in N_2O have retained their properties after stressing while as-deposited oxides are substantially degraded.

The properties reported here for oxides annealed in N_2O are comparable and in some cases superior when compared with properties of oxides obtained using complicated methods of deposition such as regular CVD,[7] plasma enhanced CVD, PECVD[2] and electron cyclotron resonance PECVD, (ECR)-PECVD.[4] For example, Kim *et al.*[4] have obtained dielectric constant of 24 and breakdown electric field of 2.3 MV/cm for samples prepared using ECR-PECVD and annealed in O_2 for 30 min at about 800 °C and in the samples prepared by plasma-enhanced CVD, the leakage current increased at annealing temperatures 600–800 °C (annealing time was 60 s) and dielectric constant decreased in the annealing temperatures 600–800 °C.[2] Such increase in leakage current and decrease in dielectric constant with annealing temperature were not observed in this work.

IV. CONCLUSION

It has been determined that a simple method of reactive deposition of tantalum oxide when rapidly annealed in N_2O after deposition produces a high quality oxide films. The dielectric constant increases to about 24, oxide breakdown electric field increases to about 3 MV/cm, and leakage current decreases considerably. Annealing in N_2O also improves the stress properties of the oxide film.

[1]K. Chen, M. Nielsen, G. R. Yang, E. J. Rymaszewski, and T. M. Lu, J. Electron. Mater. **26**, 397 (1997).
[2]S. R. Jeon, S. W. Han, and J. W. Park, J. Appl. Phys. **77**, 5978 (1995).
[3]R. A. B. Devine, Appl. Phys. Lett. **68**, 1924 (1996).
[4]I. Kim, J.-S. Kim, O.-S. Kwon, S.-T. Ahn, J. S. Chun, and W.-J. Lee, J. Electron. Mater. **24**, 1435 (1995).
[5]R. A. B. Devine, L. Vallier, J. L. Autran, P. Paillet, and J. I. Leray, Appl. Phys. Lett. **68**, 1775 (1996).
[6]D. Laviale, J. C. Oberlin, and R. A. B. Devine, Appl. Phys. Lett. **65**, 2021 (1994).
[7]T. Aoyama, S. Saida, Y. Okayama, M. Fujisaki, K. Imai, and T. Arikado, J. Electrochem. Soc. **143**, 977 (1996).
[8]G. Eftekhari, J. Electrochem. Soc. **141**, 3222 (1994).
[9]S. C. Sun and T. F. Chen, IEEE Electron Device Lett. **17**, 355 (1996).
[10]S. C. Sun and T. F. Chen, IEEE Trans. Electron Devices **44**, 1027 (1997).
[11]K. C. Kao and W. Hwang, *Electrical Transport in Solids* (Pergamon, New York, 1981), pp. 90–92 and 314–321.
[12]Y. Matsui, K. Torii, M. Hirayama, Y. Fujisaki, S. Iijima, and Y. Ohji, IEEE Electron Device Lett. **17**, 431 (1996).

Tungsten silicide and tungsten polycide anisotropic dry etch process for highly controlled dimensions and profiles

R. Bashir,[a),b)] A. E. Kabir,[b)] F. Hebert,[c)] and C. Bracken[b)]
National Semiconductors, Santa Clara, California 95051

(Received 31 December 1997; accepted 22 May 1998)

[S0734-211X(98)09904-1]

I. INTRODUCTION

Highly anisotropic dry etching of tungsten silicide and tungsten polycides is required for the realization of submicron low resistance gates and interconnects for use in high performance complementary metal–oxide–semiconductor (CMOS) and BiCMOS technologies.[1] The current etch chemistries are not anisotropic, i.e., lateral etching of the tungsten silicide takes place which results in undesirable CD loss. In many applications a spacer needs to be formed on the polycide sidewall (Fig. 1). Undesirable undercutting can result in nonideal spacer formation for further device fabrication.

Tungsten silicide etching has been described in literature using mixtures of SF_6, Cl_2, or CF_4.[2–4] However, all these chemistries have excessive undercutting, no end point detection, and poor control of sidewall profile. It is the purpose of this brief to describe an etch chemistry using C_2F_6, Cl_2, and O_2 which forms polymer on the sidewall of the tungsten silicide during the etch in order to avoid undercutting.[5] In addition, the etch chemistry allows increase of the critical dimension by controlling the polymer deposition on the sidewall.

II. EXPERIMENTS

All etch experiments were performed in a LAM 384T Triode etcher. Power was applied to the lower electrode (RIE mode). The temperature of the electrodes was held between 15 and 30° C. Chamber pressure was maintained at 150 mT and the He backside cooling pressure was 8 T for all experiments. The C_2F_6, Cl_2, and O_2 system was explored under two different plasma power conditions. Detailed experiments were performed to determine the etch rates of WSi_x, polysilicon, and oxide under various gas plasmas and rf power.

WSi_x was deposited in a varian 3190 sputtering system from a composite target. Poly-silicon was deposited at 625 °C using low pressure chemical vapor deposition (LPCVD). The first step in the etch is an oxide etch which is to be used only if there is an insulator on top of the WSi_x. The end point can be detected when the silicide is exposed at the change in the CO emission at 450 ± 25 nm when using a CHF_3:C_2F_6 chemistry. The second step is main WSi_x etch, the characterization of which is described in detailed next. An end point can be detected at the change in the emission at 289 ± 25 nm. The third step is the poly-silicon etch which results in an endpoint at changes in emission at 289 ± 25 nm. Resist loss for OCG 825 in a typical process is about 5000 Å.

III. RESULTS AND DISCUSSION

A mixture of C_2F_6 and Cl_2 was first tried for the WSi_x etch. It was discovered that the etch builds up polymer on the WSi_x sidewall, which is bombarded and removed during the etch. Since the bombardment is anisotropic, the polymer on the sidewall is not removed and a "foot" of WSi_x will be produced after resist is stripped. As the C_2F_6/Cl_2 ratio is increased, the selectivity between the WSi_x and poly etch rates is increased but more polymers are formed on the sidewall, resulting in a tapered profile. When the C_2F_6/Cl_2 ratio is decreased, the polymer formation is also decreased but the WSi_x and poly etch rates increase at a very rapid rate due to increased chlorine radicals in the ambient. High $\%Cl_2$ mixtures are thus not desirable since a decrease in poly-silicon etch rate is needed once the WSi_x is removed. Figure 2 shows the cross sections of the stack with two different C_2F_6/Cl_2 ratios. The corresponding etch rates are depicted in Fig. 3.

Addition of O_2 in the C_2F_6/Cl_2 mixture, however, results in a decrease of the polymer buildup and a highly anisotropic, straight sidewall. The polymer can be postulated to be a tungsten-carbon based compound that reacts with oxygen to form volatile compounds. SF_6 or CF_4 based chemistries described in prior arts do not provide enough carbon to form polymers on the sidewall and hence have resulted in undercut of the silicide. Figure 4 shows the etch rate variation with change in $\%Cl_2$ in the $C_2F_6/Cl_2/O_2$ system. It is interesting to note that in our study, the over all etch rates of oxide, silicide and poly-silicon does not change much within the variation of O_2 studied as shown in Fig. 5. Excessive O_2, however, can result in excessive side etching of the resist, unwanted increase of CD, and undercutting of WSi_x. The final mixture results in an optical end point at poly-Si inter-

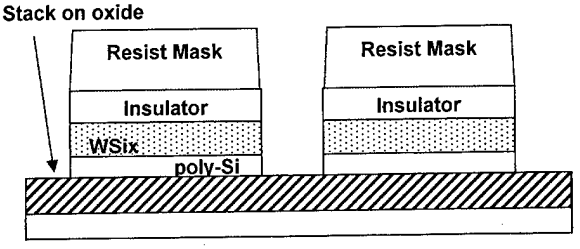


FIG. 1. Cross section of the polycide stack.

[a)]Electronic mail: rashid.bashir@nsc.com
[b)]Analog Process Technology Development.
[c)]Presently at: Spectrian, Mountain View, California.

 0734-211X/98/16(4)/2118/3/$15.00 ©1998 American Vacuum Society

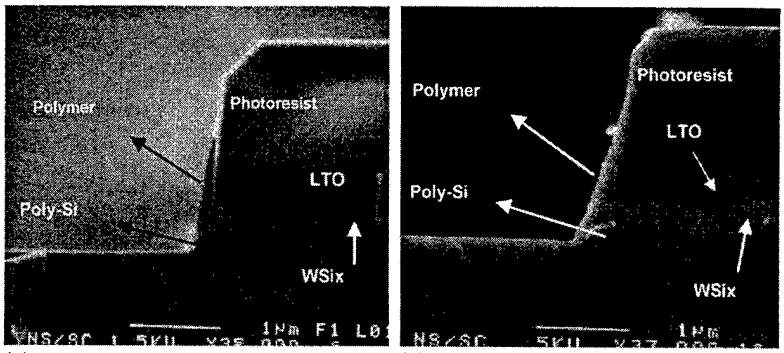


FIG. 2. (a) Stack cross section with WSi_x etch using $C_2F_6:Cl_2=85:50$, rf=400 W. (b) Stack cross section with WSi_x etch using $C_2F_6:Cl_2=100:35$, rf=400 W.

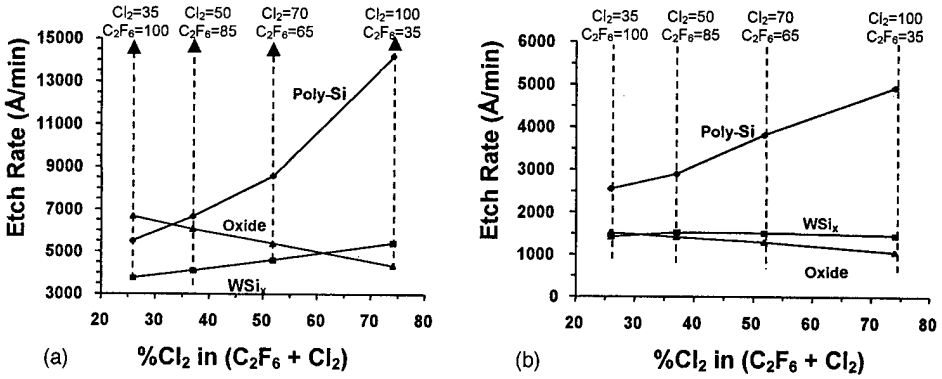


FIG. 3. (a) Etch rate variation in $C_2F_6:Cl_2$, rf=400 W. (b) Etch rate variation in $C_2F_6:Cl_2$, rf=100 W.

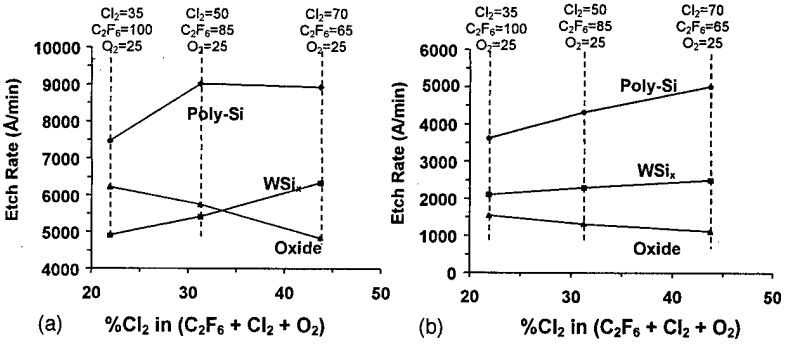


FIG. 4. (a) Etch rates for variation in %Cl at rf=400 W. (b) Etch rates for variation in %Cl at rf=100 W.

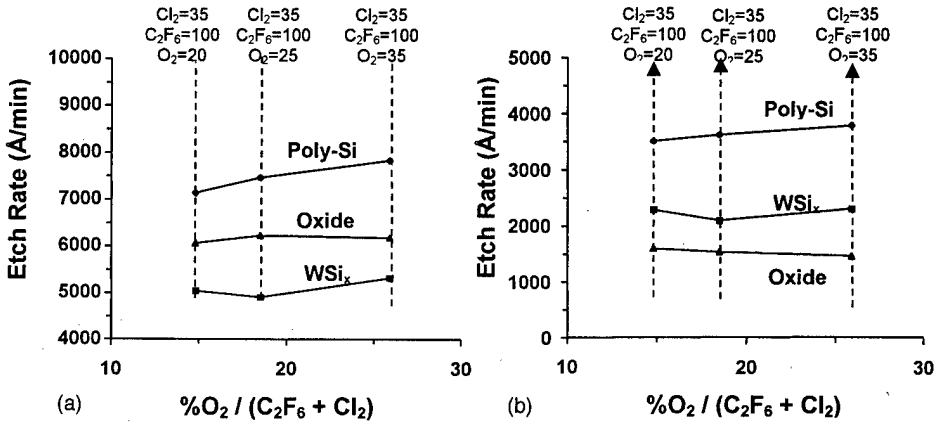


FIG. 5. (a) Etch rates for variation in $\%O_2$ at rf=400 W. (b) Etch rates for variation in $\%O_2$ at rf=100 W.

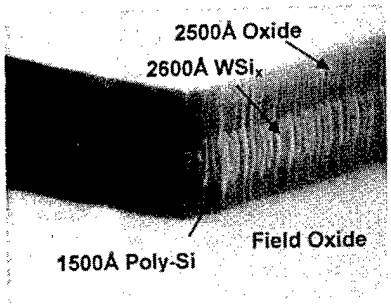


FIG. 6. Angled SEM of the stack with WSi_x etched with the final process. (WSi_x etch using $C_2F_6:Cl_2:O_2=100:35:20$ at 400 W, 150 mT, and poly-Si etch using $C_2F_6:Cl_2=85:50$ at 100 W, 120 mT.)

face. The end point is detected by a change in the emission at wavelengths of 289 ± 25 nm. There is no known reference of such an optical end point. Figure 6 shows an angled high magnification scanning electron microscopy (SEM) showing the anisotropic profile of the stack etched using the optimized WSi_x etch process with a gas mixture of $C_2F_6:Cl_2:O_2=100:35:20$ at 400 W, 150 mT, and poly-Si etch using $C_2F_6:Cl_2=85:50$ at 100 W, 120 mT.

The polymer-forming recipe can also be used to make self-aligned butted contact structures.[6] For example, simultaneous contact to the buried poly or silicide film and underlying substrate can be formed without a mask.

IV. CONCLUSIONS

A new etch chemistry has been developed to etch tungsten silicide and tungsten polycide films with high control of the sidewall profile using a $C_2F_6/Cl_2/O_2$ mixture in a LAM 384T system. The O_2 in the ambient helps control the etching of sidewall polymer and hence provides an additional degree of freedom in controlling the slope of the sidewall without significantly affecting the etch rates.

[1]M. Wong and K. Saraswat, Symposium on VLSI Technology, Ninth Symposium on VLSI Technology, 1989, pp. 101–102.
[2]W. Beinvogl, U.S. Patent No. 4,473,436 (1984).
[3]T. Hwang, S. Rogers, and M. Coe, U.S. Patent No. 4,443,930 (1984).
[4]C. Meiqiao, F. Shaoyun, and L. Jianzhon, Chin. J. Semicond. **11**, 355 (1990).
[5]R. Bashir, A. E. Kabir, and F. Hebert, U.S. Patent pending.
[6]S. Wolf and Tauber, Silicon Processing for the VLSI Era, Vol. II, p. 161.

Perturbation solution for the interfacial oxidation of silicon

Vijay R. Mhetar and L. A. Archer[a)]
Department of Chemical Engineering, Texas A&M University, College Station, Texas 77843-3122

(Received 16 October 1997; accepted 17 April 1998)

A new perturbation solution for the oxidation of silicon that properly accounts for diffusion of molecular oxygen and reaction kinetics at the silicon–silicon dioxide interface is presented. The perturbation treatment is based on the assumption of a dilute solution of O_2 in SiO_2 phase. We show that the solution obtained by Deal and Grove [J. Appl. Phys. **36**, 3770 (1965)] is the zeroth perturbation solution and formally prove that it is a quasi-steady-state solution. In the limiting case of instantaneous reaction, the perturbation solution is shown to be equivalent to the result obtained by Peng, Wang, and Slattery [J. Vac. Sci. Technol B **14**, 3316 (1996)]. The proposed perturbation scheme is developed here for first-order reaction kinetics, but can easily be extended to other rate expressions. The higher order corrections offered to Deal and Grove model predictions might provide a needed method for assessing the error incurred when these predictions are used to estimate oxide film thickness in practical situations. © *1998 American Vacuum Society.*
[S0734-211X(98)02504-9]

I. INTRODUCTION

Thermal oxidation of silicon (Si) has been studied extensively both theoretically and experimentally due to its importance in silicon-device technology. Some time ago, Deal and Grove[1] presented a simple model for the oxidation process which takes into account diffusion of molecular oxygen through the oxide film as well as reaction occuring at the silicon–silicon dioxide (SiO_2) interface. However their analysis was based on a quasi-steady-state approximation and thus neglected the movement of the Si–SiO_2 phase interface while computing the distribution of molecular oxygen (O_2), leading to a linear distribution of oxygen in the oxide phase. Although physical justification for a quasi-steady-state is tenuous, we show later that this assumption nonetheless yields the correct zero-order solution to the problem.[2]

Recently Peng *et al.*[3] analyzed the silicon oxidation problem by explicitly accounting for coupling of the concentration profile in the oxide phase with the rate of movement of the Si–SiO_2 interface. This coupling yields an inherently unsteady process. Peng *et al.*[3] assumed the oxidation process to be "diffusion controlled," however, and discarded any role of interfacial reaction kinetics. Interface reaction kinetics is believed to play an important role particularly in the growth of thin-oxide layers.[1] Also under given conditions, the rate of oxidation is known to depend on the crystal lattice orientation, an effect generally incorporated into a model through the dependence of interfacial rate constant on surface density of Si atoms.[1] Since Peng *et al.* do not account for interfacial reaction kinetics, the orientation dependence of oxidation rate is also lost. Additionally, the analytical solution provided by Peng *et al.* is valid only if there is no oxide layer present at the start up of the oxidation. This is a serious limitation because most practical situations involve multilayer oxidation schemes.

In this brief article we develop a new regular perturbation solution scheme for the oxidation of silicon that properly takes into account diffusion of molecular oxygen, interface kinetics, and the presence of an initial oxide layer prior to start up of the oxidation process. The motivation for such analysis is twofold. First, the accuracy and the range of validity of the zero-order results (Deal and Grove model) can be extended by the addition of higher-order perturbation terms. Second, utilizing the quasi-steady-state approximation as a first step in an orderly scheme of successive approximations is mathematically more rigorous than intuitively eliminating certain terms. Our perturbation scheme consists of: (a) immobilizing the moving boundary by Landau's transformation;[4] (b) change of time variable from t to $h(t)$, where $h(t)$ is the position of the moving interface at time t; (c) application of regular perturbation method.[2]

II. MODEL

The kinetics of thermal oxidation of Si is analyzed using the simple model illustrated in Fig. 1. We use the oxidation mechanism put forward by Deal and Grove.[1] In analyzing this problem we shall assume that thermodynamic equilibrium prevails at the interface and that the reaction at the interface can be described as a simple first-order reaction,

$$\mathrm{Si}+\mathrm{O_2}\rightarrow\mathrm{SiO_2}. \tag{1}$$

Assuming that the solution of O_2 in SiO_2 is dilute ($C \ll C_{SiO_2}$) and Fick's first law holds for the diffusion of O_2 in the oxide phase we write the differential mass balance[3,5] for O_2:

$$\frac{\partial C}{\partial t}=D\,\frac{\partial^2 C}{\partial z^2}, \tag{2}$$

where C and D are the concentration and binary diffusion coefficient of O_2 in the SiO_2 phase, respectively.

Jump mass balances for species O_2 and SiO_2 at the Si–SiO_2 interface require, at $z=0$,

$$N_{O_2}=\frac{r_{O_2}}{M_{O_2}}=-kC, \tag{3}$$

and, at $z=0$,

a)Electronic mail: laa@tam2000.tamu.edu

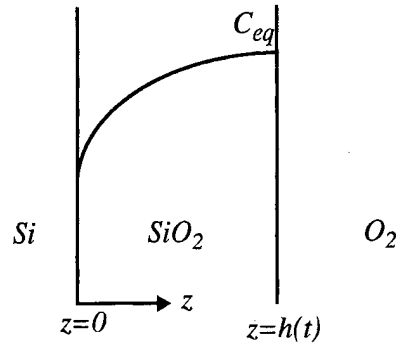


FIG. 1. Model for oxidation of silicon.

$$N_{SiO_2}=\frac{r_{O_2}}{M_{O_2}}=kC=C_{SiO_2}\frac{dh}{dt}, \tag{4}$$

where k is the reaction constant for Eq. (1). The last equality in Eq. (4) simply states that the rate of SiO_2 formed by the reaction at the interface is identical to the rate of SiO_2 grown over the Si.

Thus we wish to solve Eq. (2) consistent with following initial and boundary conditions; at $z=h(t)$,

$$C=C_{eq}, \tag{5}$$

and at $z=0$,

$$D\frac{\partial C}{\partial z}=kC=C_{SiO_2}\frac{dh}{dt}, \tag{6}$$

and at $t=0$,

$$h=h_i. \tag{7}$$

No exact solution exists for this system of equations. To facilitate a perturbation approach we introduce following dimensionless variables;

$$C^*\equiv\frac{C-C_{eq}}{C_{eq}};\quad t^*\equiv\frac{tk^2}{D};\quad z^*\equiv\frac{zk}{D};\quad h^*\equiv\frac{hk}{D};$$

$$\text{and } h_i^*\equiv\frac{h_ik}{D}. \tag{8}$$

Equation (2) along with Eqs. (5)–(7) can be reduced to the following system of equations,

$$\frac{\partial C^*}{\partial t^*}=\frac{\partial^2 C^*}{\partial z^{*2}}, \tag{9}$$

at $z^*=h^*$,

$$C^*=0, \tag{10}$$

at $z^*=0$,

$$\frac{\partial C^*}{\partial z^*}=C^*+1, \tag{11}$$

$$\frac{dh^*}{dt^*}=\left.\frac{C_{eq}\partial C^*}{C_{SiO_2}\partial z^*}\right|_{z^*=0}. \tag{12}$$

It is clear that since h^* is an unknown function of t^* the boundary conditions involving h^* [Eqs. (10) and (12)] are nonlinear. To remove this nonlinearity we introduce a Landau transformation

$$u\equiv\frac{z^*}{h^*}. \tag{13}$$

This transformation removes the nonlinearities in the boundary conditions and places them in the governing Eq. (9). From now on for the sake of clarity we shall drop the superscript *. In view of Eq. (13) the system of Eqs. (9)–(12) can be rewritten as

$$\phi\left(h\frac{\partial C}{\partial h}-u\frac{\partial C}{\partial u}\right)\left.\frac{\partial C}{\partial u}\right|_{u=0}=\frac{\partial^2 C}{\partial u^2}, \tag{14}$$

$$C(u=1,h)=0, \tag{15}$$

$$\left.\frac{\partial C}{\partial u}\right|_{u=0}=h(C+1), \tag{16}$$

$$\frac{\phi}{h}\frac{dt}{dh}=\left(\left.\frac{\partial C}{\partial u}\right|_{u=0}\right)^{-1}, \tag{17}$$

$$h(t=0)=h_i. \tag{18}$$

Here we have introduced

$$\phi\equiv\frac{C_{eq}}{C_{SiO_2}}. \tag{19}$$

In most practical cases O_2 forms a dilute solution in SiO_2, i.e., ($C_{eq}\ll C_{SiO_2}$) or $\phi\ll 1$, which suggests a perturbation approach for solving Eq. (14). To facilitate the perturbation solution we begin by expressing the dependent variable in Eq. (14) as a power series in the parameter ϕ,

$$C(u,h)=C_0(u,h)+\phi C_1(u,h)+\phi^2C_2(u,h)+\cdots. \tag{20}$$

Substituting Eq. (20) into Eq. (14) and equating terms of order ϕ^0 we obtain

$$\frac{\partial^2 C_0}{\partial u^2}=0 \tag{21}$$

with boundary conditions,

$$C_0(u=1,h)=0, \tag{22}$$

$$\left.\frac{\partial C_0}{\partial u}\right|_{u=0}=h\{C_0(u=0,h)+1\}. \tag{23}$$

The solution to Eq. (21) consistent with Eqs. (22) and (23) is

$$C_0=\frac{h(u-1)}{1+h}. \tag{24}$$

Substituting Eq. (23) in Eq. (17) yields

$$\frac{\phi}{h}\frac{dt}{dh}=\left(\left.\frac{\partial C_0}{\partial u}\right|_{u=0}\right)^{-1}, \tag{25}$$

which governs the thickness of the oxide phase as a function of time. Integrating Eq. (25) we get

$$\phi t = h + \frac{h^2}{2} - h_i - \frac{h_i^2}{2} \equiv \Gamma_0, \tag{26}$$

where h_i is the dimensionless thickness of the initial oxide layer, and ϕ is now just a parameter of known value. Using Eq. (8), Eq. (26) can be cast into dimensional form as

$$h^2 + Ah = B(t + \tau_0), \tag{27}$$

where

$$A \equiv \frac{2D}{k}, \quad B \equiv \frac{2DC_{eq}}{C_{SiO_2}}, \quad \tau_0 \equiv \frac{(h_i^2 + Ah_i)}{B}. \tag{28}$$

Equation (27) is evidently identical to the equation obtained by Deal and Grove [Eq. (12) of Ref. 1]. Notice that the zeroth-order solution $C_0(u,h(t))$ is the quasi-steady-state solution which can be obtained from Eqs. (9)–(11) by discarding the time derivative of C_0.

Next we look for the first order perturbation solution. Substituting Eq. (20) into Eq. (14) and equating terms of order ϕ^1 we obtain

$$\frac{\partial^2 C_1}{\partial u^2} = \left(h \frac{\partial C_0}{\partial h} - u \frac{\partial C_0}{\partial u} \right) \frac{\partial C_0}{\partial u} \bigg|_{u=0} \tag{29}$$

with the boundary conditions,

$$C_1(u=1,h) = 0, \tag{30}$$

$$\frac{\partial C_1}{\partial u} \bigg|_{u=0} = hC_1(u=0,h). \tag{31}$$

Solution to Eq. (29) consistent with Eqs. (30) and (31) is

$$C_1(u,h) = \frac{h^2}{6(1+h)^3} \left[3(1-u^2) + h(1-u^3) + \frac{h(3+h)(u-1)}{1+h} \right]. \tag{32}$$

Again using Eq. (17) we obtain

$$\phi dt = \frac{h\, dh}{(\partial C_0)/\partial u \,|_{u=0} + \phi(\partial C_1)/\partial u|_{u=0} + \cdots}, \tag{33}$$

which upon integration yields

$$\phi t = \Gamma_0 + \phi \Gamma_1 + \cdots, \tag{34}$$

where Γ_0 is defined in Eq. (26) and

$$\Gamma_1 \equiv \frac{1}{6} \left[\frac{h_i^2 - h^2}{2} + (h_i - h) + 2\left(\frac{1}{1+h} - \frac{1}{1+h_i} \right) + 3 \log \frac{1+h}{1+h_i} \right].$$

Equation (34) describes the thickness of the oxide layer as a function of time corrected up to the first-order perturbation and can easily be extended to include the higher orders. However due to the limitation of space we shall restrict our analysis up to the first order. The term $\phi\Gamma_1$ represents the first-order correction to the zeroth-order solution $\phi t = \Gamma_1$.

In the limit of instantaneous reaction, $k \to \infty$, and no initial oxide layer, $h_i = 0$, Eq. (34) in the dimensional form reduces to

$$h^2 = 2\phi D \left(1 + \frac{\phi}{6} + \cdots \right) t. \tag{35}$$

We shall compare Eq. (35) with the solution obtained by Peng *et al.*[3] In their analysis the thickness of the oxide layer was given as [Eq. (17) of Ref. 3]

$$h^2 = 4\lambda^2 Dt, \tag{36}$$

where λ is the solution of [Eq. (21) of Ref. 3]

$$\sqrt{\pi}\lambda \, \mathrm{erf}(\lambda) = \phi. \tag{37}$$

Expanding erf (λ) as a power series in λ we obtain[6]

$$\phi = 2\lambda^2 - \tfrac{2}{3}\lambda^4 + \tfrac{1}{5}\lambda^6 - \tfrac{1}{21}\lambda^8 + \cdots. \tag{38}$$

Inversion of the power series in λ into a power series in ϕ results in[6]

$$\lambda^2 = \tfrac{1}{2}\phi + \tfrac{1}{12}\phi^2 + \tfrac{11}{720}\phi^3 + \cdots. \tag{39}$$

From Eqs. (36) and (39) we obtain

$$h^2 = 2\phi D \left(1 + \tfrac{1}{6}\phi + \cdots \right) t, \tag{40}$$

which is identical to Eq. (35).

III. DISCUSSION

Comparing the first-order solution [Eq. (34)] with the zero-order solution [Eq. (26)] it becomes apparent that the term $\phi\Gamma_1$ represents the correction to the zeroth-order solution. For most practical situations $\phi = C_{eq}/C_{SiO_2}$ is of the order of 10^{-6}. Since the magnitude of Γ_1 is of about the same order as Γ_0 the first-order correction $\phi\Gamma_1$ is negligible when compared with the zero-order term Γ_0 and the zero-order solution should be accurate enough for most practical situations. The first-order correction becomes significant, however, at high O_2 concentrations at the O_2–SiO_2 interface, when $\phi\Gamma_1$ becomes comparable to Γ_0.

Recently Peng *et al.*[3] offered a comparison of their predictions [Eqs. (36) and (37)] with the predictions of the zero-order approximation [Eq. (26)] and experimental data at several different conditions of temperature and pressure. Comparing Eq. (27) with Eq. (40) we would expect the error in using the zero-order approximation (Deal and Grove model) to be of the order of $100\phi/6 \approx 10^{-5}\%$. Yet Peng *et al.*[3] contend that Deal and Grove model predictions were significantly different from their predictions and the experimental data. However, while estimating the parameter B of the Deal and Grove model [Eq. (28)] they did not use the same values of D and C_{eq} that were used to solve Eqs. (36) and (37). Thus we believe that the better agreement of Peng *et al.*[3] predictions with the experimental data is entirely due to accurate estimation of parameters and better available experimental values of D and C_{eq}. If the same values of parameters are used, Deal and Grove predictions are almost

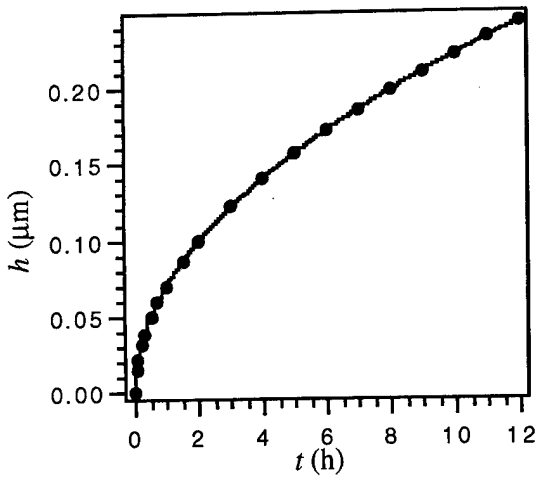


FIG. 2. Comparison of zeroth-order perturbation solution [Eq. (27)] (data points) with the prediction of Peng *et al.* (Ref. 3) (the unbroken curve). In this calculation we have taken $D=3.53\times10^{-14}$ m^2/s, $C_{eq}=7.11\times10^{-14}$ kg mole/m^3, $C_{SiO_2}=36.67$ kg mole/m^3 and $k\rightarrow\infty$. These parameters correspond to the Fig. 5 of Ref. 2.

identical to the prediction of Peng *et al.* as can be seen in Fig. 2. Moreover since Peng *et al.* ignore surface reaction kinetics in their analysis, their prediction of $h\sim\sqrt{t}$ at all times is not consistent with the experimental observation that $h\sim t$ at short times.[1] This short-time $h\sim t$ behavior is attributed to the reaction controlled regime at early stages of oxidation and is consistent with the prediction of Eq. (34).[1]

IV. CONCLUSIONS

A new perturbation solution to the oxidation of silicon is presented that properly accounts for diffusion of molecular oxygen and the interface kinetics. Dilute solution of O_2 in SiO_2 phase forms the basis for the perturbation treatment. It is shown that the Deal and Grove model is the zeroth-order approximation and formally proved that it is a quasi-steady-state solution. It is also shown that, in the limit of instantaneous reaction, the perturbation solution reduces to the recent prediction by Peng *et al.* Disagreement between Deal and Grove model predictions and the predictions of Peng *et al.* is shown to be due to inconsistency in the use of values of physical parameters D and C_{eq}.

For most practical values of physical parameters, the zeroth-order solution (Deal and Grove model) predictions are shown to be accurate enough for practical use. The perturbation scheme presented here, however, can be exploited to improve the accuracy of the predictions to the desired extent or as a guide to establish the magnitude of error in the predictions of Deal and Grove model. Although our solution is based on the assumption of first-order reaction kinetics at the $Si-SiO_2$ interface, the solution scheme can be extended to other surface reaction mechanisms that has been proposed recently.[7]

ACKNOWLEDGMENTS

V. M. thanks Professor Mark Weichold for introducing the problem during his ELEN 472 class at Texas A&M University, and Professor John Slattery, K.-Y. Peng, and Pranab Dhori for valuable discussions during the study.

NOMENCLATURE

A:	defined in Eq. (28)
B:	defined in Eq. (28)
C:	molar concentration of O_2 in SiO_2 phase
C_{eq}:	molar concentration of O_2 in SiO_2 phase at O_2-SiO_2 interface
D:	binary diffusion coefficient of O_2 in SiO_2 phase
h:	position of the O_2-SiO_2 interface
h_i:	initial thickness of the oxide layer
k:	first-order reaction rate constant for reaction (1)
M_{O_2}:	molecular weight of oxygen
N_A:	z component of the molar flux of species A
r_A:	rate of production of species A at the $Si-O_2$ interface
t:	time
u:	defined in Eq. (13)
z:	rectangular Cartesian coordinate.
Greek letters λ:	defined in Eq. (37),
ϕ:	defined in Eq. (19)
Γ_0:	defined in Eq. (26),
Γ_1:	defined in Eq. (34),
τ_0:	defined in Eq. (28).
Subscripts A:	species SiO_2 or O_2
0:	corresponds to zeroth perturbation
1:	corresponds to first perturbation
Superscript	
*:	dimensionless variables.

[1]B. Deal and A. Grove, J. Appl. Phys. **36**, 3770 (1965).
[2]M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (Academic, New York, 1964).
[3]K.-Y. Peng, L.-C. Wang, and J. C. Slattery, J. Vac. Sci. Technol. B **14**, 3316 (1996).
[4]S. G. Bankoff, Adv. Chem. Eng. **5**, 75 (1964).
[5]J. C. Slattery, *Momentum, Energy, and Mass Transfer in Continua*, 2nd ed. (Krieger, Malabar, FL, 1981).
[6]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[7]T. K. Whidden, P. Thanikasalam, M. J. Rack, and D. K. Ferry, J. Vac. Sci. Technol. B **13**, 1618 (1995).

Atomic force microscopy observation of the ferroelectric domain structure on the (010) cleavage surface of K_2ZnCl_4

D.-Y. Kim,[a] J. W. Hong, and S.-I. Kwun
Department of Physics, Seoul National University, Seoul 151-742, Korea

S.-Y. Jeong
Department of Physics, Pusan National University, Pusan 609-735, Korea

(Received 2 September 1997; accepted 10 April 1998)

The ferroelectric domain structure of K_2ZnCl_4 was studied by atomic force microscopy (AFM). The domain feature was observed on the (010) cleavage surface which is parallel to the polar *a* axis. This finding demonstrates that the image of ferroelectric domains observed by AFM is not due to the interaction between the tip and the spontaneous polarization, but due to a real deformation of the crystal surface. © *1998 American Vacuum Society.* [S0734-211X(98)00504-6]

After Saurenbach *et al.*'s pioneering study,[1] much attention has been paid to atomic force microscopy (AFM) observation of the domain structure in ferroelectric crystals.[2–7] AFM is a very useful means of studying the domain structure of ferroelectric materials because it does not require any modification of sample surfaces.[2]

The detailed mechanism by which the ferroelectric domain is visualized by AFM is not clearly understood yet. Lüthi *et al.* suggested that AFM domain images resulted from electrostatic interaction between the permanent polarization of the tip and the spontaneous polarization of the ferroelectric domain.[3] On the other hand, Kolosov *et al.* observed topographical variations corresponding to the ferroelectric domains, and asserted that the AFM domain images came from a real piezoelectric deformation of the crystal surface.[4]

To investigate the ferroelectric domain feature by AFM, single crystals of K_2ZnCl_4 (KZC) were chosen as samples. KZC has a ferroelectric phase with the spontaneous polarization along the (1, 0, 0) direction (the polar *a* axis) at room temperature, and can be cleaved only along the (010) surface.[8] It is important for this AFM study that the cleavage surface of KZC is not perpendicular but parallel to the polar *a* axis. Up to now, most AFM studies have focused on the cleavage plane normal to the spontaneous polarization, and the electrostatic interaction between the tip and the spontaneous polarization had to be deliberated.[2–7] By choosing KZC, we can observe a cleavage surface which does not have the spontaneous polarization, so that we can exclude the effect of the electrostatic interaction between the tip and the spontaneous polarization from AFM observation.

KZC single crystals of good quality were grown by the Czochralski method, and cleaved along the (010) surface with a razor blade. AFM images were obtained in the static contact mode at room temperature in an atmosphere within 30 min of the cleavage operation. (The static contact mode means a mode of imaging where the tip remains in contact with the crystal surface.) A commercially available AFM apparatus (Park Scientific Instruments, Auto Probe CP) was used for this investigation. The horizontal scanning rate was 10 μm/s. The spring constant of the boron-doped Si (Ultralever) cantilever was 0.16 N/m.

Figure 1 shows an AFM image of the ferroelectric domain on the (010) surface of KZC in the static contact mode. The image pattern in Fig. 1 is quite similar to the results of the previous AFM studies,[2–7] and also similar to the domain feature of $[N(CD_3)_4]_2ZnCl_4$, which is an isomorphus crystal of KZC, observed by the x-ray topography method.[9] (The slip in Fig. 1 is thought to be the noise effect on the AFM image.) The domain contrast did not reverse when the image was scanned in the opposite direction. In Fig. 2, a cross-sectional profile taken along the solid line of Fig. 1 can be found. It is easily seen that there exist two height levels, which are characteristic in the domain feature.[2,4]

It is very interesting that the ferroelectric domain structure can be observed by AFM even on the surface parallel to the polar axis. On the basis of Lüthi *et al.*'s assertion, the visualization of the ferroelectric domain by AFM is due to the electrostatic interaction between the tip and the spontaneous polarization, and the contrast of opposite domains is not a

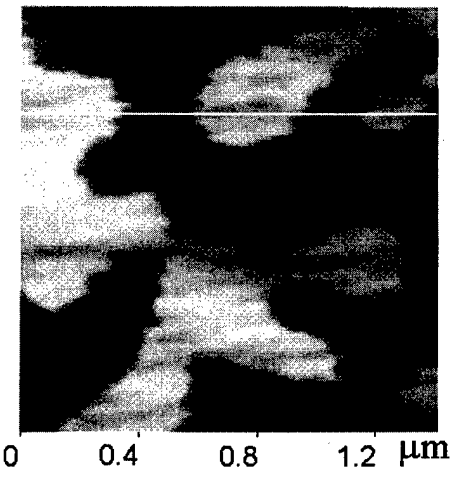


FIG. 1. Topographical image of the (010) surface of K_2ZnCl_4 observed by AFM in the static contact mode. The horizontal white solid line corresponds to a line along which a cross-sectional profile of Fig. 2 was taken.

[a]Electronic mail: rocky@phya.snu.ac.kr

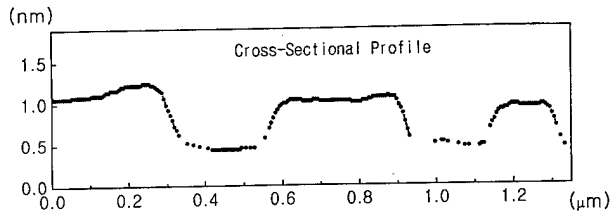


FIG. 2. Cross-sectional profile taken along the white solid line in Fig. 1. The difference in height between two opposite domains is about 0.6 nm.

real topographical variation.[3] If their explanation was accepted, the ferroelectric domain feature could not be obtained by AFM on the (010) surface of KZC because the (010) surface does not have spontaneous polarization.

An explanation for the result of this investigation can be given by Kolosov *et al.*'s piezoelectric deformation model.[4] They suggested that the AFM image of the ferroelectric domain is a real topographical variation which is ascribed to the piezoelectric deformation of the crystal surface. There exists the depolarization field E_a^* along the polar a axis inside the ferroelectric domain. Assuming the piezoelectric constant d_{aa}, the depolarization field can cause lattice deformation as $h_a=2DE_a^*d_{aa}=2DLP_ad_{aa}$.[4,10] (Here, h_a is the topographical variation along the a axis, D sample thickness, L the depolarization factor, and P_a the spontaneous polarization along the a axis.) The piezoelectric constant changes its sign with the reversal of the domain orientation, so that the ferroelectric domain structure results in the piezoelectric deformation of the crystal surface.

The domain feature on the (010) surface of KZC in Fig. 1 can be explained by introducing an off-diagonal piezoelectric constant d_{ab} to the piezoelectric deformation model. For KZC, the constant d_{ab} is about 3×10^{-12} m/V, and the piezoelectric deformation can be induced on the (010) surface by the spontaneous polarization along the polar a axis.[11]

In Fig. 2, the difference in height between the two opposite domains is about 0.6 nm, which is somewhat smaller than the results of the previous studies.[2,4] The difference is understood by the small spontaneous polarization of KZC because the piezoelectric deformation is proportional to the spontaneous polarization.[4,8] The spontaneous polarization of KZC is 0.15 μC/cm^2, which is about 1000 times smaller than that of materials studied previously.[2,4,8]

In summary, we have observed the ferroelectric domain structure on the (010) cleavage surface of KZC by AFM. Because the (010) plane does not have spontaneous polarization, our results demonstrate that the visualization of the ferroelectric domain by AFM is not due to the electrostatic interaction but due to the real deformation of the crystal surface. The deformation can be induced not only along the polar a axis but also along the b axis of KZC by the off-diagonal piezoelectric constant.

This work was supported by the Korea Science and Engineering Foundation through the Research Center for Dielectric and Advanced Matter Physics, and by the Basic Science Research Institute Program, Ministry of Education.

[1]F. Saurenbach and B. D. Terris, Appl. Phys. Lett. **56**, 1703 (1990).
[2]M.-K. Bae, T. Horiuchi, K. Hara, Y. Ishibashi, and K. Matsushige, Jpn. J. Appl. Phys., Part 1 **33**, 1390 (1994).
[3] R. Lüthi, H. Haefke, K.-P. Meyer, E. Meyer, L. Howald, and H.-J. Güntherodt, J. Appl. Phys. **74**, 7461 (1993).
[4]O. Kolosov, A. Gruverman, J. Hatano, K. Takahashi, and H. Tokumoto, Phys. Rev. Lett. **74**, 4309 (1995).
[5]L. Eng, M. Friedrich, J. Fousek, and P. Günter, J. Vac. Sci. Technol. B **14**, 1191 (1996).
[6]A. Correia, J. Massanell, N. Garcia, A. P. Levanyuk, A. Zlatkin, and J. Przeslawski, Appl. Phys. Lett. **68**, 2796 (1996).
[7]H. Bluhm, U. D. Schwarz, and R. Wiesendanger, Phys. Rev. B **57**, 161 (1998).
[8]K. Gesi, J. Phys. Soc. Jpn. **46**, 1431 (1978), and references therein.
[9]P. Saint-Gregoire, M. Ribet, and T. Tchakpele, Ferroelectrics **97**, 277 (1989).
[10]M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon, Oxford, 1982).
[11]Z. Tylczyński, Ferroelectrics **100**, 69 (1989).

In situ ultraviolet illumination of porous silicon during scanning tunneling microscopy

D. Schwall[a)]
Department of Metallurgy and Materials Engineering, University of Connecticut, Storrs, Connecticut 06269-3136

F. A. Otter
Department of Physics, University of Connecticut, Storrs, Connecticut 06269-3136

J. M. Galligan
Department of Metallurgy and Materials Engineering, University of Connecticut, Storrs, Connecticut 06269-3136

(Received 21 August 1997; accepted 22 May 1998)

Studies of porous silicon before and during ultraviolet excitation were performed using scanning tunneling microscopy. Images taken during ultraviolet excitation compared with images taken immediately prior to ultraviolet excitation show changes in the size of selective surface features. There is an increase in feature height and a decrease in feature width between the images of nonluminescing and luminescing porous silicon. These drastic effects are a consequence of an increase in available charge carriers in porous silicon, namely, in the quantum wires. © *1998 American Vacuum Society.* [S0734-211X(98)09804-7]

I. INTRODUCTION

Intense interest in the mechanism of visible light emission from silicon is rooted in the quest for silicon-based optoelectronic devices that will be compatible with the existing very-large-scale-integration technology. Suitably etched silicon,[1] called porous silicon, illuminated with ultraviolet (UV) light, photoluminesces in the visible region. Explanations for this phenomenon are centered on quantum confinement[2] of charge carriers in nanoscale surface features, chemically[3,4] and defect[5] related effects and embedded crystallites in amorphous material.[6] The quantum confinement process is expected to shift the photoluminescence (PL) spectra to shorter wavelengths as the confining feature size is reduced.[2] This relationship has been experimentally observed.[7] There is, however, a lack of experimental proof that the surface features observed are the sites of the photoluminescence. To this end, analyses of the surface morphology and electronic structure of porous silicon were performed simultaneously using scanning tunneling microscopy (STM). The porous silicon samples were analyzed while emitting visible light, providing invaluable and direct information on the visible light emission mechanism in porous silicon.

By altering the number of charge carriers present during the STM imaging process, the experiment establishes the relevant electronic structure of porous silicon. This is accomplished by measuring the effect of UV illumination on the sample between consecutive STM images. By measuring the variations in the surface features, before and during illumination, we establish those surface features in which charge carriers are confined and so contribute to the PL behavior of porous silicon.

[a)]Electronic mail: schwall@maxwell.rl.plh.af.mil

II. EXPERIMENT

P-type boron doped (111) and (100) silicon wafers (0.002–0.08 and 6–8 Ω cm, respectively) were etched in a 1:4:2 solution of hydrofluoric acid, water and ethanol.[8,9] Such etched samples are hydrogen passivated, as determined by x-ray photoelectron spectroscopy and secondary ion mass spectrometry.[7] These samples photoluminesce in the red and orange spectral regions, for (111) and (100) substrates, respectively, Fig. 1. Immediately after etching, the samples were transported in an argon purged cell to the Digital Instruments' Nanoscope II scanning tunneling microscope. The Pt–Ir STM tips were atomically sharp as determined by atomic resolution of graphite. The UV source, a 4 W handheld 365 nm lamp with an intensity of ~230 μW cm^{-2},

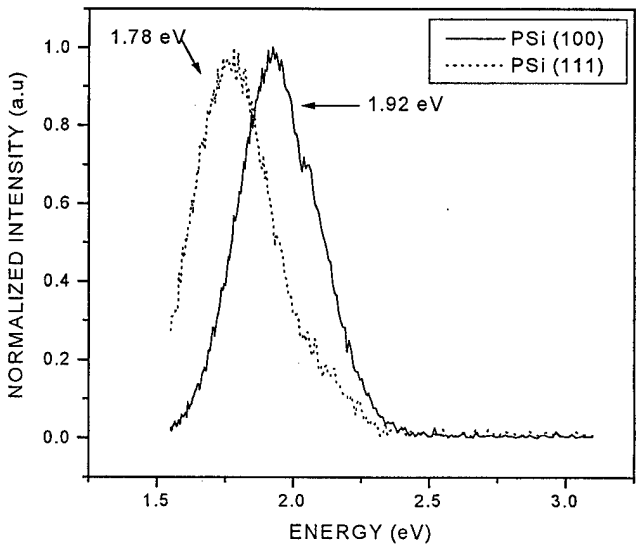


FIG. 1. Representative PL spectra of (100) and (111) porous silicon.

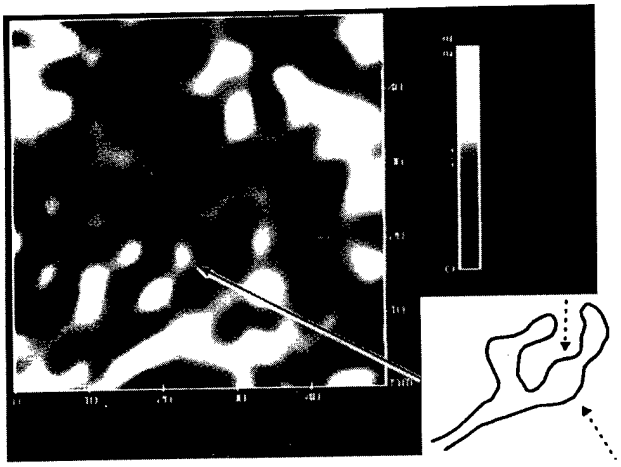
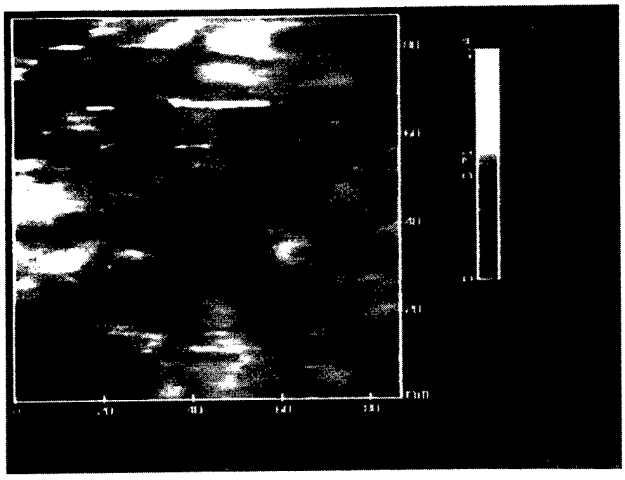
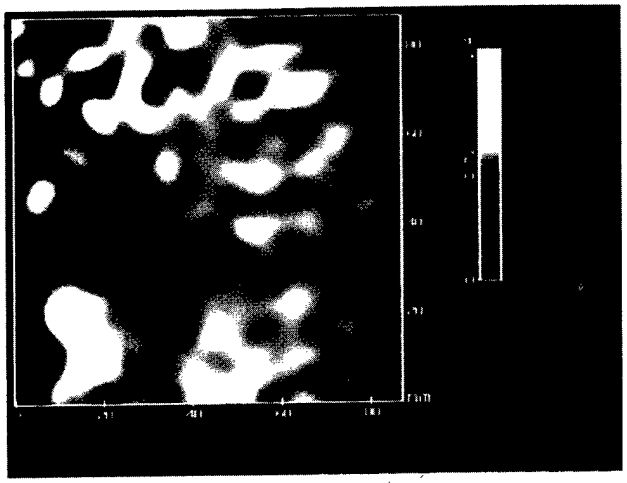


FIG. 2. Representative STM image of (100) porous silicon, inset depicts wires with nodules (50 nm×50 nm, z=2.2 nm).



(a)



(b)

FIG. 3. Consecutive STM images of (111) porous silicon (a) before UV light excitation and (b) under UV light excitation (85 nm×85 nm, z=1.4 nm).

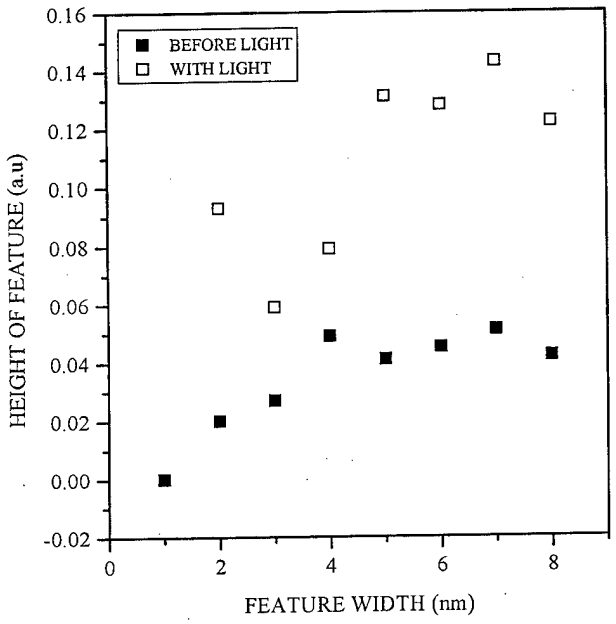


FIG. 4. Effect of *in situ* UV excitation during STM imaging on the height of the surface features as a function of width.

illuminated the surface at a distance of 3 in. The STM and the UV lamp are located inside a controlled inert gas chamber. The purged sample cell is placed in the chamber under He gas flow. The chamber is pumped out using a sorption pump to remove contaminants and then backfilled with He to create a positive pressure in the chamber.

III. RESULTS AND DISCUSSION

The morphology of the porous silicon was determined by the STM; Fig. 2 shows a typical STM image of the porous silicon surface. The relevant nanoscale features are illustrated in the inset of Fig. 2, namely, wishbone-shaped features comprised of wires with nodules. Similar features observed using atomic force microscopy (AFM)[7] are consistent with those features observed with transmission electron microscopy,[10] and are predicted by the quantum confinement model.[10] The electronic structure of porous silicon also was investigated with STM using *in situ* illumination of the sample with UV light. A series of consecutive images were then taken before and during UV excitation, Figs. 3(a) and 3(b), respectively.

Two important differences were observed as a result of UV illumination. First, there is an increase in brightness of the image, which is established by measuring the observed increase in feature height:[11] Fig. 4 shows this result for a typical experiment. This increase in brightness is solely related to an increase in tunneling current due to the presence of additional charge carriers at particular features in the image. Second, there is a shift in the measured feature width distribution such that with UV illumination the distribution has a peak at smaller feature widths, Fig. 5.[11] The observed shift in the feature width distribution, with UV illumination,

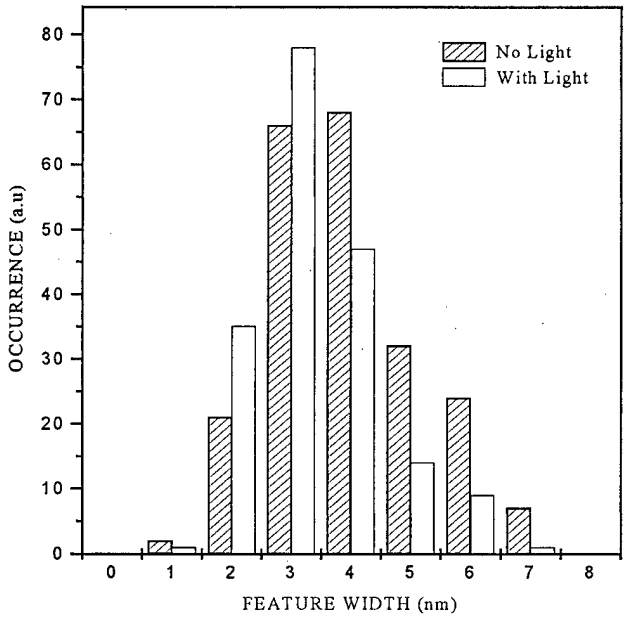


FIG. 5. Feature width distributions of (111) porous silicon from STM images before and during UV light excitation.

established that this increase in available charge carriers only occurs in the wires with nodules, such as previously observed.[7]

Similar experiments were performed on other semiconductor surfaces, graphite and HF cleaned silicon, under the same conditions to ensure that these effects were real and only occurred on the porous silicon surface. The UV light had no effect on either the graphite or silicon images. In addition, a change in the number of available charge carriers of a Pt–Ir tip would have a short lifetime, not detectable with the eye. The increase in tunneling current is, therefore, only due to an increase in the number of available charge carriers in the porous silicon.

A further enhancement in the STM images occurred when a negative bias compared with a positive bias was applied to the sample. Under negative bias, there was a larger increase in the image brightness, i.e., feature height. In this case, the tunneling electrons originated in the occupied states in the sample. With a negative bias, the current increases but, again, only in the smaller surface features such as wires with nodules. This also demonstrates unequivocally that the size of the feature specifically determines where the photoluminescence originates: charges are preferentially retained at smaller surface features, consistent with a quantum confinement mechanism.

The features, wires with nodules, observed by the STM and also the AFM[7] are consistent with the quantum confinement model.[10] The size of these features under UV excitation is consistent with the PL wavelength using an empirical second-neighbor tight-binding energy-band approximation as shown as a solid line in Fig. 6.[12] The experimental data

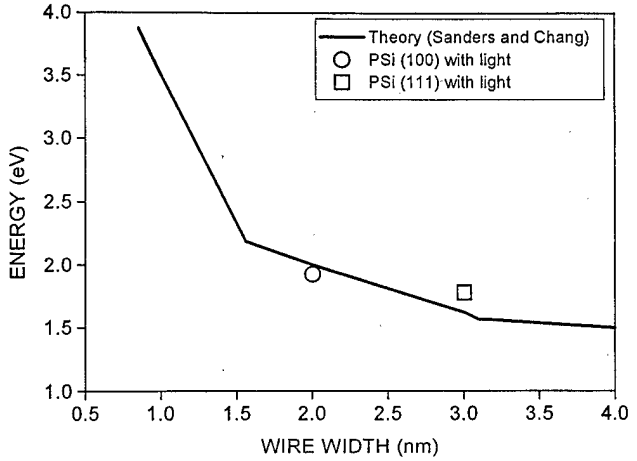


FIG. 6. Comparison of experimental data with theoretical prediction (after Sanders and Chang, Ref. 12).

points for Fig. 6 were determined by plotting the wire width corresponding to the peak in the feature width distribution (Fig. 5) versus the band-gap value for that sample from the PL data (Fig. 1). The data taken under UV light correspond to the theoretical curve; this provides further support that the visible light emission originates from these nodules.

IV. CONCLUSION

It has been shown that surface features, such as wishbone-shaped wires, trap charge carriers, leading to photoluminescence from these surface features. This is only consistent with a quantum confinement model.

ACKNOWLEDGMENTS

The authors wish to thank Motorola for donating the wafers, as well as, F. C. Jain and J. Robinson for the PL measurements. The authors acknowledge support from the Institute of Materials Science and Research Foundation at the University of Connecticut and the Connecticut Dept. of Economic Development.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]V. Lehmann and U. Gosele, Appl. Phys. Lett. **58**, 856 (1991).
[3]H. D. Fuchs, M. S. Brandt, M. Stutzmann, and J. Weber, Mater. Res. Soc. Symp. Proc. **256**, 159 (1992).
[4]R. P. Chin, Y. R. Shen, and V. Petrova-Koch, Science **270**, 776 (1995).
[5]W. E. Carlos and S. M. Prokes, J. Vac. Sci. Technol. B **13**, 1653 (1995).
[6]A. G. Cullis and L. T. Canham, Nature (London) **353**, 335 (1991).
[7]D. Schwall, F. A. Otter, and J. M. Galligan, Philos. Mag. B **75**, 887 (1997).
[8]R. Behrensmeier, F. Namavar, G. Amisola, F. A. Otter, and J. M. Galligan, Appl. Phys. Lett. **62**, 2048 (1993).
[9]G. B. Amisola, R. Behrensmeier, J. M. Galligan, F. A. Otter, F. Namavar, and N. M. Kalkoran, Appl. Phys. Lett. **61**, 2595 (1992).
[10]A. G. Cullis, L. T. Canham, and O. D. Dosser, Mater. Res. Soc. Symp. Proc. **256**, 7 (1992).
[11]The feature sizes were determined by using the STM software, each feature was measured several times and the average was computed. These averages were used to compile the distribution.
[12]G. D. Sanders and Y.-C. Chang, Appl. Phys. Lett. **60**, 2525 (1992).

Papers from the 25th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

18–21 January 1998
University Park Hotel
Salt Lake City, Utah

Sponsored by
The American Vacuum Society
The Army Research Office
The Office of Naval Research

Special Editor for the Conference Proceedings:
Jack E. Rowe

Published by the American Vacuum Society through
the American Institute of Physics, New York, 1998

ORGANIZING COMMITTEE

Jack E. Rowe
Conference Proceedings Editor

Randall M. Feenstra
Conference Chairperson

PROGRAM COMMITTEE

D. E. Aspnes (North Carolina State University)
L. R. Cooper (Office of Naval Research)
J. D. Dow (Arizona State University)
K. L. Kavanagh (University of California San Diego)
G. Lucovsky (North Carolina State University)
C. J. Palmstrom (University of Minnesota)
J. E. Rowe (Army Research Office)
H. W. M. Salemink (IBM Research Laboratories, Zurich, Switzerland)
J. S. Speck (University of California Santa Barbara)
D. J. Wolford (Iowa State University)

LOCAL ARRANGEMENTS CHAIR

C. Williams (University of Utah)

Conference Administrator

Carole Dow

PREFACE

This volume contains papers presented at the 25th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces (PCSI-25). The Conference was held on 18–22 January 1998 at the University Park Hotel in Salt Lake City, Utah. Approximately 110 persons attended the Conference, which included sessions on Oxides, Group III-Nitrides, Magnetic Materials, Compliant Substrates, Schottky Barriers, Epitaxy and Heteroepitaxy, and Novel Techniques. There were 19 invited talks and 65 short oral presentations. All presentations included posters, which were displayed for the full duration of the Conference. Time was allotted for extensive viewing of all the posters (with refreshments included), which contributed to the workshop atmosphere of the Conference.

A very interesting and informative rump session was held on Monday evening, chaired by Colin Wood of the Office of Naval Research. The main topic of the session was Lateral Epitaxial Overgrowth (LEO) of the Nitrides, a new development in the growth of GaN which permits a several order-of-magnitude decrease in the dislocation density of the films. Keynote speakers at this session were T. Zheleva from North Carolina State University and Steve DenBaars from UC Santa Barbara. Both of these groups have achieved breakthroughs in the LEO material, producing significantly improved device performance for blue light-emitting lasers made from GaN and related alloy materials. Other GaN related presentations at the Conference included a report from K. Ploog of the Paul Drude Institut in Berlin on their progress in cubic GaN growth on GaAs and Si, as well as new advances in theoretical understanding of interfaces formation for AlN on SiC and sapphire by Rosa DeFelice of Xerox Park. Other exciting advances reported at the Conference include: observation of long range order in SiO_2 (N. Herbots, Arizona State University), discussion of new magnetic semiconductors (M. Tanaka, University of Tokyo), and theoretical prediction of superlattice self-assembly during growth (J. Tersoff, IBM Research). All of the presentations were well received by the audience, and lively discussion occurred following each talk.

The Conference excursion on Tuesday afternoon was to the Alta Ski Resort. A chartered bus took about 35 of the Conference participants to the ski slope, where they all enjoyed a beautiful afternoon skiing on fresh powder snow that had fallen the previous day. A wonderful time was had by all. The Conference facilities at the University Park Hotel were quite suitable for the Conference, and many persons expressed the opinion that the Conference should be held at the same location again in a few years. The Conference succeeded through the efforts of many people. We would like to thank particularly the many anonymous reviewers who performed their work carefully and thoroughly, Jack Dow for conference management, Carole Dow for registration management, and Becky York, who provided a smooth interface between the Editor and the *Journal of Vacuum Science and Technology B*. The commitment of Gary McGuire for making this collection a part of *JVST B* is also appreciated, along with that of Galen Fisher, who provided an effective link with the American Vacuum Society. Financial support from Larry Cooper of the Office of Naval Research and Jack Rowe of the Army Office of Scientific Research are gratefully acknowledged. The Conference was held under the sponsorship of the American Vacuum Society through the Electronic and Materials Processing Division, the Office of Naval Research, and the Army Research Office.

Jack E. Rowe
Conference Proceedings Editor
Randall M. Feenstra
Conference Chairperson

What can electron paramagnetic resonance tell us about the Si/SiO_2 system?

P. M. Lenahan[a)]
The Pennsylvania State University, University Park, Pennsylvania 16802

J. F. Conley, Jr.
Dynamics Research Corporation, Beaverton, Oregon 97006

(Received 21 January 1998; accepted 21 May 1998)

Electron paramagnetic resonance (EPR) measurements of Si/SiO_2 systems began over 30 years ago. Most EPR studies of Si/SiO_2 systems have dealt with two families of defects: P_b centers and E' centers. Several variants from each group have been observed in a wide range of Si/SiO_2 samples. Some of the most basic aspects of this extensive, body of work remain controversial. EPR is an extraordinary powerful analytical tool quite widely utilized in chemistry, biomedical research, and solid state physics. Although uniquely well suited for metal–oxide–silicon (MOS) device studies, its capabilities are not widely understood in the MOS research and development community. The impact of EPR has been limited in the MOS community by a failure of EPR spectroscopists to effectively communicate with other engineers and scientists in the MOS community. In this article we hope to, first of all, ameliorate the communications problem by providing a brief but quantitative introduction to those aspects of EPR which are most relevant to MOS systems. We review, critically, those aspects of the MOS/EPR literature which are most relevant to MOS technology and show how this information can be used to develop physically based reliability models. Finally, we briefly review EPR work dealing with impurity defects in oxide thin films. © *1998 American Vacuum Society.* [S0734-211X(98)08004-4]

I. INTRODUCTION

Electron paramagnetic resonance (EPR)[1] investigations of metal–oxide–silicon (MOS) systems were begun in earnest by Nishi and co-workers[2,3] who identified a paramagnetic defect called the P_b center as a "trivalent silicon at or very near" the Si/SiO_2 interface. Nishi *et al.* argued rather persuasively that P_b centers are quite important Si/SiO_2 interface state centers in as processed Si/SiO_2 systems. Later studies refined and reinforced this conclusion.[4–12] P_b centers are silicon "dangling bond" centers dominating interface traps at the the Si/SiO_2 boundary. Studies by at least four independent groups indicate a dominating role for P_b centers in several technologically relevant device instabilities.[6–12]

Quite a few MOS oxide centers have also been identified with EPR. The most important centers are E' defects,[8,9,11–14] usually holes trapped at oxygen vacancies. At least five independent groups[8–15] have identified E' defects as dominating deep hole traps in a wide range of oxides. Quite recently, physically based models with considerable predictive power have been developed linking E' defects and molecular hydrogen to P_b dominated Si/SiO_2 interface instabilities.[15,16] In addition to the E' defects, about a dozen nitrogen, phosphorous, and boron related defects have also been identified in MOS oxide systems.[17]

Recently, Stathis and co-workers have strenuously objected to the conclusions drawn in the earlier P_b work.[18,19] On the technologically important (100) Si/SiO_2 interface, the P_{bo} center variant dominates. Stathis and Dori[18] argue that "the defect responsible for the P_{bo} resonance either is fundamentally different from a dangling bond or lies deeper inside the silicon away from the interface." Recently Cartier and Stathis[19] wrote that "prior to these studies, it was widely accepted that the silicon dangling bond defect, which gives rise to the well known P_b signal in electron spin resonance (ESR) is the microscopic defect causing the fast interface state. As will be outlined in this contribution, we cannot support this view." They go on to argue that "silicon dangling bonds, as detected by ESR measurements, account for only a small fraction of the electrically detected interface states."

Why should we think that the P_{bo} center is a silicon dangling bond? Why should we think that P_b centers play important dominating roles in Si/SiO_2 instabilities? Should we think otherwise?

In order to answer these and other questions, one might simply ask a specialist in the area. However, with a rudimentary understanding of EPR spectroscopy, one may draw conclusions for oneself. In this article we present a brief but quantitative introduction to those aspects of EPR most relevant to MOS studies and a critical review of MOS EPR studies. We show how information gleaned from EPR studies may be utilized to develop physically based predictive models of oxide reliability problems.

II. EPR AND MOS TECHNOLOGY

Advances in MOS technology have resulted in extremely complex integrated circuits with remarkably small device dimensions. With ever greater complexity and with device dimensions approaching the "atomic" scale, an approach called building in reliability (BIR) has grown in technologi-

[a)]Electronic mail: pmlesm@engr.psu.edu

cal importance. BIR involves identifying those device processing parameters that are involved in device failure phenomena and adjusting these parameters in ways that ameliorate or eliminate the failures. This approach holds great promise but for full realization, it requires physically based models of the effects of processing parameters on reliability limiting mechanisms.

With regard to MOS systems, the reliability limiting mechanisms largely involve point defects and point defect interactions: hole trapping, Si/SiO_2 interface trap generation, defect/hydrogen interactions, nitrogen, boron, and phosphorous impurity center responses. These point defects and point defect interactions are generally amenable to studies involving EPR.

Widely utilized by chemists, EPR is an analytical tool which can provide fairly detailed chemical and structural information about trapping centers.[1] It can also provide moderately precise measurements of the densities of these centers, provided that they are paramagnetic. For the relatively simple electrically active defect centers of relevance to MOS device technology, the requirement of paramagnetism is a great advantage. Most trapping centers will capture a single electron or a single hole. Thus, with the capture of either an electron or a hole, an initially diamagnetic center will be rendered paramagnetic and EPR "active." An initially paramagnetic center can be rendered diamagnetic and EPR "inactive" with the capture of either an electron or a hole. In this way, EPR can identify the response of a defect to charge carriers, measure its density, and identify its chemical and structural nature.

With a fundamental understanding of the physical and chemical nature of the defects which limit device performance, one may apply the techniques of the statistical mechanics of solids to predict and to manipulate their numbers. EPR is thus directly applicable to the development of physically based BIR models and, arguably, quite technologically useful.

III. EXPERIMENTAL TECHNIQUES

In EPR measurements, the sample under study is exposed to a large slowly varying magnetic field and a microwave frequency magnetic field oriented perpendicularly to the applied field.[1] Usually the measurements are made at an X band: a microwave frequency $\nu \cong 9.5$ GHz.

An unpaired electron has two possible orientations in the large applied field and thus two possible orientation dependent energies. (From classical electricity and magnetism, the energy of a magnetic moment $\boldsymbol{\mu}$ in a magnetic field $\mathbf{H}$ is $-\boldsymbol{\mu}\cdot\mathbf{H}$.) Magnetic resonance occurs when the energy difference between the two electron orientations is equal to Planck's constant, h, times the microwave frequency. For the very simple case of an isolated electron, the resonance requirement may be expressed as

$$h\nu = g_0 \beta_e H, \tag{1}$$

where $g_0 = 2.002319$ and β is the Bohr magneton, $eh/4\pi m_e$, where e is electronic charge and m_e is the electron mass. The Bohr magneton is 9.274015×10^{-28} J/G.

Expression (1) describes the resonance condition for an electron which does not otherwise interact with its surroundings. The structural information provided by EPR is due to deviations from this simple expression. For the relatively simple trapping centers studied in MOS systems, these deviations are due to spin-orbit coupling and electron–nuclear hyperfine interactions.

A. Spin-orbit coupling

The deviations from expression (1) due to spin orbit coupling come about because a charged particle, the electron, traveling in an electric field due to the nuclear charge, experiences a magnetic field $\mathbf{B} = \mathbf{E} \times \mathbf{v}/c^2$, where $\mathbf{E}$ is the electric field, $\mathbf{v}$ is the velocity, and c is the speed of light.[1] The spin-orbit interaction may be understood *qualitatively* (and *only* qualitatively) in terms of the Bohr picture: an electron moves about the nucleus in a circular orbit. It would appear to an observer on the electron that the positively charged nucleus is in a circular orbit about the electron. (It appears to an unsophisticated observer on earth that the sun is in a circular orbit about the earth.) The nucleus thus generates a local magnetic field which would scale with the electron's orbital angular momentum, $\mathbf{r} \times \mathbf{p}$, and with the nuclear charge. One would thus correctly surmise that spin-orbit coupling interactions increase with increasing atomic number and orbital angular momentum quantum number.

In solids, the spin-orbit interaction is "quenched" but a second order effect appears from excited states. This effect scales with the applied magnetic field and depends on the orientation of the paramagnetic defect in the applied magnetic field. The spin-orbit coupling may thus be included in the EPR resonance condition by replacing the constant g_0 of expression (1) with a second rank tensor g_{ij}. The symmetry of this tensor reflects the symmetry of the paramagnetic center. Under some circumstances, the symmetry of the tensor may permit identification of the defect under study.

Perturbation theory allows calculation (with modest accuracy) of the g tensor for the simple defects so far studied in MOS systems.[1] The components of the g tensor are given by

$$g_{ij} = g_0 \delta_{ij} - 2\lambda \sum_k \frac{\langle \alpha | L_i | k \rangle \langle k | L_j | \alpha \rangle}{(E_k - E_\alpha)}. \tag{2}$$

Here, g_0 is the free electron value, λ the atomic spin-orbit coupling constant, L_i and L_j are angular momentum operators appropriate for the x, y, or z directions, and the summation is over all excited states k. State $|\alpha\rangle$ and energy E_α correspond to the paramagnetic ground state of the system.

B. Electron–nuclear hyperfine interactions

The other important source of deviation from expression (1) is the hyperfine interaction of the unpaired electron with nearby nuclei.[1,20] Certain nuclei have magnetic moments; in metal/insulator/silicon systems, the significant magnetic nuclei are ^{29}Si (spin 1/2), ^{1}H (spin 1/2), ^{31}P (spin 1/2), and ^{14}N

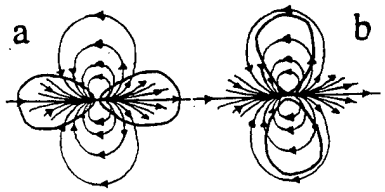


FIG. 1. Schematic illustration of an electron in a p orbital interacting with a magnetic nucleus (a) for the nuclear moment parallel to the symmetry axis and (b) perpendicular to the symmetry axis.

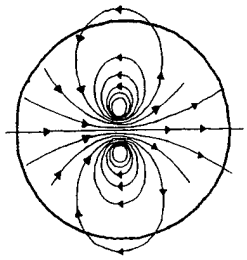


FIG. 2. Schematic of the isotropic interaction of an s orbital electron with a magnetic nucleus.

(spin 1). (Boron, with two stable magnetic nuclei, has been a minor factor in EPR studies of insulating films on silicon.) A spin 1/2 nucleus has two possible orientations in the large applied field; a spin 1 nucleus three possible orientations. Each nuclear moment orientation corresponds to one local nuclear moment field distribution.

We envision the nuclear moment interacting with an unpaired electron residing in a wave function which is a linear combination of atomic orbitals (LCAOs). For the defects of interest in MOS systems, we need only consider s- and p-type wave functions. The LCAO for an unpaired electron can be written as

$$|\alpha\rangle=\sum_n a_n\{c_s|s\rangle+c_p|p\rangle\}, \tag{3}$$

where $|s\rangle$ and $|p\rangle$ represent the appropriate atomic orbitals corresponding to the nth site, a_n^2 represents the localization on the nth site, and c_s^2 and c_p^2 represent, respectively, the amount of s and p character of the wave function on the nth atomic site.

For the most important MOS oxide and interface sites, $a_1^2\cong 1$; that is, the unpaired electron is reasonably well localized at a single nuclear site. (For all but one of the defects discussed in this article, $0.6<a_1^2<1$.) To first order then, we can interpret EPR spectra in terms of s/p hybridized atomic orbitals localized at a central site.

The electron nuclear interaction of an electron in a p orbital is anisotropic: a classical magnetic dipole interaction is schematically illustrated in Fig. 1. The interaction is strongest when the field is parallel to the symmetry axis. The sign of the interaction changes and the magnitude is decreased by one half when the field is perpendicular to the symmetry axis.

When the electron and nuclear moments are aligned by a strong magnetic field in the z direction, a reasonable assumption for work discussed in this article, only the z component of the dipolar field is important, because the interaction energy involves a dot product. $-\boldsymbol{\mu}\cdot\mathbf{H}$. This z component, $g_n\beta_n(1-3\cos^2\theta)r^{-3}$, is averaged over the electronic wave function to produce the dipolar contribution,

$$\text{Dipolar contribution}=-g_n\beta_n\left\langle\frac{1-3\cos^2\theta}{r^3}\right\rangle. \tag{4}$$

The electron–nuclear hyperfine interaction of an electron in an s orbital is isotropic. This interaction is illustrated in Fig. 2. The spherical symmetry of the orbital results in zero interaction with this field except for a spherical region about the nucleus with a radius of an imaginary current loop generating the nuclear moment's field. Since the s orbital has a nonzero probability density at the nucleus, a large isotropic interaction results. The s orbital hyperfine interaction can also be computed from an elementary electricity and magnetism calculation:[21] the magnetic field at the center of a current loop of radius a is given by $2\mu/a^3$, where μ is the magnetic moment of the current loop. The probability density of the electron varies little over the volume of the nucleus; take it to be constant, $|\alpha(0)|^2$. Considering then only the fraction of the electron wave function at the nucleus, to be $\frac{4}{3}\pi a^3|\alpha(0)|^2$, the interaction would be

$$A(\text{isotropic})=\left(\frac{4}{3}\pi a^3|\alpha(0)|^2\right)\left(\frac{2\mu}{a^3}\right). \tag{5}$$

The magnetic moment of the nucleus is the nuclear g factor, g_n, times the nuclear Bohr magneton, β_n. Thus, the isotropic or Fermi contact interaction is given by

$$A_{\text{iso}}=\frac{8\pi}{3}g_n\beta_n|\alpha(0)|^2, \tag{6}$$

where $|\alpha(0)|^2$ represents the unpaired electron probability density at the nucleus.

Both isotropic and anisotropic hyperfine interactions are present for nearly all the paramagnetic centers studied in amorphous thin films on silicon. (The sole exception is atomic hydrogen.) This is so because the unpaired electron wave functions generally involve both p-orbital and s-orbital character. The hyperfine interactions, like the spin-orbit interactions, are expressed in terms of a second rank tensor. To a pretty good approximation, the centers in these films have axially symmetric wave functions and thus an axially symmetric tensor is appropriate.

With the magnetic field parallel to the p orbital symmetry axis, the anisotropic coupling of Eq. (4) yields $(4/5)g_n\beta_n\langle r^{-3}\rangle$; the field perpendicular to the symmetry axis results in an interaction of half the magnitude and opposite sign $-(2/5)g_n\beta_n\langle r^{-3}\rangle$. This result is intuitively satisfying and consistent with the sketches of Fig. 1.

The components of the hyperfine tensor correspond to sums of the isotropic and anisotropic interactions for the applied field parallel and perpendicular to the unpaired electron's orbital symmetry axis:

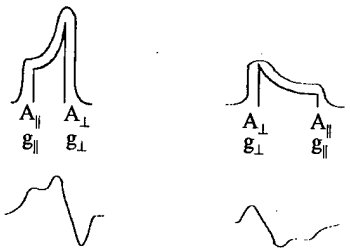


FIG. 3. Schematic sketch of the EPR spectrum of a spin-1/2 nucleus system with axial symmetry. (a) illustrates a random array of identical defects with a solid line and a random array in an amorphous matrix with a dotted line. (b) illustrates the derivative of the amorphous absorption pattern.

$$A_{\parallel}=A_{\mathrm{iso}}+2A_{\mathrm{aniso}}; \tag{7}$$

$$A_{\perp}=A_{\mathrm{iso}}-A_{\mathrm{aniso}}, \tag{8}$$

where

$$A_{\mathrm{aniso}}=\tfrac{2}{5}g_N\beta_N\langle r^{-3}\rangle. \tag{9}$$

For a paramagnetic center with a specific orientation (designated by the angle θ between the symmetry axis and the applied field vector) the resonance condition is

$$H=H_0+M_1A, \tag{10}$$

where $H_0=h\nu/g\beta_e$, and M_1 is the nuclear spin quantum number,

$$g=(g_{\parallel}^2\cos^2\theta+g_{\perp}^2\sin^2\theta) \tag{11}$$

and

$$A=(A_{\parallel}^2\cos^2\theta+A_{\perp}^2\sin^2\theta). \tag{12}$$

Equations (10)–(12) provide a very straightforward basis for analyzing EPR results for defects with a specific orientation with respect to the applied magnetic field. The "dangling bond" centers at the Si/SiO$_2$ boundary yield spectra readily described by expressions (10)–(12) since the crystallinity of the silicon substrate provides a fixed relationship between the applied field and defect orientation.

The description of EPR spectra of defects within an amorphous film is more complex. All defect orientations are equally likely and, due to the lack of long range order, slight differences in local defect geometry may be anticipated. The presence of defects at all orientations leads to the continuous distribution of both g and A values from $g_{\parallel}$ and $A_{\parallel}$ to $g_{\perp}$ and $A_{\perp}$. The differences in local geometry lead to slight defect-to-defect variations in $g_{\parallel}$, $g_{\perp}$, $A_{\parallel}$, and $A_{\perp}$.

Both of these complications are relatively easy to deal with. The random distribution of defect orientation can be dealt with easily in terms of analytical expressions found in most EPR textbooks. (For axially symmetric centers, far fewer centers will have the symmetry axis parallel to the applied field than perpendicular to it; thus the EPR spectrum intensity will be far stronger at the $A_{\perp}$ and $g_{\perp}$ values than at $A_{\parallel}$ and $g_{\parallel}$.) The slight defect-to-defect variations in g and A values lead to broadening of the line shapes anticipated for unbroadened tensor components. (The process is illustrated in Fig. 3.)

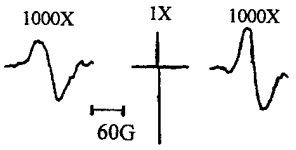


FIG. 4. EPR trace of the E' center showing a very narrow center line trace corresponding to the ≈95% abundant spin-zero ^{28}Si nuclei and two broad lines corresponding to the ≈5% abundant spin-1/2 ^{29}Si nuclei.

The evaluation of EPR hyperfine tensor components allows for a reliable and moderately precise identification of the unpaired electron's wave function.

For reasonably "clean" MOS oxides, we anticipate significant concentrations of only silicon, oxygen, hydrogen, and under certain circumstances, nitrogen, phosphorous, and boron. The nuclear moments of these atoms are all quite different.[1] Over 99% of oxygen atoms have nuclear spin zero; 95% of silicon atoms also have spin zero but about 5%, those with ^{29}Si nuclei have a nuclear spin of one half. This 95% spin zero/5% spin one half ratio is unique among elements of the periodic table. Thus, a three line pattern with two side peaks, each about 2.5% the integrated intensity of the much more intense center line, can be convincingly linked to an unpaired electron on a silicon atom. Hydrogen nuclei (99.9% of them) have a nuclear spin of one half and thus produce a two line spectrum. Nitrogen nuclei (99.6% of them) have a nuclear spin of one and therefore produce three lines of equal intensity.

A little common sense usually allows one to identify the magnetic nuclei involved in observed hyperfine interactions. Having identified the nuclear species involved, a first order analysis of the unpaired electron wave function is extremely straightforward in terms of the LCAO picture. For defects in a crystalline environment, Eq. (12) can be fit to the EPR spectrum for several values of θ. For defects in an amorphous (or polycrystalline) environment one may fit the appropriately broadened analytical expressions to the EPR spectra to yield $A_{\perp}$ and $A_{\parallel}$. (This process is illustrated in Fig. 3.) Using Eqs. (7) and (8) one then obtains the isotropic and anisotropic coupling constants A_{iso} and A_{aniso}.

Tabulated values[1] of A_{iso} and A_{aniso} calculated for 100% occupation probability can then be utilized to determine the hybridization and localization of the electronic wave functions. For example, the isotropic and anisotropic coupling constants for an electron 100% localized in a silicon s and p orbital are, respectively, $a_o=1639.3$ G and $b_o=40.75$ G. In Fig. 4, we illustrate an EPR trace of the E' center, the dominating deep hole trap in high quality thermally grown oxides on silicon. An application of the analysis schematically indicated in Fig. 2 indicates that $A_{\mathrm{iso}}=439$ G and $A_{\mathrm{aniso}}=22$ G. If the electron were 100% localized in a silicon s orbital, we would expect an isotropic coupling constant of $a_o=1639.3$ G. We measured 439 G; thus the orbital has $439/1639\cong 27\%s$. If the electron were 100% localized in a silicon p orbital we would expect $A_{\mathrm{aniso}}=b_o=40.75$ G. We measured $\cong$ 22 G; thus, the orbital has $22/40.75\cong 54\%p$. The analysis indicates a localization on the center silicon of

about (54+27)=81%. Although the crude analysis just discussed is not extremely precise it is, to first order, *quite* reliable. One should realize that the isolated atomic values obtained for a_o and b_o are themselves only moderately accurate and that placing a silicon atom in an oxide matrix will inevitably alter the constants somewhat. Nevertheless, a straightforward analysis of hyperfine parameters provides moderately accurate measurement of hybridization and localization.

IV. MEASUREMENT OF DEFECT DENSITIES

A. Accuracy and sensitivity

The EPR is typically measured by placing the sample and a calibrated spin standard in a high Q microwave cavity. EPR is detected via changes in Q: Comparing sample and standard responses, relative defect densities can be determined to a precision of better than $\pm 10\%$; absolute precision is better than a factor of 2. With considerable effort, as few as $\cong 10^{10}$ defects/cm^2 may be observed using standard EPR. As discussed later in the text, the EPR detection technique called spin dependent recombination is about $\sim 10^7$ times more sensitive than conventional EPR. Unfortunately quantitative spin counting measurements are not yet possible via spin dependent recombination.

V. PARAMAGNETIC CENTERS IN MOS SYSTEMS

A. Si/SiO_2 interface defects: P_b centers

1. Analysis of the structure

The chemical and structural nature of P_b centers has been established by several independent, consistent, and mutually corroborating studies. Three P_b variants have been consistently observed: at (111) Si/SiO_2 interfaces a defect called simply P_b, at (100) Si/SiO_2 interfaces two defects called P_{b0} and P_{b1}. The structure of both the (111) Si/SiO_2 P_b and the (100) Si/SiO_2 P_{b0} are reasonably well understood; only a rudimentary understanding of P_{b1} exists at this time.

P_b centers were first observed by Nishi and co-workers[2,3] in a study initiated more than 30 years ago. Their work focused primarily on the (111) Si/SiO_2 system. They showed that the P_b centers were at or very near to the Si/SiO_2 interface and that they possess an axially symmetric g tensor $g_\parallel \cong 2.000$ and $g_\perp \cong 2.01$ with the symmetry axis corresponding to the (111) direction. On the basis of this information Nishi *et al.* concluded that P_b centers are "trivalent silicon" centers at or very near the Si/SiO_2 boundary.

Later, Poindexter *et al.*[4,5] obtained more precise g tensor parameters. For the (111) Si/SiO_2 P_b, they found $g_\parallel \cong 2{,}0014$ and $g_\perp \cong 2.0081$ with, as Nishi had observed previously, the symmetry axis corresponding to the Si(111) direction. For the (100) Si/SiO_2 system, they found two P_b variants; P_{b0} and P_{b1}. Although hampered by overlapping spectra, Poindexter *et al.* were able to show that the P_{b0} g tensor was virtually identical to that of the (111) Si/SiO_2 P_b; they obtained $g_\parallel = g_3 = 2.0015$ and $g_\perp \cong g_1 \cong g_2$, with $g_1 = 2.0087$ and $g_2 = 2.0080$. This very close correspondence between the (111) P_b and (100) P_{b0} g tensors was confirmed in later observations by Kim and Lenahan.[9] Poindexter *et al.*[4,5] found that P_{b1} exhibits lower symmetry: $g_1 = 2.0076$, $g_2 = 2.0052$, and $g_3 = 2.0012$. Traces of the P_b center taken with the magnetic field parallel and perpendicular to the (111) symmetry axis are shown in Fig. 5.

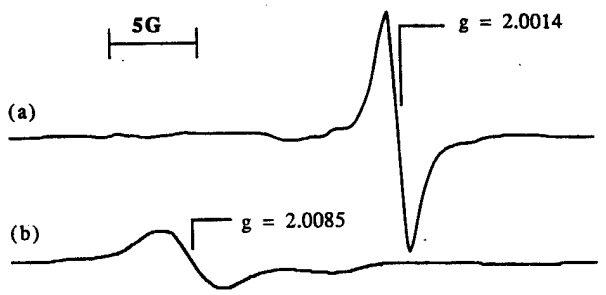


FIG. 5. EPR traces of the P_b center for the magnetic field (a) parallel and (b) perpendicular to the (111) symmetry axis.

The close similarity between P_b and P_{b0} g tensors lead Poindexter *et al.* to suggest that the two centers are essentially identical: silicon "dangling bond" defects in which the unpaired electron resides on a silicon backbonded to three other silicon atoms at the Si/SiO_2 boundary. This identification makes physical sense: a silicon dangling bond defect *should* have a g tensor with (111) axial symmetry. [Silicon bonds point in (111) directions.] Also, as pointed out by Caplan, Poindexter and co-workers, the g tensors are consistent with those of other silicon dangling bond centers.[4] For a (111) Si/SiO_2 interface this g tensor symmetry axis should be normal to the interface (it is). Serious doubts about the basic structure of the (111) P_b and (100) P_{b0} centers should have been resolved by measurements of the hyperfine interactions with ^{29}Si nuclei. The (111) P_b hyperfine tensor was first measured by Brower[22] who found $A_\parallel \cong 152$ G and $A_\perp \cong 89$ G. The similar (100) P_{b0} hyperfine tensor was first measured by Jupina and Lenahan[23] in radiation damage studies and later by Gabrys and Lenahan[24] in hot carrier damaged transistors and (probably) quite recently by Cantin *et al.*[25] in porous silicon films. Brower's (111) results have also been confirmed by Jupina and Lenahan[23] and (probably) Cantin *et al.*[25] In all cases, the approximately 5%/95% ratio of side (^{29}Si) peaks to center (^{29}Si) peaks was reported. The $\cong$5%/95% ratio of "hyperfine lines" to center line intensity for both (111) P_b and (100) P_{b0} *unequivocally* establishes that both centers are *silicon* dangling bonds. The (111) P_b and (100) P_{b0} have nearly identical hyperfine tensors which demonstrate, as had been indicated by earlier g tensor results, that P_b and P_{b0} are essentially identical defects. [Of course quite subtle differences *inevitably* exist; the (111) P_b center's unpaired electron is in an orbital directed along the (111) surface normal.] The (100) P_{b0} is also directed along a (111) direction but this direction is obviously not normal to the (100) surface. The structures of the (111) P_b and the (100) P_{b0} are shown in Figs. 6 and 7.

The *detailed* structure of the P_{b1} center remains something of a mystery; although it too is clearly a silicon "dan-

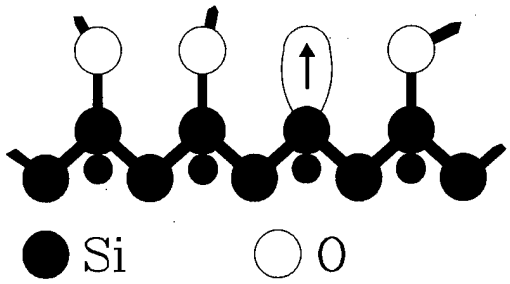


FIG. 6. Schematic illustration of a P_b center at the (111) Si/SiO_2 interface.

gling bond'' at the interface. Brower observed the P_{b1} ^{29}Si hyperfine interactions for a single orientation.[26] His results clearly indicate an unpaired electron localized on a silicon atom in an orbital of high p character. Brower observed P_{b1} (as well as P_{b0} and P_b) in oxides grown on both (111) and (100) silicon substrates in a ^{17}O enriched atmosphere. Since ^{17}O possesses a nuclear magnetic moment, if the P_b center silicon were bonded to an oxygen, the ^{17}O nuclear moment would greatly broaden the spectrum. None of the three P_b variants is broadened enough to indicate nearest neighbor oxygens. However, *all* are slightly broadened, indicating that they are *all at the interface*. Since nitrogen and hydrogen also possess magnetic moments we know that the P_{b0} and P_{b1} silicons could not be bonded to them either. Since no other atoms are consistently present in numbers sufficient to account for the P_b centers we may conclude that *all* three varieties of P_b center are silicons back bonded to silicons at the respective Si/SiO_2 boundaries.

It should be noted that Stathis and Dori[18] strenuously object to the conclusions above regarding both the structure of P_{b0} and its close similarity to the (111) P_b. They argue that ''the defect responsible for the P_{b0} resonance either is fundamentally different from a dangling bond, or lies deeper inside the silicon away from the interface.'' They also argue that ''the structure of P_{b1} is a silicon dangling bond similar to the P_b on (111) and that P_{b0} is a *fundamentally* different defect.''

Stathis and Dori[18] draw these conclusions from an experiment for which they reported on two oxide samples—one grown on a (111) surface and one on a (100) surface. The oxides were exposed to a post oxidation anneal in argon at 700 °C in a sealed oxygen free silicide anneal furnace (SAF). The (111) samples received 5 min anneals, the (100) samples received 30 min anneals. After the anneal in the SAF both samples were placed in a furnace in an argon ambient at 760 °C and then pulled out after 30 s. Some samples received a blanket aluminum deposition after the annealing steps ''to ensure that these samples saw the same processing sequence as similar samples that were used for electrical characterization.'' (Stathis and Dori did not discuss these electrical characterizations.)

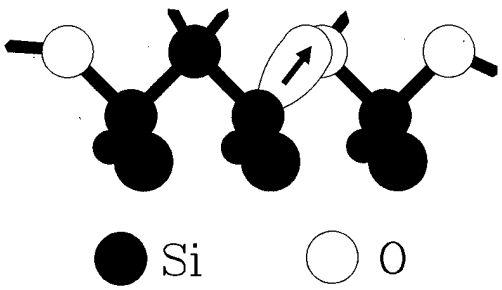


FIG. 7. Schematic illustration of a P_{b0} center at the (100) Si/SiO_2 interface.

The Stathis/Dori conclusions with regard to P_{b0} not being a dangling bond, for example, apparently arise from differences they observed in the P_b, P_{b0}, and P_{b1} responses to the growth and annealing steps which were themselves somewhat different for the two samples. What conclusions may be drawn from the Stathis/Dori work are not obvious. However, in view of the enormous amount of consistent and mutually corroborating evidence generated by many other groups investigating a vastly wider range of oxides, one must conclude that the overwhelming preponderance of evidence indicating that the (111) P_b and (100) P_{b0} are virtually identical defects is convincing. Both defects are silicon dangling bonds; in this regard the hyperfine results are utterly conclusive. Both involve silicons backbonded to three other silicons and both are at the Si/SiO_2 boundary.

B. Electronic levels of P_b centers

Several measurements of the population of paramagnetic P_b density versus Si/SiO_2 interface Fermi level provide approximate but unequivocal information about the center's electronic density of states.

The first attempt at this measurement by Poindexter and co-workers[27] was not successful. They were able to show that the P_b center spin lattice relaxation time was gate bias dependent but were unable to determine whether or not the charge state is also bias dependent. (The spin lattice relaxation time is the time it takes for an electron ''flipped'' via EPR to return to its previous orientation in the magnetic field.)

Another attempt at this measurement by Brunstrom and Svenson,[28] although qualitative in nature, showed that the application of a very large positive or negative bias across the oxide could suppress much of the P_b amplitude. From this observation, Brunstrom and Svenson inferred that P_b centers can accept both electrons and holes.

The first *semiquantitative* and *quantitative* information about P_b levels within the Si band gap was reported by Lenahan and Dressendorfer[7,8] who measured the ESR amplitude of the unsaturated P_b absorption spectra as a function of the interface Fermi level using a TE_{104} double resonant cavity and a calibrated spin standard. Their results are shown in Fig. 8. Investigating irradiated thermal oxides, they found that the distribution of P_b centers is broadly peaked. Below mid-gap the P_b center is a donorlike interface state defect ($P_b = P_b^+ + e$; $P_b + h^+ = P_b^+$). As the Fermi level moves toward mid-gap, the positively charged P_b center accepts an electron and becomes paramagnetic and neutral ($P_b^+ + e^- = P_b$). As the Fermi level moves from mid-gap to the conduction band edge, the P_b center picks up another electron,

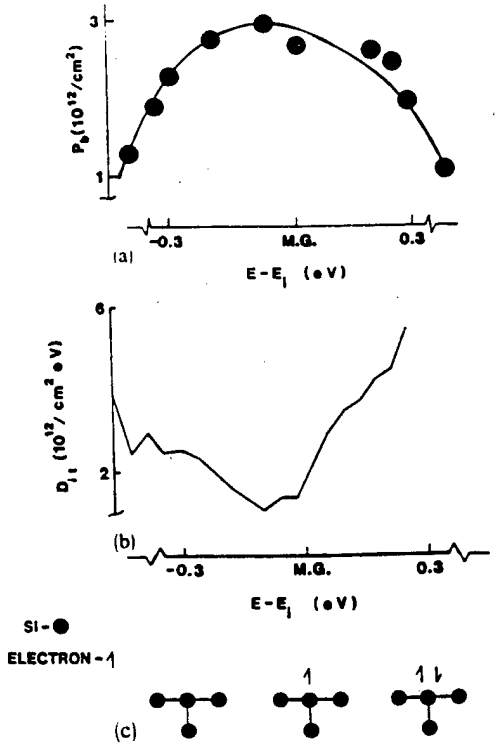


FIG. 8. (a) Population of paramagnetic P_b centers vs the position of the Fermi energy at the Si/SiO$_2$ boundary (b) Density of interface states. (c) Schematic illustration of P_b occupation vs Fermi energy.

becoming negatively charged and again diamagnetic. In the upper part of the band gap, the P_b center is thus an acceptorlike interface trap ($P_b + e^- = P_b^-$; $P_b^- + h^+ = P_b$). The P_b center is paramagnetic when it has an unpaired (one) electron; it becomes diamagnetic by either accepting or donating an electron. The distribution of paramagnetic P_b centers is consistent with the "U-shaped" distribution of interface traps that is generally reported. Lenahan and Dressendorfer proposed that the P_b density of states could be roughly approximated by the absolute value of the derivative of the P_b versus energy curve, yielding a U-shaped distribution. The spin state of the P_b center corresponding to its rough position in the gap is also displayed in Fig. 8.

Lenahan and Dressendorfer's results showed that P_b centers are amphoteric: P_b centers are positively charged and diamagnetic when the Fermi level is near the valence band and negatively charged and diamagnetic when the Fermi level is near the conduction band. When the Fermi level is near the center of the gap, P_b centers are paramagnetic and have no net charge.

Lenahan and Dressendorfer pointed out that if P_b centers account for most of the interface traps, then one can separate the effects of oxide space charge from interface charge by measuring capacitance versus voltage shifts corresponding to the Fermi level at mid-gap, where the P_b centers are electrically neutral.

Poindexter *et al.*[29] subsequently confirmed this P_b versus energy result. They published an essentially identical energy distribution plot for P_b in unirradiated high temperature as processed oxides. A recent study by Semon and Lenahan[28] also confirms the original results; they found that the distribution of P_b centers in a wide variety of oxides is roughly the same and that the electron–electron correlation energy of the P_b is about 0.7 eV. Gerardi *et al.*[30] established that the energy levels and correlation energy for P_{b0} are nearly identical to the P_b (111). They also report that the P_{b1} has a smaller correlation energy than the P_{b0} and a distribution skewed towards the top of the gap. Kim and Lenahan[9] later obtained P_{b0} versus Fermi energy results nearly identical to those off Gerardi *et al.*; they also showed that the P_{b0} center (and not the P_{b1}) is primarily responsible for ^{60}Co radiation induced interface traps (D_{it}).

Although Lenahan and Dressendorfer and the Poindexter group published virtually identical P_b amplitude versus energy curves, they interpreted them in slightly different ways. Whereas Lenahan and Dressendorfer proposed that the absolute value of the derivative of P_b amplitude versus energy $|dP_b/dE|$ would *roughly* correspond to the P_b density of states, Poindexter *et al.* presented a plot of $|dP_b/dE|$ *as* the density of states. Although Poindexter *et al.* also clearly viewed this as an approximation, the idea that $|dP_b/dE|$ represents the P_b density of states has caused some misunderstandings. The $|dP_b/dE|$ curve goes to zero near mid-gap; if $|dP_b/dE|$ very accurately represented the P_b density of states, there would be zero P_b levels at mid-gap. It is clear that this is not the case from at least two types of measurements: (1) spin dependent recombination studies of irradiated and hot carrier stressed metal–oxide–semiconductor field-effect transistors (MOSFETs) and (2) a comparison of mid-gap interface state densities and P_b densities in many as processed (and relatively poor) Si/SiO$_2$ samples.

In spin dependent recombination (SDR)[31] one detects EPR via spin dependent changes in a recombination current in a semiconductor device. SDR was discovered by Lepine,[31] who showed that the spin dependent capture of charge carriers at paramagnetic deep levels could be modulated via EPR and that EPR induced changes in capture probability could be measured in a recombination current. Lepine proposed a very simple model which, although not completely correct, provides qualitative understanding. A strong applied magnetic field polarizes the spins of both the paramagnetic trapping centers and the paramagnetic charge carriers. With both trap and charge carrier spin systems polarized, imagine the capture of an electron at a P_b center dangling bond site. If both the P_b center and conduction electron spin point in the same direction, the recombination event will be forbidden. Thus, the application of a large magnetic field will reduce charge carrier capture and thus will also reduce a recombination current. However, in EPR, the application of a microwave frequency field satisfying the resonance condition will "flip" the trap electron. The trapping event is no longer forbidden and the recombination current increases. Thus, by measuring the recombination current versus applied field while simultaneously applying microwave radiation, one may identify the defects dominating recombination. The current versus field SDR spectrum will almost exactly match the

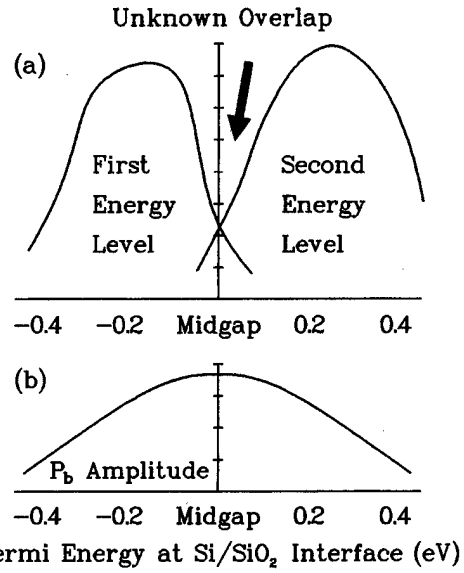


FIG. 9. Schematic density of states for P_b centers.

EPR spectrum of the defect (or defects) involved in recombination.

Grove, Fitzgerald, and co-workers[32–34] extensively investigated recombination currents in gate controlled MOS diodes. They showed that near mid-gap interface states are almost entirely responsible for recombination currents when these devices are biased to yield the maximum recombination current. Both the pioneering SDR study of Vranch *et al.*[10] and the later more extensive SDR measurements of Jupina and Lenahan[23] reported quite strong P_b SDR spectra under these biasing conditions, results essentially impossible to reconcile with a near zero P_b density of states in the vicinity of mid-gap.

Almost equally compelling objections to the notion of near zero P_b density at mid-gap come from an examination of the early collaboration between Caplan, Poindexter, Deal and Razouk (CPDR).[4,5] In the CPDR studies P_b and mid-gap interface state densities were evaluated on a variety of as-processed Si/SiO$_2$ samples with relatively high mid-gap interface state densities in the range of $\sim 1\times 10^{11}/\text{cm}^2$ eV to $2\times 10^{12}/\text{cm}^2$ eV. (CPDR reported a virtually one-to-one correlation between the P_b densities and 1 eV times the mid-gap interface state densities.) If the P_b density of states were virtually zero, Poindexter *et al.* could not have obtained the results they reported. (Obviously, if the P_b densities at mid-gap were very low, the mid-gap interface trap density /P_b density ratio *has* to be very high.)

A schematic illustration of the P_b density of states is shown in Fig. 9. As recently demonstrated for P_b-like dangling bonds in amorphous hydrogenated silicon,[35] this schematic curve should be viewed as exactly that—a *schematic*. The shape of the density of states curve will depend to some extent upon the density of P_b centers.[35] The P_b energy levels couple to the band tail states which they themselves (in part) create. P_b defects close to one another will have a coupling between their own levels.[35] Nevertheless the schematic illustration is a reasonable zero order description of P_b levels. Two broad levels separated by a correlation energy of ~0.6 eV; a quite significant P_b density of states exists at mid-gap.

VI. ROLE OF P_b IN Si/SiO$_2$ INTERFACE INSTABILITIES

The Si/SiO$_2$ interface can be damaged when MOS devices are stressed in a variety of ways. Among the stressing phenomena, ionizing radiation has been the most extensively investigated, although a considerable body of work also exists regarding hot carrier and high electric field effects.

A. Radiation damage

Several early studies by Lenahan, Dressendorfer *et al.*[6–8,36] using (111) Si/SiO$_2$ structures, established that when MOS devices are subjected to ionizing radiation, P_b centers are generated in densities which approximately match the average of the interface state densities generated in the mid-half band gap. They showed that the annealing characteristics of the P_b centers and radiation induced interface state densities are virtually identical. They furthermore demonstrated that the densities of radiation induced P_b centers could be strongly influenced by processing variations. Processing yielding low P_b generation also produced low yields of radiation induced interface state densities. Processing yielding higher P_b densities resulted in proportionately larger interface state densities. A later study by Kim and Lenahan[9] extended these results to the more technologically important (100) Si/SiO$_2$ system. Some results from these studies are shown in Figs. 10 and 11. As mentioned previously, an interesting aspect of the Kim/Lenahan study was the observation that the radiation induced P_b centers were primarily (although not exclusively) P_{b0} defects. (Weaker P_{b1} spectra were also generated.) The results of these studies have been confirmed and extended by many other groups.

Miki *et al.*[11] compared the response of ultradry and steam grown oxides to ionizing radiation. They found higher densities of radiation induced interface states and higher densities of radiation induced P_b centers in the steam grown oxides. They also found a rough numerical correspondence between the ratios of induced P_b and interface state defects as well as in the absolute numbers of P_b centers and interface states. However since the EPR measurements were made on soft x-ray irradiated devices and the electrical measurements were made on devices in which holes were avalanche active injected into the oxide (to simulate the radiation) precise numerical comparisons were not possible.

Awazu *et al.*[12] have also studied the role of processing parameters on the generation of P_b centers by ionizing radiation. Among the oxide processing parameters investigated were those used in the first study of Lenahan and Dressedorfer. For these oxides, Awazu *et al.* obtained the same P_b versus dose curve reported in the original 1981 study.

Awazu *et al.*[12] studied oxides grown on both (111) and (100) substrates. They found that if an as-processed interface had low P_b density then technologically relevant irradiation levels (<10 Mrad) generated very large ($\sim 10^{12}/\text{cm}^2$) P_b

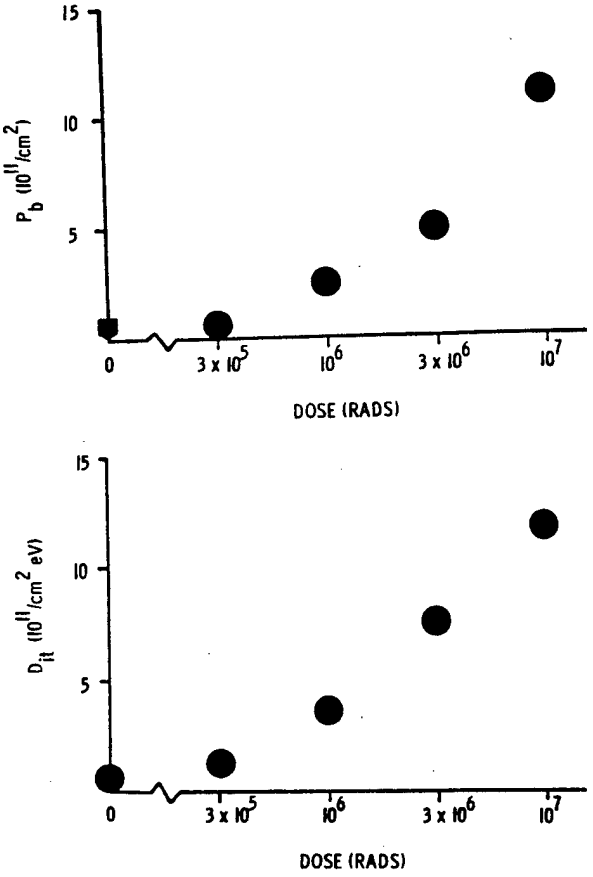


FIG. 10. Plot of P_b density (above) and interface state density in the mid-half band gap as a function of ionizing radiation dose.

densities. However, if the as-processed P_b densities were extremely high ($\sim 2-3\times 10^{12}/cm^2$) the P_b density was reduced. They argued that this result should be expected from elementary reaction theory.

Vranch *et al.*[10] have also investigated the effects of ionizing radiation on the silicon–silicon dioxide. They showed that dose levels of several megarad could generate $\sim 10^{13}$ P_b centers/cm^2 as measured in a conventional EPR measurement. Their studies included conventional EPR as well as spin dependent recombination. They found (as did several other groups) a strong preponderance of P_{b0} centers generated by radiation stressing.

B. High and low oxide field electron injection

In addition to ionizing radiation, several other oxide stressing mechanisms have been investigated by EPR. Mikawa and Lenahan[37] found that P_b centers could be generated by injecting electrons into an oxide at low field by internal photoemission. Warren and Lenahan[38] showed that P_b centers could also be generated by high field stressing oxides. In both the Mikawa/Lenahan and Warren/Lenahan studies a rough (about one to one) correspondence was observed between the densities of P_b centers generated and the densities of interface states in the middle half of the band gap.

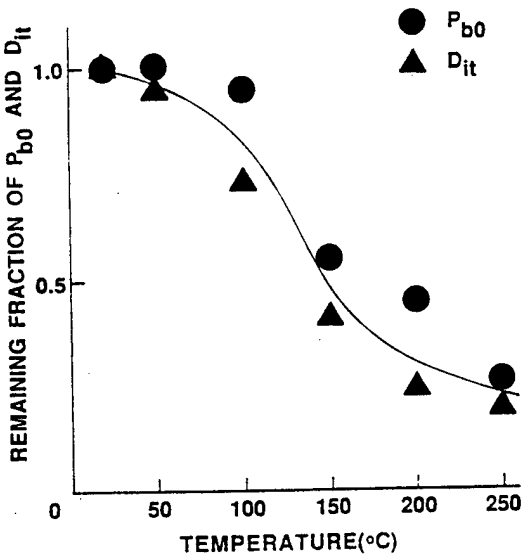


FIG. 11. Plot of the annealing behavior of P_b centers and the interface state density in the middle half of the silicon band gap vs ionizing radiation dose.

C. Hot carrier stress of short channel MOSFETs

Krick *et al.*[39] and Gabrys *et al.*[24] have used SDR to study hot carrier damage centers created near the drain of short channel MOSFETs. To the best of our knowledge, these studies included the highest sensitivity electron spin resonance measurements ever made in condensed matter. Both Krick *et al.* and Gabrys *et al.* generated strong P_b SDR signals by hot hole stressing the drain regions of the transistors. Krick *et al.* reported that the P_{b0} SDR signal scaled with increasing interface state density as measured by charge pumping. The Gabrys measurements were sufficiently sensitive to allow detection of the ^{29}Si hyperfine side peaks as well as an evaluation of the P_{b0} ^{29}Si hyperfine tensor.

D. Conclusion regarding P_b centers and Si/SiO_2 instabilities

At least four independent groups[6–12,23,24,37–39] have reported essentially incontrovertible measurements on many oxides demonstrating a strong generation of P_b centers in device stressing. For radiation induced instabilities, all four groups concluded that P_b centers play a dominating role. Although these studies have been most extensive for ionizing radiation, one may reasonably conclude that P_b centers do indeed play dominating roles in several technologically important instabilities. The earlier studies of Nishi *et al.*[2–5] and Poindexter *et al.*[4,5] established that P_b centers play dominating roles in as-processed Si/SiO_2 structures with relatively poor interfaces.

E. Dissenting opinion

It should probably be noted that Cartier and Stathis[19] have strenuously objected to this conclusion. In a recent paper, they wrote that "prior to these studies, it was widely accepted that the silicon dangling bond defect, which gives rise

to the well known P_b signal in ESR is the microscopic defect causing the fast interface states. As will be outlined in this contribution, we cannot support this view." They argue that "silicon dangling bonds, as detected by ESR measurements account for only *a small* fraction of the electrically detected interface states" (emphasis added).

Cartier and Stathis draw this conclusion on studies in which they bombard Si/SiO_2 structures with *extremely* high fluences of atomic hydrogen (up to 10^{21} hydrogen atoms/cm^2) to generate extremely high densities of interface states ($>5\times10^{12}$/cm^2 eV).

There are a few problems with the Cartier/Stathis study. At least five bear mentioning.

(1) The $\approx10^{21}$ hydrogen atoms/cm^2 used in their study corresponds to about *one million monolayers*. A typical oxide ~100 Å thick with ~0.1 at. % hydrogen would have about 10^{14} hydrogen atoms/cm^2. Thus, the amount of hydrogen involved in their model experiments is many orders of magnitude greater than that present in the process they are attempting to model.

(2) Ionizing radiation, in particular, has been well studied. Atomic hydrogen is dimerized in a fraction of a second at room temperature.[40] At room temperature, generation of interface states proceeds for *many seconds* after a device is exposed to ionizing radiation.[41,42] Since atomic hydrogen is not present during nearly all the time involved in interface state generation, atomic hydrogen by itself *cannot* be responsible for most of the process.

(3) Johnson *et al.*[43] have shown that atomic hydrogen is extremely effective in annihilating silicon dangling bonds. Testing a silicon dangling bond generation model with a process known to annihilate silicon dangling bonds is a less than an optimal approach.

(4) A fourth problem with the work of Cartier and Stathis involves the extremely high amounts of energy which would be required to generate fluences of 10^{21} hydrogen atoms/cm^2. For example, if one were actually to flood an interface with 10^{21} hydrogen atoms/cm^2, say via ionizing irradiation, one would necessarily have to break 10^{21} hydrogen atoms bonds/cm^2. If each bond energy is ~2 eV, $\sim2\times10^{21}$ eV/cm^2 would be absorbed by the ~100 Å oxide involved. (Stathis *et al.* reported results of P_b generation in 97.5 Å oxides.) Radiation dose is typically reported in rads: 1 rad$=10^2$ ergs/gram. For Si/SiO_2 then,

$$1\ \text{rad}=(10^2\ \text{erg/g})\left(\frac{10^{-7}J}{\text{erg}}\right)\left(\frac{\text{eV}}{1.6\times10^{-19}J}\right)$$

$$\times\left(\frac{2.2\ \text{g}(SiO_2)}{\text{cm}^3}\right)\cong1.4\times10^{14}\ \text{eV/cm}^3.$$

Thus, for a 100 Å thick oxide, a fluence of 10^{21} hydrogen atoms/cm^2 would correspond to a dose of *at least* 1.3×10^{13} rad. This is about a million times higher dose level than the highest levels utilized in the earlier studies and, not coincidentally, a million times higher than the upper limit of technological relevance. Indeed, this amount of energy is orders of magnitude higher than that required to vaporize the oxide sample in question.

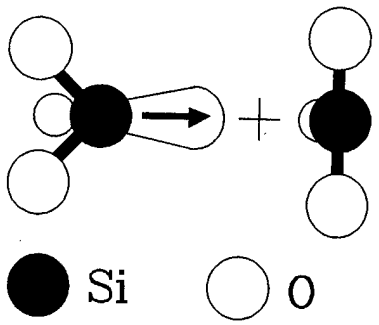


FIG. 12. Schematic illustration of the E' center.

(5) There is another, slightly more subtle problem with the Cartier *et al.* study. Their fundamental result, gross differences in P_b and interface state densities, involved EPR measurements of P_b centers on 97 Å oxides and capacitance versus voltage measurements on 495 Å oxides.[44] An extremely large number of studies involving radiation damage have consistently shown that interface state generation is a very strong fraction of oxide thickness.[45–47] Since very large differences in interface state generation are consistently observed for different oxide thicknesses, one would not expect a 97 Å oxide and a 495 Å oxide to exhibit comparable interface state densities in the process which Cartier *et al.* attempt to model. (Indeed, the thinner oxide would be expected to exhibit much lower interface state density.[45–47])

Although the Stathis/Cartier atomic hydrogen studies may be of some general interest, their conclusions are not strongly supported by their own data and are contradicted by much more relevant and extensive data generated by many other groups.

VII. OXIDE CENTERS: NEAR Si/SiO_2 CENTERS

A. E' centers structure

The most important oxide trapping centers are E' centers, which involve an unpaired electron localized on a silicon backbonded to three oxygens. Usually, though not always, the paramagnetic silicon site is coupled to a positively charged diamagnetic silicon as shown in Fig. 12.

E' centers have been studied in cubic centimeter sized samples for quite some time.[48,49] In these large volume samples, it is quite easy to measure the hyperfine interactions of the unpaired electron with the single silicon atom on which it primarily resides. Although the magnetic ^{29}Si nuclei are only ≅5% abundant, the number of centers present in large volume samples (typically $\sim10^{16}$–10^{17}/cm^3 for amorphous SiO_2) is more than sufficient to generate quite strong ^{29}Si spectra. An EPR spectrum taken on such a large volume amorphous sample is shown in Fig. 4. As discussed previously one may, by inspection, obtain a rough estimate of the hybridization and localization of the electron from this spectrum. Also, as discussed previously, one notes that a second integration of the two side peaks (corresponding to spin 1/2

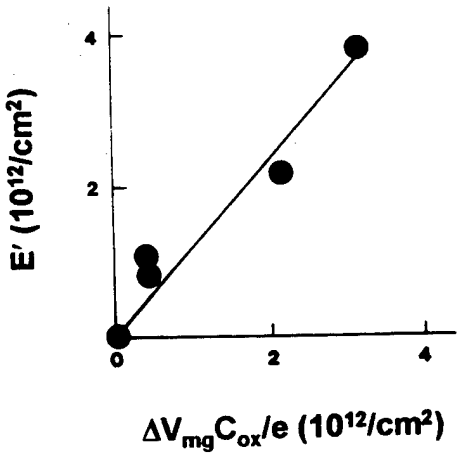


FIG. 13. Plot of E' density vs trapped hole density in MOS oxides subjected to ionizing radiation.

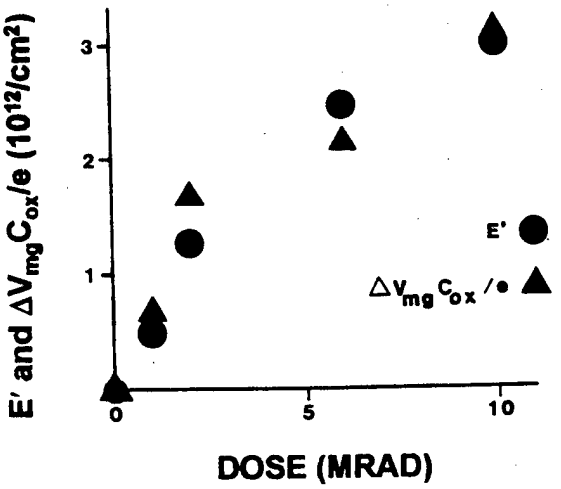


FIG. 14. Plot of trapped hole density and E' density vs dose for oxides irradiated under positive gate bias.

nuclei) yields an intensity of about 5% of the center line—unambiguously identifying the center as an unpaired spin localized on a silicon atom.

In crystalline SiO_2, Feigl *et al.*[49] argued that E' centers are holes trapped in oxygen vacancies. The unpaired electron resides on a neutral silicon on one side of the vacancy. The silicon on the other side of the vacancy is positively charged. Bonded to just three other atoms, it adjusts its position to a flat-planar arrangement (expected for sp^2 bonding) with its three neighboring oxygens. In amorphous SiO_2, E' centers can be positively charged or neutral.

Quite a number of studies indicate that E' centers are dominating trapped hole centers in technologically relevant thermally grown oxide films.[8,9,11–16,36,50] These E' centers are primarily the positively charged Feigl Fowler Yip E' defects shown in Fig. 12.

B. MOS oxide E' centers: Electronic properties

The role of E' centers in high quality thermally grown oxides is fairly well understood at the present time. These centers were probably first detected in thermally grown oxides by Marquardt and Sigel[51] who studied quite thick (up to 11 000 Å) oxides subjected to quite high (up to 220 Mrad) doses of ionizing radiation. They observed weak signals in these films which they attributed to E' centers. Although they did not report results of electrical measurements, they proposed (correctly) that E' centers are thermal oxide hole traps.

The electronic properties of E' centers and their significance in MOS device operation were first demonstrated by Lenahan and Dressendorfer.[8,36] They made a series of observations which clearly established that E' centers are dominating hole trap centers in a variety of MOS oxides. (1) They found a rough one-to-one correspondence between E' density and the density of hole traps in relatively hard and relatively soft oxides grown in both steam and dry oxygen (see Fig. 13). (2) They found a rough one-to-one correspondence between E' density and trapped hole density in oxides irradiated under positive gate bias over a technologically meaningful range of ionizing radiation dose (see Fig. 14). (3) They showed that MOS oxide E' centers and oxide trapped holes have the same annealing response in air (see Fig. 15). (4) They demonstrated that the distribution of E' centers and trapped holes are virtually identical in oxides subjected to ionizing irradiation under positive gate bias: both the trapped holes and E' centers are quite close to the Si/SiO_2 boundary (see Fig. 16).

The results of the early studies have been confirmed and extended in quite a few later studies. (1) Takahashi and co-workers[52,53] also reported an approximately one-to-one correspondence between E' centers and trapped holes; they also reported that the distribution of E' centers and trapped holes were the same in their irradiated oxides. (2) Lipkin

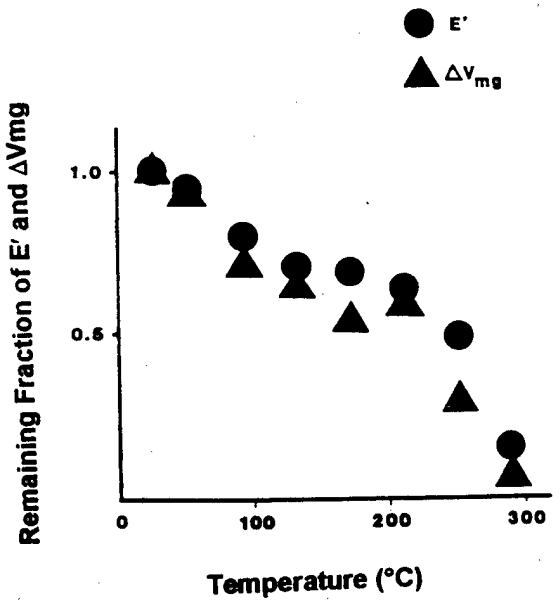


FIG. 15. Annealing response of E' centers and holes trapped in the oxide.

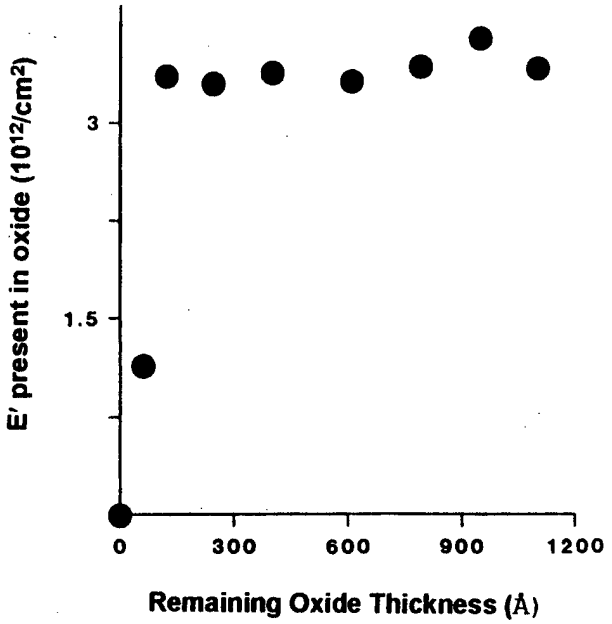


FIG. 16. Plot of E' centers remaining in an oxide after a series of etch back steps. The results show that nearly all the E' centers, like nearly all the trapped holes, are near the Si/SiO$_2$ boundary.

et al.[54] also measured an approximately one-to-one correspondence between E' density and the density of trapped holes generated in oxides subjected to 10–20 Mrad of gamma radiation (3) Miki *et al.*[11] compared both E' generation and trapped hole generation in ultradry and steam grown oxides. They found that their ultradry oxides contained twice as many E' centers as the steam grown oxides and that the ultradry oxides also had twice as many trapped holes as the steam grown oxides. In addition, they found a rough numerical correspondence between the E' densities and trapped hole densities in the samples investigated. However, since Miki *et al.*[11] made electrical measurements on oxides subjected to avalanche injection of holes and EPR measurements on x-ray irradiated oxides, a precise numerical comparison between E' density and trapped hole density was not possible. (4) Awazu *et al.*[12] have explored the role of processing parameters on E' generation. As previously noted by Lenahan and Dressendorfer,[8,36] as well as by Miki *et al.*,[11] the densities of E' centers are strongly processing dependent. A point of particular interest in the Awazu study is their observation that the cooling rate after high temperature processing strongly affects E' generation. Awazu *et al.* concluded that the E' centers in their oxides were holes trapped in oxygen vacancies $O_3{\equiv}Si+\cdot Si{\equiv}O_3$. Since they observed $\sim 1-3\times 10^{12} E'$ centers/cm^2 after modest ($\sim$1–11 Mrad) doses of ionizing radiation, these centers would inevitably be the dominant hole trap centers in the oxides of their study. (5) Results of a more qualitative nature by Carlos[14] also support the correspondence between oxide trapped holes and E' centers. He reported a significant ($\sim 10^{12}$/cm^2) density of E' centers in oxides subjected to technologically relevant radiation levels. Although he did not report results of any electrical measurements, he did report a gate polarity dependence of E' generation, suggesting (as the other studies had shown) that E' generation was due to the capture of a charged particle.

The results of at least five independent EPR studies all indicate a dominant role for E' centers in oxide hole trapping. On the basis of this mutually corroborating work we conclude that E' centers do indeed dominate oxide hole trapping in a wide variety of thermally grown oxide films on silicon. However, the current understanding of oxide hole trapping is moderately complex and still incomplete. At least five E' variants have been observed in thermally grown oxide films. Two E' variants are E' defect/hydrogen complexes for which detailed and convincing models have yet to be established. A third variant called E'_δ or EP is almost certainly closely related to the conventional E' site but exhibits significantly different g and hyperfine tensor components and somewhat different capture cross sections than the conventional E' site. A fourth variant, called EP2 is even less well characterized. It is positively charged when paramagnetic and exhibits a large capture cross section for holes. A fifth variant is the neutral E' center, observed in plasma-enhanced chemical vapor deposition (PECVD) oxide films and in thermally grown oxides exposed to very high doses of ionizing radiation.

VIII. HYDROGEN COMPLEXED E' CENTERS

Two hydrogen complexed E' variants may play important roles in device reliability. The hydrogen complexed centers are called the 74 G doublet and the 10.4 G doublet. EPR traces of both defects are shown in Fig. 17. Both defects were first observed in cubic centimeter size samples,[55,56] and both centers clearly involve an E' hydrogen complex. It has been proposed[55] that the 74 G doubled involves an unpaired electron on a silicon back bonded to two oxygens and one hydrogen and that the 10.4 G doublet defect involves an unpaired electron on a silicon back bonded to three oxygens with one of the oxygens bonded to a hydrogen.[55] (Although these models seem quite reasonable and are undoubtedly correct to the extent that the defects are E'/hydrogen complexes, in detail, the models should probably be viewed as provisional.)

The 74 G doublet was first observed in thin oxide films by Takahashi *et al.*[52,53] who generated them at somewhat elevated temperatures ($\cong$100 °C) in irradiated oxides. More recently, Conley *et al.*[57,58] observed the room temperature generation of both the 74 G doublet centers and 10.4 G doublet centers in oxides subjected to either vacuum ultraviolet ($hc/\lambda \leqslant 10.2$ eV) or gamma irradiation.

Takahashi *et al.*[52,53] suggested that the hydrogen complexed E' defects might play an important role in Si/SiO$_2$ interface state generation. Conley *et al.*[57,58] provided strong circumstantial evidence linking E'/hydrogen coupled centers to interface trap generation. Several (purely electrical measurement) studies[59,60] had shown that a molecular–hydrogen-containing ambient leads to an enhancement in radiation induced interface state generation. Conley *et al.*[57,58] showed that exposing an oxide previously flooded with holes (to gen-

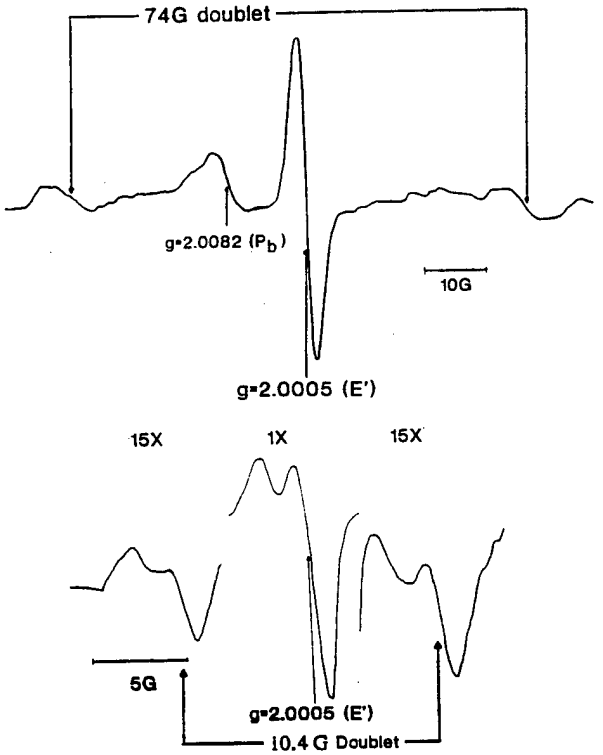


FIG. 17. EPR traces of (a) the 74 G doublet and (b) the 10-4 G doublet centers.

erate E' centers) to an H_2/N_2 ambient leads to a conversion of conventional E' centers to 74 G doublet centers as well as generation of interface state centers. The number of E' centers converted to hydrogen complexed centers is approximately equal to the number of interface traps generated. The time period involved in interface trap formation is approximately equal to the time required to saturate the E'/hydrogen complexing process.

IX. EP OR E'_δ VARIANT (ONE DEFECT, TWO NAMES)

The EP or E'_δ variant has been observed in separation by implanted oxide (SIMOX) buried oxides,[61–65] bond and etch back (BESOI) buried oxides, and thermally grown oxides. It is generally observed simultaneously with the conventional E' center. Its spectrum is quite narrow with a zero crossing $g \cong 2.002$. Vanheusden and Stesmans[64] estimate that the E'_δ/EP variant accounts for about 20% of the E' centers in SIMOX buried oxides. Conley *et al.*[61,63,65,66] have observed EP/E'_δ centers in a variety of thermally grown oxides as well as in SIMOX buried oxides and demonstrated that they are positively charged.

Vanheusden and Stesmans and Warren *et al.*[63] had initially proposed that the EP/E'_δ center involved a five silicon atom microcluster. Conley and Lenahan[67] compared the response of the EP/E'_δ centers to molecular hydrogen and found nearly identical responses for both centers: virtually identical time scales for the room temperature response, nearly identical hyperfine coupling constants for both hydrogen complexed centers, and nearly identical g values for both hydrogen complexed centers. On the basis of those observations, they argued that the EP/E'_δ center must be very closely related to conventional E' centers, and almost certainly *not* a five atom silicon cluster. Recent studies of the EP/E'_δ ^{29}Si hyperfine spectrum indicated that the unpaired electron is shared by two equivalent silicons.[67]

Quite recently, Chavez *et al.* have proposed that the EP/E'_δ center is an oxygen vacancy but with the positive charge and the unpaired electron equally shared by the two silicons. The model proposed by Chavez *et al.*[68] is consistent with many experimental observations: (1) as observed experimentally, the center is positively charged, (2) as observed experimentally the center behaves almost exactly like the conventional E' center in response to molecular hydrogen. (3) The predicted hyperfine coupling is almost exactly what is observed experimentally. (4) The model is consistent with recent observations indicating that the unpaired electron is shared by two silicons. There is extensive agreement between the Chavez model and experimental observations; it's probably correct.

X. NEAR Si/SiO_2 INTERFACE E' CENTERS: ELECTRONIC PROPERTIES

One could reasonably divide electrically active MOS defects into three categories: Interface state traps, which can communicate readily with charge carries in the silicon, oxide traps which do not communicate with charge in the silicon, and very near interface traps which can, on fairly long time scales, communicate with charge carriers in the silicon.

The near Si/SiO_2 interface traps go by many names: slow states, border traps, switching traps,.... It is possible, even likely, that more than one type of near Si/SiO_2 interface trap can exist in certain oxides under certain conditions. One near Si/SiO_2 interface trap has been directly identified via EPR the E' center.

Many studies (involving only electrical measurements) show that when some MOS devices are subjected to ionizing radiation capacitance versus voltage and current versus voltage, characteristics experience a negative voltage shift, ΔV, indicating the capture of positive charge in the oxide.[69,70] However, if a positive gate bias is applied, the magnitude of ΔV decreases logarithmically in time, indicating the annihilation of some of the positive charge. If the applied bias is reversed from positive to negative some of the previously annihilated charge returns. The charge that returns is said to be in switching traps.

An EPR study by Conley *et al.*[71] clearly demonstrates that some E' centers can act as switching traps. These centers are presumably very close to the Si/SiO_2 boundary. Results from the Conley *et al.*[71] study are shown in Fig. 18. The bias voltages and bias switching times approximately match those of earlier purely electronic measurements. The first point (HOLES) indicates the E' density initially after holes were injected into the oxide. The second point (Zero 1) was taken after 10^5 s with no bias across the oxide. Point (Neg 1) was taken after a negative gate voltage corresponding to an aver-

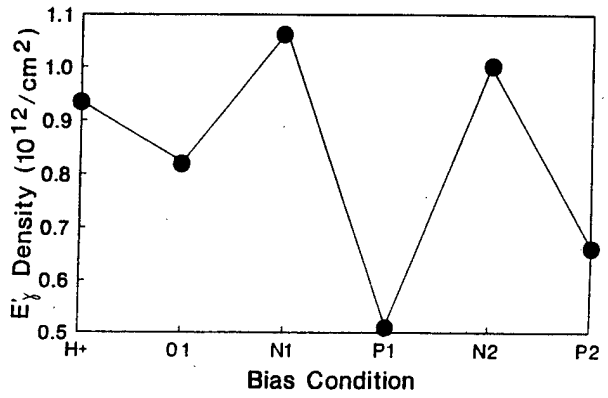


FIG. 18. Plot of E' spectrum amplitude vs the biasing sequence discussed in the text.

age oxide field of 3.5 MV/cm was applied for 24 h. The negative bias increased the number of paramagnetic E' sites. Point (POS 1) was taken on the same sample after an additional 24 h under positive gate bias, also corresponding to an average oxide field of 3.5 MV/cm. The positive bias substantially decreases the density of paramagnetic E' centers. Two additional biasing points indicate the repeatability of this process. Clearly, the "spin state" and thus "charge state" of these E' centers can be repeatably switched with bias;thus E' centers can act as oxide switching traps.

The results of Conley *et al.*[71] are consistent with, and clearly confirm, the basic premise of the switching trap model proposed earlier by Lelis *et al.*[69,70] after hole capture, subsequent electron capture does not always return the E' site involved to its original condition. This irreversibility leads to the switching behavior. The results of Conley *et al.* do not, of course, preclude the possibility that defects other than E' centers may act as switching traps.

The Conley *et al.*[71] results extend earlier results generated by Jupina and Lenahan who reported the SDR detection of E' centers.[21] Since SDR can only detect defects which in some way "communicate" with Si/SiO$_2$ interface charge carriers, their results strongly indicated that some near Si/SiO$_2$ interface E' centers did indeed behave in this way. The Conley *et al.* study is also consistent with a recent ESR study by Warren *et al.*[72] which suggested, but did not demonstrate, that E' centers may act as switching traps.

XI. INTRINSIC DEFECTS AND DEVICE RELIABILITY: PHYSICALLY BASED PREDICTIVE MODELS

Although the many EPR studies of MOS systems are of some general interest as physics, chemistry, and materials science, their ultimate significance must relate to their utility: can these studies help design better, more reliable, integrated circuits? The answer to this question is almost certainly yes. Yes, if the results can be utilized to predict and manipulate defect densities.

It is clear that two families of point defects, E' centers and P_b centers, play dominating roles in a number of MOS reliability problems. Materials scientists and engineers have well developed and widely verified methods of manipulating intrinsic point defect populations. These methods are based upon the fundamental principles of the statistical mechanics of solids as well as on basic principles of physical chemistry.

One should be able to ameliorate device reliability problems by applying these well established principles to E' and P_b centers.

A. Predicting oxide hole trapping

Lenahan and Conley[15] used the standard approach of statistical mechanics[73,74] to calculate the density of oxygen vacancies in MOSFET oxides, calibrated the parameters of the expression with EPR measurements, and then tested the validity of calibrated (quantitative) expression on several oxide films. They found good correspondence between the calibrated expression and experimental results.

A consideration of the basic principles of statistical thermodynamics tells us that equilibrium occurs when the Gibbs free energy G of a solid is minimized.[73,74] It can be shown that, for the simplest cases, the minimization of Gibbs free energy leads to an equilibrium density of vacancy sites given by

$$n = Ne^{\Delta S_f/k - \Delta H_f/kT}, \tag{13}$$

where ΔS_f represents the nonconfigurational entropy contribution per defect site, ΔH_f represents the enthalpy of formation of a defect site, k is the Boltzmann constant, and N represents the density of available sites. For the purposes of this discussion, the important points here are that the nonconfigurational entropy contribution is large and essentially temperature independent, and the ΔH_f essentially represents the increase in system energy caused by vacancy creation of an unstressed lattice site minus the strain energy lost by removal from a compressed SiO$_2$ matrix. (This reduction in ΔH_f would be a strain energy $\sim PdV$ caused by the effective volume change resulting from the removal of the atom from its particular location.)

As pointed out by Ohmameuda *et al.*[75] this strain energy reduction will be greatest for sites near the Si/SiO$_2$ boundary; this energy contribution should amount to several tenths of an electron volt.[6] One thus expects and finds[8] that the E' centers are primarily located close to the Si/SiO$_2$ boundary.

Anticipating then an oxygen vacancy/E' precursor density of the form

$$n = \alpha e^{-\beta/T}, \tag{14}$$

where the temperature independent constant α is given by $\alpha = Ne^{\Delta S_f/k}$ and $\beta = \Delta H_f/k$, we may evaluate the relevant "thermodynamic" constants by making measurements on devices exposed to various high-temperature anneals. With a knowledge of E' center hole capture cross section[14] and the standard analysis of charge capture in oxide films, we would anticipate that, for a given fluence of holes through the oxides,

$$N_{\text{th}} = \alpha e^{-\beta/T}(1 - e^{-\sigma\eta}), \tag{15}$$

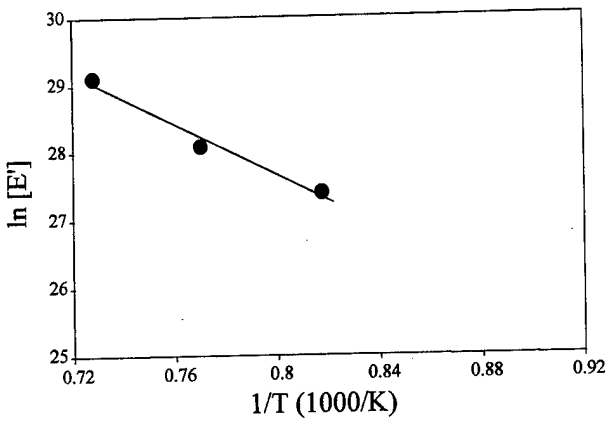


FIG. 19. Plot of the natural logarithm of E' density vs the reciprocal of annealing temperature. Slope of the line yields an activation energy of to be about 1.5 eV.

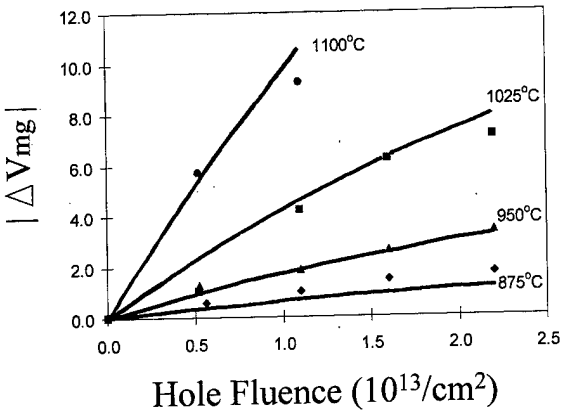


FIG. 20. Solid lines represent Eq. (16) evaluated for the various indicated temperatures. Dots represent experimental results.

where N_{th} is the density of trapped holes and η is the fluence of holes through the oxide. With α, β, and σ evaluated from spin resonance measurements the expression provides a no-adjustable-parameter prediction of oxide hole trapping. (However, due to the modest absolute precision of EPR measurements, the value of α as determined strictly from EPR could be in error by almost a factor of 2.)

The potential validity of Eq. (15) was assessed[15] through a series of measurements on MOS oxides subjected to anneals at 875, 950, 1025, and 1100 °C. The oxides were all grown at 825 °C and then a polysilicon gate was deposited. After gate deposition the anneals were carried out for 30 min in a dry N_2 atmosphere. After the anneals the capacitors were rapidly pulled from the furnace in order to "quench in" the defect densities at the annealing temperatures.

The poly gates were removed and two sets of measurements were made on the samples, both after subjecting the oxides to hole flooding. To evaluate the E' precursor enthalpy of creation, oxides of the three higher temperature annealing samples were each flooded with approximately 2×10^{13} holes/cm^2. The enthalpy was determined from the slope of a plot of the natural logarithm of E' density versus reciprocal temperature, shown in Fig. 19; the activation enthalpy is approximately 1.5 ± 0.1 eV. To test the predictive capability of Eq. (15), holes were injected into samples subjected to each of the four annealing steps; mid-gap capacitance versus voltage shifts, ΔV_{mg}, were plotted versus injected hole fluence. Using expression (15) and taking the trapped holes to be close to the Si/SiO$_2$ boundary, our model[15] predicts mid-gap shifts of

$$\Delta V_{mg} = \frac{q \alpha e^{-\beta/T}}{C_{ox}} (1 - e^{\sigma \eta}), \tag{16}$$

where q is electronic charge, C_{ox} is oxide capacitance, and all other parameters are as previously defined.

Figure 20 compares the experimental results and the predictions of Eq. (16). The correspondence between prediction and experiment is quite close. It clearly demonstrates something new and almost certainly useful: It is possible to predict the response of an oxide from an essentially no-adjustable-parameter fit of a physically based model.

B. Predicting interface trap formation

Straightforward concepts from the equilibrium thermodynamics of chemical reactions also allow one to make some predictions about interface trap formation.[16]

As discussed previously at least four independent groups have demonstrated that significant (greater than or approximately equal to 10^{12}/cm^2) generation of P_b centers occurs when MOS oxides are subjected to technologically relevant levels of ionizing radiation. More limited results indicate P_b generation resulting from high or low oxide electric field injection of electrons as well as the injection of hot holes into the oxide from the near drain region of short channel MOSFETS. The previously discussed studies of Conley *et al.*[57,58] show that E' centers react with molecular hydrogen at room temperature and that this reaction is accompanied by the simultaneous generation of Si/SiO$_2$ interface traps (these are P_b centers).

In our study,[57,58] only about 25% of the E' centers disappeared as about an equal number of hydrogen complexed E' centers (which we term E'H) appeared: the loss in E' density was accompanied by an approximately equivalent gain in interface trap density. No increase in interface trap density occurred with H_2 exposure if the positively charged E' centers were absent. These observations are significant because when SiO_2 is subjected to ionizing radiation, atomic hydrogen is created: above 110 °K it very rapidly dimerizes leaving behind H_2 in the oxide.[76]

It is well established that silicon dangling bond sites at the Si/SiO$_2$ interface (P_b centers) are passivated by hydrogen.[3,77] Assuming that the interface trap creation process involves the breaking of silicon–hydrogen bonds at P_b center precursor sites (P_bH) we proposed a reaction of the following form:[16]

$$H_2 + P_b H + E' \rightleftarrows H_2 + P_b + E'H. \tag{17}$$

In this reaction, H_2 plays the formal role of a catalyst. When a hole drifting to an E' precursor site is captured, a positively charged silicon dangling bond site (E' center) is

created, which, as Conley *et al.*[58] have shown, can react with radiolytic H_2 to form a complex which we term $E'H$. After irradiation, the H_2 is eventually dissipated, but for a short time the system will *approach* equilibrium. Elementary statistical mechanics tells us that, if the system were to reach equilibrium, one could write[78]

$$\frac{[P_b][E'H]}{[P_bH][E']}=K, \tag{18}$$

where $K=\exp(-\Delta G/kT)$ and ΔG is the difference in Gibbs free energy of the reactants and products. Since expression (18) involves the transfer of a hydrogen atom from a silicon at the interface (P_b) to a silicon in the oxide (E'), one would reasonably conclude that ΔG is small. Thus, $K\cong 1$, at least within about an order of magnitude.

In order to solve Eq. (18) for ΔP_b (the concentration of P_b centers eventually generated after the interface trap formation process is complete) define the initial (prestress) P_b concentration to be P_{bi}, the initial (prestress) P_bH concentration to be $(P_bH)_i$, and the density of E' trapped holes present immediately after irradiation (and immediately after all the holes which were not trapped are swept from the oxide) to be E_1'.

With these definitions, Eq. (19) becomes

$$\frac{[P_{bi}+\Delta P_b][\Delta P_b]}{[(P_bH)_i-\Delta P_b][E_i'-\Delta P_b]}=K. \tag{19}$$

In Eq. (19), take the number of $E'H$ complex sites created to be equal to the number of P_b sites created. Assuming it captures the essential physics of the process, Eq. (19) predicts the interface trap generation behavior in a wide range of oxides.

For the technologically important situation in which the Si/SiO_2 interface has a very low interface trap and P_b density, $P_{bi}\cong 0$, at low dose $[\Delta P_b \ll (P_bH)_i$, Eq. (19) becomes

$$\frac{[\Delta P_b][\Delta P_b]}{(P_bH)_i(E_i'-\Delta P_b)}\cong K, \tag{20}$$

yielding

$$\Delta P_b\cong \frac{K}{2}(P_bH)_i(\{1+4E_i'/[K(P_bH)_i]\}^{1/2}-1). \tag{21}$$

For a low level of initial E_i' generation ΔP_b will be almost equal (always slightly less than) the initial E_i' density. Thus, if we were to flood a very good oxide with a small number of holes, suppress the interface trap generation process, measure the initial trapped hole concentration, and then allow interface trap generation to proceed, we would expect that the eventual interface trap density (each P_b has two levels) would be roughly equal to the initial trapped hole density.

The generation of interface traps *can* be suppressed for hours by lowering the temperature of the system;[79] warming to room temperature allows the process to proceed. Many years ago, Hu and Johnson[79] subjected good oxide/silicon devices to relatively low levels of hole flooding at temperatures low enough to temporarily suppress interface trap generation. Initial oxide hole densities were evaluated, then interface traps allowed to generate, and those interface trap densities were also evaluated. As Eq. (21) predicts, Hu and Johnson found that the initial oxide hole density was approximately equal to the eventual interface trap density.

There are other semiquantitative aspects of this model which are in agreement with results in the literature.

(1) Expression (21) shows that ΔP_b should be sublinear in E_i'. Since E_i' generation should itself be sublinear in dose,[15] the model predicts a sublinear buildup of interface sites with dose. Such behavior is widely reported.[80]

(2) The model predicts that in devices with low initial interface trap density, the hole trapping and interface trap generation, P_b generation and E' generation would approximately scale together. This behavior (both cases) has been observed.[36]

(3) Since the model involves interaction of a trapped hole site with molecular hydrogen triggering a reaction at a Si/SiO_2 interface P_bH site, one would expect that the time involved in interface trap generation would be significantly increased by reversing the irradiation bias from positive gate voltage to negative gate voltage. This behavior is consistently observed.[81–83]

(4) Since the E' center precursors (oxygen vacancies) are intrinsic defects, one would expect that their number would be an exponential function of processing temperature. Thus one would expect a strong increase in interface trap generation with increasing temperature of gate oxide processing. This behavior is observed.[36]

(5) Radiolytic, molecular hydrogen will be rapidly dissipated from the oxide; in some cases this will not allow equilibrium densities of P_b interface traps to be achieved. One would thus expect that post irradiated exposure to a molecular hydrogen ambient generally increases interface trap density. This behavior is observed.[59,60]

(6) Consider a metal–oxide–semiconductor device in which the oxide has been flooded with holes for a brief period. If a positive voltage was applied to the gate electrode, E' precursors near the Si/SiO_2 boundary would be populated with holes: if a negative voltage was applied, E' precursors near the gate (usually polycrystalline Si/SiO_2 boundary) would be similarly populated. Our model, at least to zero order, predicts a similar radiation response, in that the eventual number of P_b centers created would be the same with either sign of gate bias during irradiation. (This assumes equal E' precursor density at both interfaces.) Experimental work indeed shows this to be the case "electrically" for brief bursts of irradiation *provided* that the oxide bias is positive *after* the irradiation.[83] (That is, the eventual interface state densities generated are approximately equal for both cases.)

(7) Briefly consider the technologically irrelevant case of a very high initial interface trap density and a very high initial P_b density. In such a case we would expect a reaction of the following form:

$$H_2+P_b+E'\rightleftarrows E'H+P_bH. \tag{22}$$

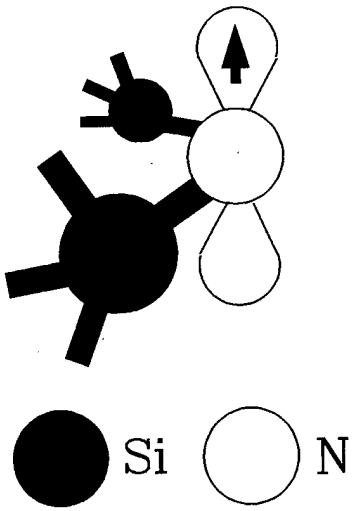


FIG. 21. Schematic illustration of the structure of the bridging nitrogen center.

Thus, the model would clearly predict an initial post irradiation *decrease* in P_b density for Si/SiO_2 structures with quite high initial P_b density. This behavior has been reported.[12]

XII. EXTRINSIC TRAPPING CENTERS

Often, impurity atoms (other than hydrogen) will be present in an oxide; nitrogen, phosphorous, and boron impurities have all been investigated to some extent in oxide films on silicon. Among these impurities, nitrogen appears to be of greatest potential significance.

A. Defect centers involving nitrogen

Chaiyesena *et al.*[84] and Yount *et al.*[85–88] have studied nitrogen defect centers in nitride and reoxidized, nitrided oxide films. Most of their work dealt with the bridging nitrogen center, N_b, illustrated in Fig. 21. An EPR trace due to this defect is illustrated in Fig. 22. A three line spectrum with each line of equal intensity is indicative of a spin-1 nucleus of approximately 100% abundance. Nitrogen is the only possibility. The center line is quite narrow, the two side peaks rather broad. This is so because quantum mechanics restricts the nitrogen nucleus to three orientations: parallel to the applied field, antiparallel to the applied field, and (any direction) perpendicular to the applied field. Since the perpendicular orientation is not further specified, the center line effectively corresponds to no nuclear moment.

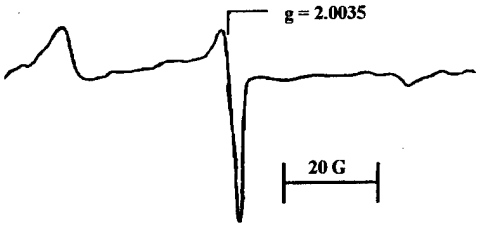


FIG. 22. EPR trace of the bridging nitrogen center.

A straightforward analysis of the spectrum of Fig. 22 indicates an unpaired electron highly localized on a single nitrogen in a nearly pure p-type wave function. (Recall that the small splitting of the side peaks indicates quite low s character; their substantial breadth high p character.) The results clearly demonstrate that this center involves a nitrogen bonded to two other atoms, almost certainly silicons. The structure of the bridging nitrogen center is shown in Fig. 21. The bridging nitrogen EPR spectrum was first studied in large volume (~ 1 cm^3) samples by Mackey *et al.*[89] who also established the center's structural nature.

It should probably be mentioned that a very different EPR spectrum has been linked to bridging nitrogens by Stathis and Kastner.[90] Later studies by Tsai *et al.*[91] pointed out that the spectrum described by Stathis and Kastner[90] was incompatible with the bridging nitrogen hybridization and localization. Recently Austin and Leisure[92] have argued that the spectrum discussed by Stathis and Kastner[90] does not involve nitrogen at all but is caused by a carbon atom bonded to two hydrogens.

Chaiyasena *et al.*[84] and Yount *et al.*[85–88] identified the bridging nitrogen center as a rather large capture-cross-section electron trap. They found quite high densities of these centers in NH_3 annealed oxides ($\sim 10^{13}$/cm^3) and that their numbers are substantially reduced but not eliminated by reoxidation. Quite high densities of electron traps ($\sim 10^{13}$/cm^3) are found in NH_3 annealed oxides; their numbers are reduced but not eliminated by reoxidation. In N_2O annealed or N_2O grown oxides one finds far lower levels of both N_B and electron traps. Yount *et al.*[85] provided quite direct evidence regarding the trapping capabilities of N_B by showing that N_B density is reduced when electrons are injected into the oxides. The N_B capture cross section is about 10^{-15} cm^2. Yount *et al.*[85] noted that N_B reacts rapidly with H_2 at room temperature. The N_B signal is annihilated if the oxides under study are briefly exposed to H_2 at room temperature.

On the basis of these observations, Yount *et al.*[85–88] proposed that the precursor for N_B is a nitrogen bonded to a hydrogen and two other atoms, presumably silicons and that the defect serves as an effective electron trap.

B. Defect centers involving phosphorous and boron

At least four phosphorous-related defect centers and one boron-related center have been identified in doped oxide films on silicon.[93–95] These centers have been studied in phosphosilicate glass (PSG) and borophosphosilicate glass (BPSG) films on silicon. The wide scan EPR trace of Fig. 23 illustrates the simultaneous presence of P_1, P_2, and P_4 centers along with the ubiquitous E' center and some organic free radicals in a tetraethyl orthosilicate (TEOS) based PSG thin film on silicon. A much narrower EPR trace taken on a silane based BPSG film indicated the simultaneous presence of the phosphorous oxygen hole center (POHC) in addition to the previously discussed P_{b0} and E' centers.

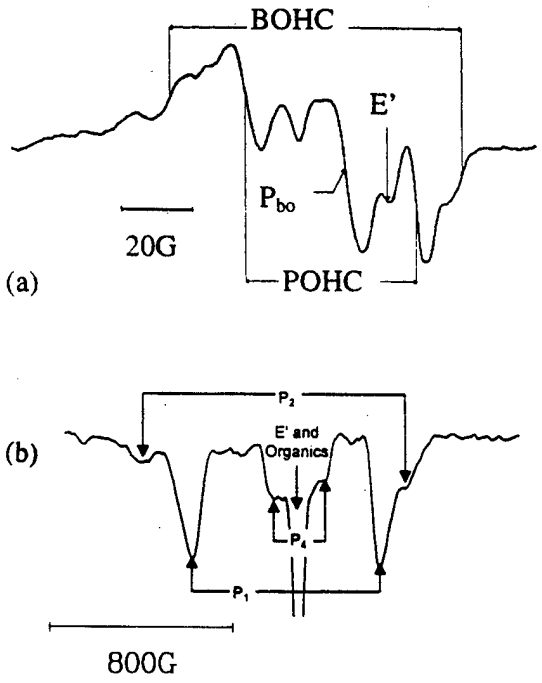


FIG. 23. EPR traces taken in (a) borophosphosilicate (BPSG) and (b) phosphosilicate (PSG) films showing (a) the boron oxygen hole center (BOHC) and phosphorous oxygen hole center (POHC) spectra superimposed upon the ubiquitous E' and P_b spectra and (b) the P_1, P_2, and P_4 spectra along with E' and organic radical spectra.

EPR parameters of the five centers are shown in Table I. Provisional defect structures (following Griscom *et al.*[96,97]) are illustrated in Fig. 24.

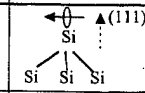
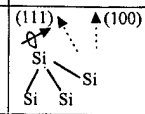
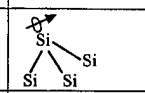
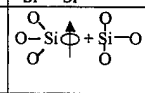
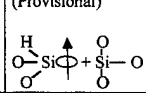
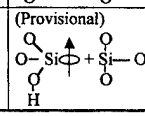
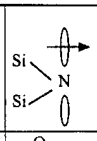
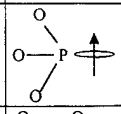
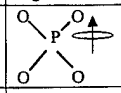
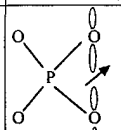
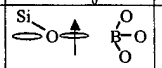
The literature on these phosphorous and boron centers in thin films is fairly limited.[93–95] They appear to play a role in instabilities in certain integrated circuits even though the BPSG and PSG films are typically interlevel dielectrics. Along with several organic centers, the phosphorous centers can capture significant amounts of space charge in the TEOS films.[95] Some of the phosphorous centers are also important charge trapping centers in silane based films.[93,94]

Limited studies of the response of P_1, P_2, and P_4 to the injection of charge carriers[94,95] into the oxide indicate that significant changes in the density of paramagnetic P_1, P_2, and P_4, and POHC centers occur when charge carriers are injected into the oxides in question; P_1, P_2, and POHC have large capture cross sections for holes, P_4 a large capture cross section for electrons.

C. Other paramagnetic centers in Si/SiO$_2$ systems

Several other paramagnetic centers have been observed in thin oxide films on silicon: unpaired electrons on oxygens,[98] (probably) nonbridging oxygens or peroxy centers, organic radicals,[95] and atomic hydrogen,[76] and electrically neutral E' centers.[65] Although radiolytic atomic hydrogen is clearly important since it is almost instantly dimerized to form molecular hydrogen at room temperature, studies of its presence in SiO_2 films on silicon are very limited. Studies of organic free radicals and the paramagnetic oxygen radical, have also been too limited to warrant further discussion.

TABLE I. EPR centers in oxide films: intrinsic and extrinsic defects.

Defect Name	EPR Parameters	Structure	Electronic Properties
(111) P_b	$g_{\|\|} = 2.0014$ $g_\perp = 2.008$ $A_{iso} = 110$ G $A_{aniso} = 21$ G		Two broad levels in gap $U \cong 0.6$ eV
(100) P_{bo}	$g_1 = 2.0015 \cong g_{\|\|}$ $g_2 = 2.0080 \cong$ $g_3 = 2.0087 \cong g_\perp$ $A_{iso} = 96$ G $A_{aniso} = 24$ G		Two broad levels in gap $U \cong 0.6$ eV
(100) P_{b1}	$g_1 = 2.0012$ $g_2 = 2.0076$ $g_3 = 2.0052$		Two broad levels in gap $U \cong 0.3$ eV
E' (Positively Charged)	$g_{\|\|} = 2.0018$ $g_\perp = 2.0002$ $A_{iso} = 439$ G $A_{aniso} = 22$ G		Deep hole trap Large hole capture cross section $\sigma \cong 3 \times 10^{-14}$ cm^2
EP/E'_δ	g (zero cross) $\cong$ 2.0002 $A_{iso} = 120$ G A_{aniso} = "small"	Provisional unrelaxed E' center	Deep hole trap Large hole capture cross section ($\sigma > \sigma(E')$)
E' (74 G doublet)	g (zero cross) $\cong$ 2.0016 A (H) $\cong$ 72 G	(Provisional) 	At least sometimes positively charged. Created by H_2 interaction with E'
E' (10.4 G doublet)	$g_{\|\|} = 2.0018$ $g_\perp = 2.0002$ A (H) $\cong$ 13 G	(Provisional) 	Created by H_2 interactions with E'
N_b	$g_{\|\|} = 2.0035$ $g_\perp = 2.0068$ $A_{iso} = 11$ G $A_{aniso} = 12.5$ G		Neutral electron trap $\sigma \cong 10^{-15}$ cm^2
P_1	$g_{\|\|} \cong g_\perp \cong 2.005$ $A_{iso} = 910$ G $A_{aniso} = 60$ G		When paramagnetic, it is a hole trap
P_2	$g \cong 2.0013$ $A_{iso} = 1110$ G $A_{aniso} = 50$ G		When paramagnetic, it is a hole trap
P_4	$g \cong 2.00$ $A_{iso} \cong 103$ G $A_{aniso} \cong 103$ G		Electron trap
POHC	$g_1 \cong 2.0075$ $g_2 \cong 2.0097$ $g_3 \cong 2.0179$ $A_1 = 48$ G $A_2 = 54$ G $A_3 = 51$ G		Hole trap
BOHC	$g_1 \cong 2.0025$ $g_2 \cong 2.0115$ $g_3 \cong 2.0355$ $A_1 = 13.6$ G $A_2 = 15.3$ G $A_3 = 8.7$ G		Does not have a large capture cross section for electrons or holes

XIII. CONCLUSIONS

More than a dozen paramagnetic centers have been observed in MOS systems. Several of these centers play dominant roles in MOS device instabilities. Recent studies show that predictions with regard to oxide reliability can be made by combining our EPR-derived understanding of electrically active defects with fundamental principles of statistical mechanics. The various oxide centers and their properties are summarized in Table I.

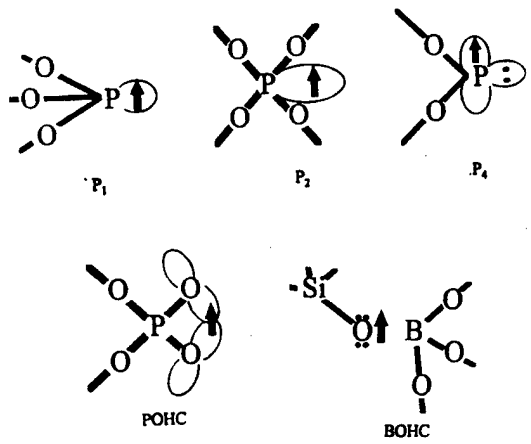


FIG. 24. Schematic sketch of phosphorous and boron centers found in PSG and BPSG thin films on silicon.

[1]J. H. Weil, J. R. Bolton, and J. E. Wertz, *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications* (Wiley, New York, 1994).
[2]Y. Nishi, Jpn. J. Appl. Phys. **10**, 52 (1971).
[3]Y. Nishi, T. Tanaka, and A. Ohwada, Jpn. J. Appl. Phys. **11**, 85 (1972).
[4]P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **50**, 879 (1979).
[5]E. H. Poindexter, P. J. Caplan, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **52**, 879 (1981).
[6]P. M. Lenahan, K. L. Brower, P. V. Dressendorfer, and W. C. Johnson, IEEE Trans. Nucl. Sci. **28**, 4105 (1981).
[7]P. M. Lenahan and P. V. Dressendorfer, Appl. Phys. Lett. **41**, 542 (1982).
[8]P. M. Lenahan and P. V. Dressendorfer, J. Appl. Phys. **55**, 3495 (1984).
[9]Y. Y. Kim and P. M. Lenahan, J. Appl. Phys. **64**, 3551 (1988).
[10]R. L. Vranch, B. Henderson, and M. Pepper, Appl. Phys. Lett. **52**, 1161 (1988).
[11]H. Miki, M. Noguchi, K. Yokogawa, B. Kim, K. Asada, and T. Sugano, IEEE Trans. Electron Devices **35**, 2245 (1988).
[12]K. Awazu, W. Watanabe, and H. Kawazoe, J. Appl. Phys. **73**, 8519 (1993).
[13]T. Takahashi, B. B. Triplett, K. Yokogawa, and T. Sugano, Appl. Phys. Lett. **26**, 1339 (1987).
[14]W. Carlos, Nucl. Instrum. Methods Phys. Res. B **1**, 383 (1984).
[15]P. M. Lenahan, J. F. Conley, Jr., and B. D. Wallace, J. Appl. Phys. **81**, 6824 (1997).
[16]P. M. Lenahan and J. F. Conley, Jr., Appl. Phys. Lett. **71**, 3128 (1997).
[17]J. F. Conley, Jr. and P. M. Lenahan in *The Physics and Chemistry of SiO_2 and the Si/SiO_2 Interface III*, Proceedings of the Electrochemical Society, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (Electrochemical Society, Pennington, NJ, 1996), Vol. 96-1, p. 214.
[18]J. H. Stathis and L. Dori, Appl. Phys. Lett. **58**, 1641 (1991).
[19]E. Cartier and J. H. Stathis, Microelectron. Eng. **28**, 3 (1995).
[20]C. P. Slichter, *Principles of Magnetic Resonance*, 2nd ed. (Springer, Berlin 1978), Chap. 4.
[21]H. A. Schafft, D. A. Baglee, and P. E. Kennedy, Proceedings of the IEEE International Reliability Symposium (IRPS), 1991, pp. 1–7.
[22]K. L. Brower, Appl. Phys. Lett. **43**, 1111 (1983).
[23]M. A. Jupina and P. M. Lenahan, IEEE Trans. Nucl. Sci. **37**, 1650 (1990).
[24]J. W. Gabrys, P. M. Lenahan, and W. Weber, Microelectron. Eng. **22**, 273 (1993).
[25]J. L. Cantin, M. Schoisswohl, H. J. von Bardeleben, V. Morazzani, J. J. Ganem, and I. Trimaille, in *The Physics and Chemistry of SiO_2 and the Si/SiO_2 Interface III*, Proceedings of the Electrochemical Society, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (The Electrochemical Society, Pennington, NJ, 1996), Vol. 96-1, p. 28.
[26]K. L. Brower, Z. Phys. Chem., Neue Folge **151**, 177 (1987).
[27]E. H. Poindexter, P. J. Caplan, J. J. Finegan, N. M. Johnson, D. K. Biegelson, and M. D. Moyer, in *The Physics and Technology of MOS Insulators*, edited by G. Lucovsky, S. T. Panelides, and F. L. Galeener (Pergamon, New York, 1980), p. 326.
[28]C. Brunstrom and C. V. Svensson, Solid State Commun. **37**, 339 (1981).
[29]E. H. Poindexter, G. J. Gerardi, M. E. Ruekel, P. J. Caplan, N. M. Johnson, and D. K. Biegelson, J. Appl. Phys. **56**, 2844 (1984).
[30]G. J. Gerardi, E. H. Poindexter, P. J. Kaplan, and N. M. Johnson, Appl. Phys. Lett. **49**, 348 (1986).
[31]D. J. Lepine, Phys. Rev. B **6**, 436 (1972).
[32]A. S. Grove and D. J. Fitzgerald, Solid-State Electron. **9**, 783 (1966).
[33]D. J. Fitzgerald and A. S. Grove, Surf. Sci. **9**, 347 (1968).
[34]E. H. Snow, A. S. Grove, and D. J. Fitzgerald, Proc. IEEE **55**, 1168 (1967).
[35]J. Dong and D. Drabold, Phys. Rev. Lett. **80**, 1928 (1998).
[36]P. M. Lenahan and P. V. Dressendorfer, IEEE Trans. Nucl. Sci. **30**, 4602 (1983).
[37]R. E. Mikawa and P. M. Lenahan, J. Appl. Phys. **59**, 2054 (1986).
[38]W. L. Warren and P. M. Lenahan, Appl. Phys. Lett. **49**, 1296 (1986).
[39]J. T. Krick, P. M. Lenahan, and G. J. Dunn, Appl. Phys. Lett. **59**, 3437 (1991).
[40]T. E. Tsai, D. L. Griscom, and E. J. Friebele, Phys. Rev. B **40**, 6374 (1989).
[41]P. S. Winokur and H. E. Boesch, Jr., IEEE Trans. Nucl. Sci. **NS-28**, 4088 (1981).
[42]P. S. Winokur, J. M. McGarrity, and H. E. Boesch, Jr., IEEE Trans. Nucl. Sci. **NS-23**, 1580 (1976).
[43]N. M. Johnson, in *Hydrogen in Semiconductors*, edited by J. Pankove and N. M. Johnson (Academic, San Diego, 1991), p. 113.
[44]E. Cartier, J. H. Stathis, and D. A. Buchanan, Appl. Phys. Lett. **63**, 1510 (1993).
[45]C. R. Viswanathan and J. Maserjian, IEEE Trans. Nucl. Sci. **NS-23**, 1540 (1976).
[46]G. F. Derbenwick and B. L. Gregory, IEEE Trans. Nucl. Sci. **NS-22**, 2151 (1975).
[47]N. S. Saks, M. G. Ancona, and J. A. Modolo, IEEE Trans. Nucl. Sci. **NS-33**, 1185 (1986).
[48]R. H. Silsbee, J. Appl. Phys. **32**, 1459 (1961).
[49]F. J. Feigl, W. B. Fowler, and K. L. Yip, Solid State Commun. **14**, 225 (1974).
[50]P. M. Lenahan and P. V. Dressendorfer, IEEE Trans. Nucl. Sci. **29**, 1459 (1982).
[51]C. L. Marquardt and G. H. Sigel, Jr., IEEE Trans. Nucl. Sci. **22**, 2234 (1975).
[52]T. Takahashi, B. B. Triplett, K. Yokogawa, and T. Sugano, Appl. Phys. Lett. **26**, 1334 (1987).
[53]B. B. Triplett, T. Takahishi, and T. Sugano, Appl. Phys. Lett. **50**, 1663 (1987).
[54]L. Lipkin, L. Rowan, A. Reisman, and C. K. Williams, J. Electrochem. Soc. **138**, 2050 (1991).
[55]T. E. Tsai and D. L. Griscom, J. Non-Cryst. Solids **91**, 170 (1987).
[56]J. Vitko, J. Appl. Phys. **49**, 5530 (1978).
[57]J. F. Conley, Jr. and P. M. Lenahan, IEEE Trans. Nucl. Sci. **39**, 2186 (1992).
[58]J. F. Conley, Jr. and P. M. Lenahan, IEEE Trans. Nucl. Sci. **40**, (1993).
[59]R. A. Kohler, R. A. Kushner, and K. H. Lee, IEEE Trans. Nucl. Sci. **35**, 1492 (1988).
[60]R. E. Stahlbush, B. J. Mrstik, and R. K. Lawrence, IEEE Trans. Nucl. Sci. **37**, 1641 (1990).
[61]J. F. Conley, Jr., P. M. Lenahan, and P. Roitman, IEEE Trans. Nucl. Sci. **39**, 2114 (1992).
[62]J. F. Conley, Jr., P. M. Lenahan, and P. Roitman, Appl. Phys. Lett. **60**, 2889 (1992).
[63]W. L. Warren, M. R. Shaneyfelt, J. R. Schwank, D. M. Fleetwood, P. S. Winokur, R. A. B. Devine, W. P. Maszara, and J. B. McKitterick, IEEE Trans. Nucl. Sci. **40**, 1755 (1993).
[64]K. Vanheusden and A. Stesmans, J. Appl. Phys. **74**, 275 (1993).
[65]J. F. Conley, Jr., P. M. Lenahan, H. L. Evans, R. K. Lowry, and T. J. Morthorst, J. Appl. Phys. **76**, 2872 (1994).
[66]J. F. Conley, Jr. and P. M. Lenahan, Microelectron. Eng. **28**, 35 (1995).
[67]L. Zhang and R. G. Leisure, J. Appl. Phys. **80**, 3744 (1996).
[68]J. R. Chavez, S. P. Karna, K. Vanheusden, C. P. Brothers, R. D. Pugh, B. K. Singaraju, W. L. Warren, and R. A. B. Devine, IEEE Trans. Nucl. Sci. **NS-44**, 1799 (1997).

[69]A. J. Lelis, H. E. Boesch, Jr., T. R. Oldham, and F. B. McLean, IEEE Trans. Nucl. Sci. **35**, 1186 (1988).
[70]A. J. Lelis, Jr., T. R. Oldham, H. E. Boesch, and F. B. McLean, IEEE Trans. Nucl. Sci. **36**, 1186 (1989).
[71]J. F. Conley, Jr., P. M. Lenahan, A. J. Lelis, and T. R. Oldham, Appl. Phys. Lett. **67**, 2179 (1995).
[72]W. L. Warren, M. R. Shaneyfelt, D. M. Fleetwood, J. R. Schwank, and P. S. Winokur, IEEE Trans. Nucl. Sci. **41**, 1817 (1994).
[73]B. Henderson, *Defects in Crystalline Solids* (Crane-Russek, New York, 1972), Chap. 1.
[74]Y. M. Chaing, D. P. Birnie III, and D. M. Kingery, *Physical Ceramics*, (Wiley, New York, 1997), Chap. 2.
[75]T. Ohmameuda, H. Miki, K. Asada, T. Sugano, and Y. Ohji, Jpn. J. Appl. Phys., Part 2 **30**, L1993 (1991).
[76]K. L. Brower, P. M. Lenahan, and P. V. Dressendorfer, J. Appl. Phys. **41**, 251 (1982).
[77]K. L. Brower, Phys. Rev. B **38**, 9657 (1988).
[78]L. Pauling, *General Chemistry* (Dover, New York, 1988).
[79]G. J. Hu and W. C. Johnson, J. Appl. Phys. **54**, 1441 (1983).
[80]T. P. Ma and P. V. Dressendorfer, *Ionizing Radiation Effects in MOS Devices and Circuits* (Wiley, New York, 1989), and references therein.
[81]F. B. McLean, IEEE Trans. Nucl. Sci. **NS-27**, 1651 (1981).
[82]N. S. Saks and M. G. Ancona, IEEE Trans. Nucl. Sci. **NS-34**, 1348 (1987).
[83]H. E. Boesch, IEEE Trans. Nucl. Sci. **NS-35**, 1160 (1988).
[84]J. A. Chaiyasena, P. M. Lenahan, and G. J. Dunn, Appl. Phys. Lett. **58**, 2141 (1991).
[85]J. T. Yount, P. M. Lenahan, and G. J. Dunn, IEEE Trans. Nucl. Sci. **39**, 2211 (1992).
[86]J. T. Yount and P. M. Lenahan, J. Non-Cryst. Solids **164–66**, 1069 (1993).
[87]J. T. Yount, P. M. Lenahan, and J. T. Krick, J. Appl. Phys. **76**, 1754 (1994).
[88]J. T. Yount, Ph.D. thesis, The Pennsylvania State University, 1995.
[89]J. H. Mackey, J. W. Boss, and M. Kopp, Phys. Chem. Glasses **II**, 205 (1970).
[90]J. H. Stathis and M. A. Kastner, Phys. Rev. B **29**, 7079 (1984).
[91]T. E. Tsai, D. L. Griscom, and E. J. Friebele, Phys. Rev. B **38**, 2140 (1988).
[92]W. R. Austin and R. G. Leisure, Phys. Rev. B **54**, 15064 (1996).
[93]W. L. Warren, M. R. Shaneyfelt, D. M. Fleetwood, P. S. Winokur, and S. Montague, IEEE Trans. Nucl. Sci. **NS-42**, 1731 (1995).
[94]R. Fuller, H. Evans, C. Gamlin, B. Czagas, M. Morrison, D. Decrosta, R. Lowry, P. M. Lenahan, and C. J. Frye, IEEE Trans. Nucl. Sci. **43**, 2565 (1996).
[95]C. A. Billman, P. M. Lenahan, R. Fuller, H. Evans, W. H. Speece, D. DeCrosta, and R. Lowry, IEEE Trans. Nucl. Sci. **44**, 1834 (1997).
[96]D. L. Griscom, E. J. Friebele, K. J. Long, and J. W. Fleming, J. Appl. Phys. **43**, 960 (1976).
[97]D. L. Griscom, G. H. Sigel, Jr., and R. J. Ginther, J. Appl. Phys. **54**, 3743 (1983).
[98]P. M. Lenahan and P. V. Dressendorfer, IEEE Trans. Nucl. Sci. **NS-29**, 1459 (1982).

Interface formation and thermal stability of advanced metal gate and ultrathin gate dielectric layers

B. Claflin[a] and G. Lucovsky
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 21 January 1998; accepted 28 May 1998)

The compatibility of metallic titanium nitride (TiN_x) films for advanced gate electrodes and remote plasma enhanced chemical vapor deposited silicon oxide (SiO_2) or silicon oxide/silicon nitride (Si_3N_4) advanced gate dielectric layers is investigated by interrupted growth and on-line rapid thermal annealing using on-line Auger electron spectroscopy. Growth of TiN_x on SiO_2 and Si_3N_4 occurs uniformly without a titanium seed layer. TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces are chemically stable against reaction for rapid thermal annealing treatments below 850 °C. Metal–oxide–semiconductor capacitors using TiN_x gate contacts and SiO_2 or SiO_2/Si_3N_4 gate dielectrics exhibit excellent $C-V$ characteristics. The measured TiN_x/SiO_2 barrier height in these devices is $\Phi_b = 3.7 \pm 0.1$ eV. The observed difference in fixed charge for SiO_2 and SiO_2/Si_3N_4 dielectrics is briefly discussed in terms of a new interface dipole model. © *1998 American Vacuum Society.* [S0734-211X(98)11704-3]

I. INTRODUCTION

The continuing effort to reduce metal–oxide–semiconductor (MOS) device dimensions has led to considerable interest in silicon oxide[1] (SiO_2) and silicon oxide/silicon nitride[2,3] (Si_3N_4) layered films with an equivalent oxide thickness $t_{ox\text{-}eq} < 3$ nm for advanced gate dielectrics. Similarly, to maintain proper device scaling and to avoid the effects of gate electrode depletion, attention is being focused on advanced gate materials such as metallic titanium nitride[4–8] (TiN_x) as a possible replacement for heavily doped polycrystalline silicon. However, before such materials can be incorporated into device production they must be shown to be compatible with the other materials and with subsequent process conditions required for fabrication. Of particular concern is the possibility of chemical reaction at the interface between these advanced dielectric and gate materials at elevated temperatures.[9,10] For example, the gate metal/dielectric interface must be able to withstand rapid thermal annealing (RTA) temperatures of at least 800–1000 °C which are required to activate dopants in source and drain structures.

In this article, the interface formation between reactively sputtered TiN_x metallic films and SiO_2 or SiO_2/Si_3N_4 dielectric layers prepared by low-temperature remote plasma enhanced chemical vapor deposition (RPECVD) is investigated as well as the chemical stability of these interfaces following on-line RTA treatments up to 850 °C. The evolution of the TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces is investigated from their initial formation at the monolayer level to the buried interface by interrupted film growth using on-line Auger electron spectroscopy (AES). The stability of TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces against reaction during RTA treatments at temperatures below 850 °C is also established by on-line AES. Finally, capacitance–voltage ($C-V$) characteristics of MOS capacitors employing TiN_x gate electrodes with ultrathin single layer SiO_2 and dual layer SiO_2/Si_3N_4 gate dielectrics are presented. These $C-V$ results are compared with similar MOS devices made using Al gate contacts.

II. EXPERIMENT

Interrupted growth and chemical stability experiments were conducted on 3 in. *p*-type Si(100) wafers on which either a single layer SiO_2 dielectric or a dual layer SiO_2/Si_3N_4 dielectric stack was deposited. In both cases, the SiO_2 layer was 5 nm thick and was prepared by a two-step plasma assisted oxidation/RPECVD process[1] which is described in detail elsewhere. The Si_3N_4 layer was 3 nm thick and was deposited by RPECVD with NH_3 and SiH_4 source gases injected downstream[2,11,12] at a ratio of 10:1. Prior to TiN_x deposition, the SiO_2 and SiO_2/Si_3N_4 dielectric layers were subjected to a 30 s RTA treatment at 900 °C to improve their electrical properties by reducing bond defects.[13,14] This bonding rearrangement was verified by Fourier transform infrared (FTIR) spectroscopy.

Metallic TiN_x films were deposited at room temperature by reactive dc magnetron sputtering of a 99.9% Ti target at 50 W with a 5:1 mixture of 99.999% Ar and 99.999% N_2 at a chamber pressure of 6 mTorr. Several ultrathin TiN_x layers were deposited sequentially and AES spectra collected until the Si_{LVV} Auger peak could no longer be observed, indicating that the interface had been buried. Thicker TiN_x layers (each about 0.5–1.0 nm) were then added until bulk characteristics were observed. To study the chemical stability of the metal/dielectric interfaces, a 0.6–1.0 nm thick TiN_x layer was deposited on SiO_2 or SiO_2/Si_3N_4 and annealed in vacuum for 180 s in 100 °C steps from 350 to 850 °C. Following each on-line RTA treatment, AES spectra were collected.

MOS capacitors were prepared on *n*-type Si(100) wafers with ultrathin SiO_2 or SiO_2/Si_3N_4 dielectrics and TiN_x or Al gate electrodes. The wafers were cleaned, a thermal field

[a]Electronic mail: claflin@ncsu.edu

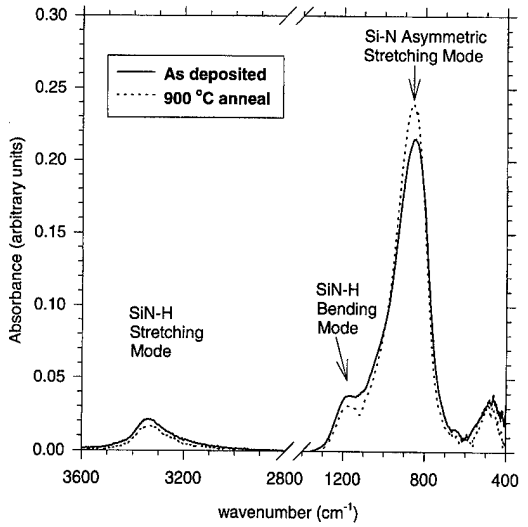


FIG. 1. FTIR spectra of RPECVD Si_3N_4 dielectric as deposited and following a 30 s 900 °C RTA treatment. The decrease of SiN–H stretch and bending modes while the Si–N asymmetric stretch increases after annealing indicates conversion of SiN–H dangling bonds to Si–N network bonds.

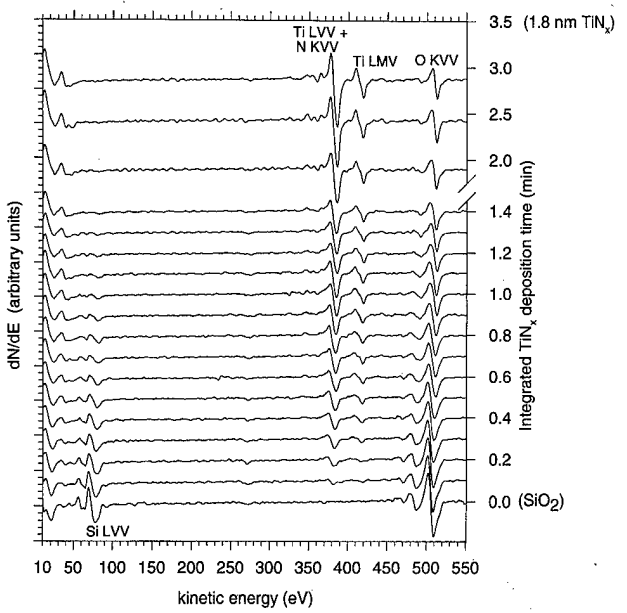


FIG. 2. Interrupted growth of TiN_x on RPECVD SiO_2 measured by AES. Si_{LVV} and O_{KVV} peaks decrease while $Ti_{LVV}+N_{KVV}$ and Ti_{LMV} peaks increase with additional TiN_x deposition. The asymmetry of $Ti_{LVV}+N_{KVV}$ and Ti_{LMV} intensities indicates no Ti seed layer is required.

oxide was grown and patterned, and the active area wet chemically etched. The wafers were cleaned again prior to RPECVD deposition of a 3.5 nm SiO_2 gate dielectric or a 2.0 nm SiO_2/2.0 nm Si_3N_4 dielectric bilayer. The wafers were patterned for liftoff and a 50 nm TiN_x or 100 nm Al gate electrode deposited by sputtering. A 100 nm thick Al layer was sputtered on top of the TiN_x gate to facilitate external contact. Liftoff was performed, Al back contacts evaporated, and the samples were annealed at 300 °C in H_2 for 1800 s. Standard high frequency (100 kHz) $C-V$ measurements were conducted to determine the device characteristics.

III. RESULTS AND DISCUSSION

To evaluate the integration of TiN_x and RPECVD dielectrics with existing Si technology and process flows it is critical to use optimal preparation conditions for these new materials. It has been shown[13] previously that a 15 s, 900 °C RTA treatment improves the dielectric quality of RPECVD SiO_2 films by reducing the amount of suboxide (SiO_x with $x<2$) at the Si/SiO_2 interface and hence the number of electrically active dangling bond defects. A similar improvement in the dielectric reliability of RPECVD Si_3N_4 films occurs[14] by annealing at 900 °C for 30 s and can be monitored with FTIR as seen in Fig. 1. The spectrum displays three main features: a Si–N asymmetric stretching mode at approximately 850 cm^{-1}, a SiN–H bending mode at 1200 cm^{-1}, and a SiN–H stretching mode near 330 cm^{-1}. Figure 1 shows a decrease in the intensity of SiN–H bending and stretching modes while the Si–N asymmetric stretch increases from as-deposited levels following a 900 °C RTA treatment. These changes in the IR spectra have recently been explained[14] as resulting from a conversion of correlated Si–H and SiN–H dangling bonds into Si–N network bonds.

SiO_2 and SiO_2/Si_3N_4 dielectric films optimized in this fashion were used to study the interface formation, growth, and chemical stability of TiN_x. Derivative AES spectra demonstrating the growth of TiN_x on RPECVD SiO_2 are provided in Fig. 2 with the relevant Auger transitions indicated. The bottom curve shows the spectrum for the bare SiO_2 dielectric layer. The vertical offset of subsequent curves corresponds to the integrated TiN_x sputtering time elapsed, indicated by the right-hand axis. Film thickness was estimated from the bulk deposition rate of 1.0 nm/min which was determined by cross-sectional scanning electron microscopy on thick samples.

A quantative difficulty which arises in the TiN_x system is that the Ti_{LVV} and N_{KVV} AES peaks overlap[15–17] at 385 eV; however the Ti_{LMV} feature at 420 eV does not suffer from this deficiency. For TiN_x, the $Ti_{LVV}+N_{KVV}$ and Ti_{LMV} features exhibit a peak to peak intensity ratio that is about 2:1. In contrast, for pure Ti the intensity of these two peaks would be nearly equal.[8] In Fig. 2 the $Ti_{LVV}+N_{KVV}$ feature is seen to form and grow on SiO_2 as additional layers of TiN_x are added and the intensity of this peak is always about twice that of Ti_{LMV}. This asymmetry in these peak intensities, which occurs even for submonolayer coverage, indicates that TiN_x grows directly on SiO_2 without the need for an intermediate Ti seed layer. The Si_{LVV} peak intensity at 76 eV decreases as successive TiN_x layers are deposited and disappears entirely for a metal thickness greater than about 0.8 nm. The O_{KVV} peak near 510 eV also initially decreases; however a small amount of parasitic oxygen is always observed even after the TiN_x/SiO_2 interface is completely buried and bulk TiN_x features are observed.

Figure 3 shows similar derivative AES spectra for interrupted TiN_x growth on Si_3N_4. Because the dielectric contains nitrogen, the initial development of the $Ti_{LVV}+N_{KVV}$ feature is impossible to observe. However, the Ti_{LMV} peak evolution

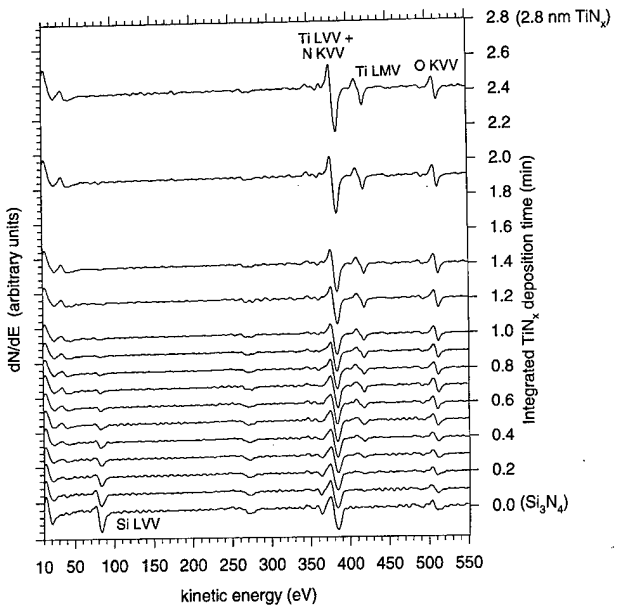


FIG. 3. Interrupted growth of TiN_x on RPECVD Si_3N_4 measured by AES. The Si_{LVV} peak decreases while the $Ti_{LVV}+N_{KVV}$ and Ti_{LMV} peaks increase with additional TiN_x deposition. The similarity in growth of the Ti_{LMV} peak to Fig. 2 suggests that no seed layer is required.

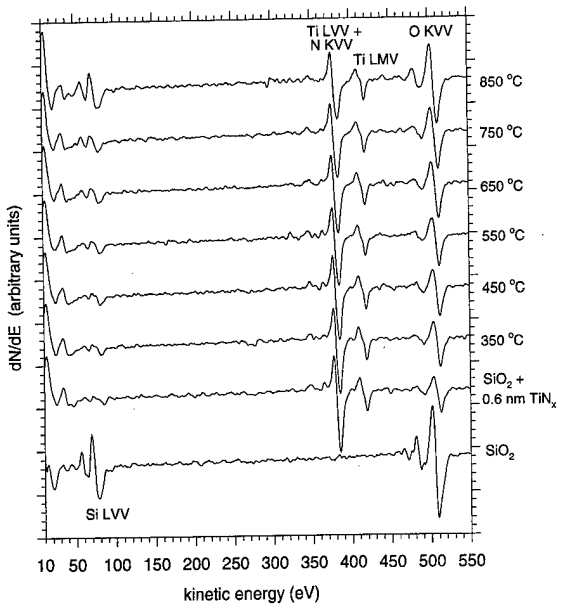


FIG. 5. Chemical stability of TiN_x/SiO_2 interface following the RTA treatments indicated. The dramatic increase of the Si_{LVV} AES peak at 850 °C indicates that a reaction has occurred.

on Si_3N_4 is similar to the behavior observed in Fig. 2 on SiO_2 and suggests that TiN_x grows directly on Si_3N_4 without a buffer layer. The Si_{LVV} peak intensity in Fig. 3 is diminished by successive TiN_x layers and disappears above about 0.8 nm as expected.

The decrease of the Si_{LVV} peak intensity normalized to its value in the bare dielectric is plotted on a log scale as a function of TiN_x deposition time in Fig. 4 for both the SiO_2 and the SiO_2/Si_3N_4 substrates. In both cases the intensity decreases exponentially with TiN_x thickness although the slopes differ by about a factor of 2. This exponential decrease of the Si_{LVV} peak intensity with TiN_x thickness indicates uniform layer growth on both SiO_2 and Si_3N_4. The difference in the slopes of these curves corresponds to different deposition rates due to target aging. Detailed kinetics of the interface formation and morphology of the TiN_x layer in each case require further investigation.

Derivative AES spectra for TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces subjected to on-line RTA treatments up to 850 °C are shown in Figs. 5 and 6, respectively. In both cases, a

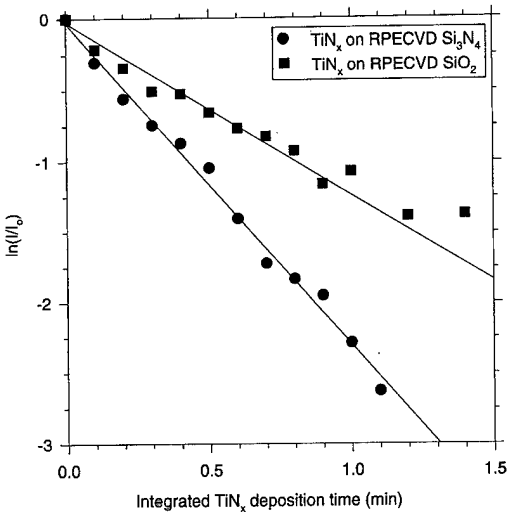


FIG. 4. Decrease of normalized Si_{LVV} AES peak intensity with increasing TiN_x deposition on SiO_2 and Si_3N_4 dielectric layers. The exponential decay indicates uniform TiN_x layer growth on both dielectrics.

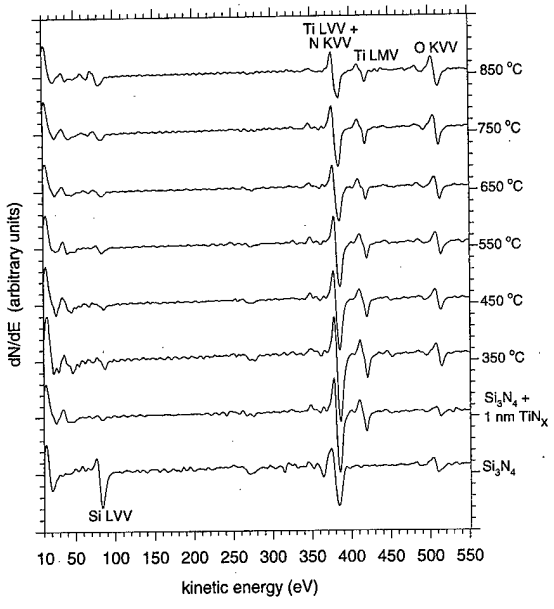


FIG. 6. Chemical stability of TiN_x/Si_3N_4 interface following the RTA treatments indicated. A slight increase of the Si_{LVV} AES peak at 850 °C indicates that the interface is chemically stable.

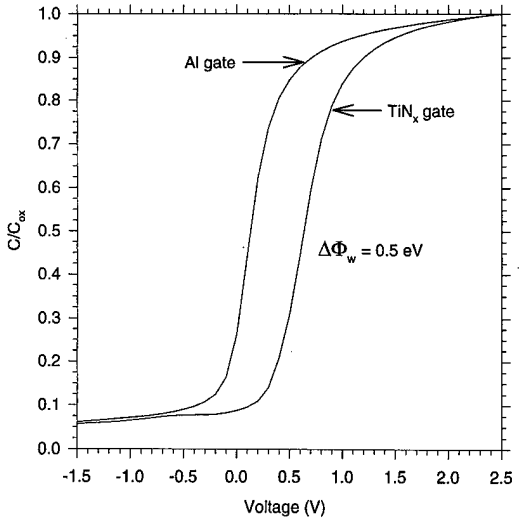


FIG. 7. High frequency $C-V$ data for MOS capacitors using TiN_x or Al gates and a 3.5 nm RPECVD SiO_2 gate dielectric. The flatband voltage shift corresponds to the work function difference between TiN_x and Al.

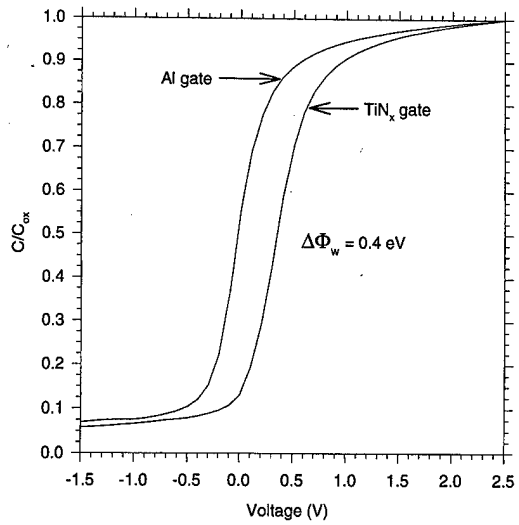


FIG. 8. High frequency $C-V$ data for MOS capacitors using TiN_x or Al gates and a 2 nm RPECVD SiO_2/2 nm RPECVD Si_3N_4 gate dielectric stack. The flatband voltage shift corresponds to the work function difference between TiN_x and Al.

minimal thickness of TiN_x was deposited to bury the interface, reducing the Si_{LVV} peak intensity to nearly zero. Only minor changes in the AES spectra for both TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces are observed below 850 °C and are most likely the result of slight variations in thickness of the TiN_x overlayer. Following the 850 °C anneal, however, the Si_{LVV} peak intensity increases substantially for the TiN_x/SiO_2 interface and modestly for TiN_x/Si_3N_4. The fact that these changes in the AES spectra occur at the same temperature for both dielectrics suggests that decomposition of TiN_x begins at 850 °C. The dramatic change in the TiN_x/SiO_2 interface compared to that of TiN_x/Si_3N_4 indicates that the metal chemically reacts with the underlying SiO_2 layer but does not interact with Si_3N_4. The details of this reaction and the chemical structure that results, including the possible formation of a titanium silicide layer, still need to be elucidated.

High frequency $C-V$ characteristics for MOS capacitors using TiN_x and Al gates on ultrathin RPECVD SiO_2 dielectrics are shown in Fig. 7. The flatband voltage for the TiN_x gate is shifted by $\Delta\Phi=0.5\pm0.1$ eV above Al. A similar shift $\Delta\Phi=0.4\pm0.1$ eV is observed in Fig. 8 for capacitors with TiN_x and Al gates on a SiO_2/Si_3N_4 bilayer. In each case, the flatband voltage shift results from the work function difference of the two metals. Using the known[18] barrier height of 3.2 eV for Al on SiO_2, the observed flatband voltage shift gives a barrier height of $\Phi_b=3.7\pm0.1$ eV for TiN_x/SiO_2, in agreement with previous studies.[8] Comparing the two devices with TiN_x gates, there is an additional flatband voltage shift $\Delta\Phi=0.3$ eV between the capacitors using SiO_2 and SiO_2/Si_3N_4 dielectrics as seen in Fig. 9. This flatband voltage shift is not due to a work function difference between the contacts since the same gate metal and substrate were used in each device. Instead, this difference in flatband voltage must result from fixed positive charge in the dielectric. However, since the wafers were processed in parallel and subjected to the same treatments at each step except for deposition of the dielectric layer, it is assumed that this fixed positive charge is intrinsic to the dielectric and does not result from device processing. The most likely source of this fixed charge is an interface dipole layer in the dielectric. Recent calculations[19] suggest that dipole layers between Si/SiO_x and Si/Si_3N_4 interfaces may account for the observed flatband voltage shift.

IV. CONCLUSIONS

As MOS device dimensions are reduced, the impact of interface physics on their operation and performance increases dramatically. Understanding, controlling, and optimizing interface properties will be critical to continued de-

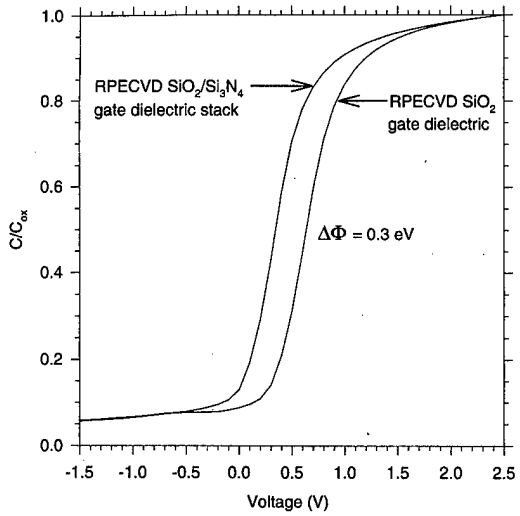


FIG. 9. Comparison of high frequency $C-V$ curves for MOS capacitors using TiN_x gates and ultrathin SiO_2 or SiO_2/Si_3N_4 gate dielectrics from Figs. 7 and 8. The flatband voltage shift due to fixed charge may result from interface dipoles.

vice scaling. In particular, issues relating to the integration of novel gate dielectrics and gate metals with existing Si technology and process flows must be investigated and verified.

Metallic TiN_x gates are compatible with advanced RPECVD SiO_2 dielectric layers and SiO_2/Si_3N_4 dielectric stacks. TiN_x grows uniformly on these dielectrics without an intermediate seed layer. TiN_x/SiO_2 and TiN_x/Si_3N_4 interfaces are stable against chemical reaction for RTA treatments below 850 °C. At 850 °C, the metal reacts with SiO_2 possibly due to decomposition of TiN_x, but is more stable on Si_3N_4.

MOS capacitors have been fabricated which integrate TiN_x gate electrodes with RPECVD SiO_2 or SiO_2/Si_3N_4 gate dielectrics and exhibit good $C-V$ characteristics. From the flatband voltage shift of the $C-V$ curves for TiN_x gates relative to Al, a barrier height of $\Phi_b = 3.7 \pm 0.1$ eV for TiN_x on SiO_2 is determined. In addition, a flatband voltage shift $\Delta\Phi = 0.3$ eV is observed between TiN_x gates on SiO_2 and SiO_2/Si_3N_4 gate dielectrics. This difference may result from a change in the charge dipole layer in the dielectric.

ACKNOWLEDGMENTS

The authors would like to thank D. M. Wolfe for assistance with RPECVD SiO_2 and Si_3N_4 deposition and M. Binger for assistance with data analysis. This work was supported by the ONR, NSF ERC, and SRC.

[1]T. Yatsuda, Y. Ma, S. Habermehl, and G. Lucovsky, Appl. Phys. Lett. **60**, 434 (1992).

[2]Y. Ma, T. Yatsuda, and G. Lucovsky, J. Vac. Sci. Technol. B **11**, 1533 (1993).

[3]Y. Ma, T. Yatsuda, and G. Lucovsky, J. Vac. Sci. Technol. A **11**, 952 (1993).

[4]M. Wittmer, J. R. Noser, and H. Melchior, J. Appl. Phys. **54**, 1423 (1983).

[5]D. H. Lee, K. H. Yeom, M. H. Cho, N. S. Kang, and T. E. Shim, 1996 Symposium on VLSI Technology, p. 208.

[6]D. H. Lee, S. H. Joo, G. H. Lee, J. Moon, T. E. Shim, and J. G. Lee, 1995 Symposium on VLSI Technology, p. 119.

[7]K. T. Kim, L. G. Kang, T. S. Park, Y. S. Shin, J. K. Park, C. J. Lee, C. G. Hwang, D. Chin, and Y. E. Park, 1990 Symposium on VLSI Technology, p. 115.

[8]B. Claflin, M. Binger, and G. Lucovsky, J. Vac. Sci. Technol. A **16**, 1757 (1998).

[9]S. Q. Wang and J. W. Mayer, J. Appl. Phys. **64**, 4711 (1988).

[10]R. Pretorius, J. M. Harris, and M.-A. Nicolet, Solid-State Electron. **21**, 667 (1978).

[11]Y. Ma, Ph.D. thesis, North Carolina State University, 1993.

[12]M. J. Williams, Ph.D. thesis, North Carolina State University, 1994.

[13]G. Lucovsky, A. Banerjee, B. Hinds, B. Claflin, K. Koh, and H. Yang, J. Vac. Sci. Technol. B **15**, 1074 (1997), and references therein.

[14]Z. Lu, P. Santos-Filho, G. Stevens, M. J. Williams, and G. Lucovsky, J. Vac. Sci. Technol. A **13**, 607 (1995).

[15]W. Pamler, Surf. Interface Anal. **13**, 55 (1988).

[16]A. J. Perry, C. Strandberg, W. D. Sproul, S. Hoffman, C. Ernsberger, J. Nickerson, and L. Chollet, Thin Solid Films **153**, 169 (1987).

[17]B. J. Burrow, A. E. Morgan, and R. C. Ellwanger, J. Vac. Sci. Technol. A **4**, 2463 (1986).

[18]B. E. Deal, E. H. Snow, and C. A. Mead, J. Phys. Chem. Solids **27**, 1873 (1966).

[19]H. Z. Massoud, G. Lucovsky, and H. Yang, J. Vac. Sci. Technol. B, these proceedings.

Formation of ultrathin oxynitride layers on Si(100) by low-temperature electron cyclotron resonance N_2O plasma oxynitridation process

Sahana Chakraborty, Toshiyuki Yoshida, Tamotsu Hashizume, and Hideki Hasegawa
Research Center for Interface Quantum Electronics, and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo 060, Japan

Takamasa Sakai
Dainippon Screen Manufacturing Co., Ltd., 322 Furukawa-cho, Hazukashi, Fushimi-ku, Kyoto 612, Japan

(Received 21 January 1998; accepted 11 May 1998)

We have undertaken the characterization of the low-temperature direct electron cyclotron resonance (ECR) N_2O plasma process on the technically important Si(100) surface by the combined use of the newly developed ultrahigh vacuum (UHV)-based contactless capacitance–voltage ($C-V$) and x-ray photoelectron spectroscopy (XPS) techniques. An UHV-based contactless $C-V$ measurement system allowed for the first time a direct *in situ* assessment of electronic properties of "free" Si surfaces and the surfaces covered with an ultrathin insulator layer. The UHV contactless $C-V$ analysis indicated that a oxynitride/Si interface with a minimum interface state density of 1×10^{11} eV^{-1} cm^{-2} could be realized by the ECR N_2O oxynitridation process at 400 °C. The angle-resolved XPS study showed that the formation of a phase-separated Si_3N_4/SiO_2 structure is important to realize such a interface at low temperatures. The electronic and the chemical properties of the interfaces were found to be strongly correlated. © *1998 American Vacuum Society.*
[S0734-211X(98)05104-X]

I. INTRODUCTION

Ever since the invention of the transistor, there has been a spectacular growth in the silicon technology as the mainstream. This is mainly because only this technology has been allowing systematic realization of higher packaging densities and therefore functional capabilities through continued scaling down of the device feature sizes. Thus, it is anticipated that the future ultimately scaled-down complementary metal–oxide–semiconductor (CMOS) devices will require circa 2010 feature sizes into the decananometer range according to the well-known SIA Roadmap.[1] Since sizes in such a range are comparable with the de Broglie wavelengths of electrons in Si, novel devices in the quantum regime such as Si-based single electron transistors[2] are emerging.

The gate dielectric/Si(100) interface has been a key issue in the development of Si-based ultralarge scale integrated (ULSI) technology. This will also be the case for the future development of a suitable technology to form advanced Si CMOS devices as well as for a high quality ultrathin insulator on Si whose interface is completely free of Fermi-level pinning and whose thickness is highly controllable down to monolayer levels, which is a crucial issue.

Currently oxynitrides/nitrided oxides are extensively studied for these purposes because of their superior electrical properties.[3–7] But, they are grown at rather high temperatures. From the viewpoint of controllability and thermal budget, a low-temperature process is favorable for the formation of such ultrathin insulators. We have been investigating in detail various low-temperature processes. We carried out a detail characterization of thick thermally oxidized surface, hydrogen terminated surface, chemically oxidized surface, and an air-exposed Si(111) surface, using contactless capacitance–voltage ($C-V$), x-ray photoelectron spectroscopy (XPS), and photoluminescence surface state spectroscopy techniques.[8,9] But realization of an unpinned ultrathin-oxide Si interface by such low-temperature processes was found to be extremely difficult.[9] It was also found that low-temperature (400 °C) thermally oxidized and subsequently N_2 plasma-treated surfaces produced SiO_2/Si interface with a high density of interface states. Only recently, we have arrived at a preliminary promising result that a low-temperature direct electron cyclotron resonance (ECR) plasma oxynitridation process gives, rather well-behaved insulator–semiconductor interface on Si(111) surface.[10] This process is also suited for manufacturing of devices.

The purpose of the present study is to further optimize our low-temperature direct ECR N_2O plasma process on the more technically important Si(100) surface through direct *in situ* characterization of the interface by the combined use of *in situ* ultrahigh vacuum (UHV)-based contactless $C-V$ and XPS methods. Oxynitridation and *in situ* characterization were carried out using an ultrahigh vacuum (UHV)-based multichamber system shown in Fig. 1(a). For this purpose, an UHV-based contactless $C-V$ measurement system was newly developed, and applied for the first time. The results indicate that there exist a strong correlation between the chemical and electronic properties of the interfaces.

II. EXPERIMENT

A. Sample preparation and oxynitridation

n-type Si(100) wafers were cut to chips of 15 mm×15 mm dimension and mounted on the standard molybdenum sample holder used for molecular beam epitaxial (MBE) growth. Before loading into the UHV multichamber system, sample surfaces were subjected to a BHF- and NH_4F-based

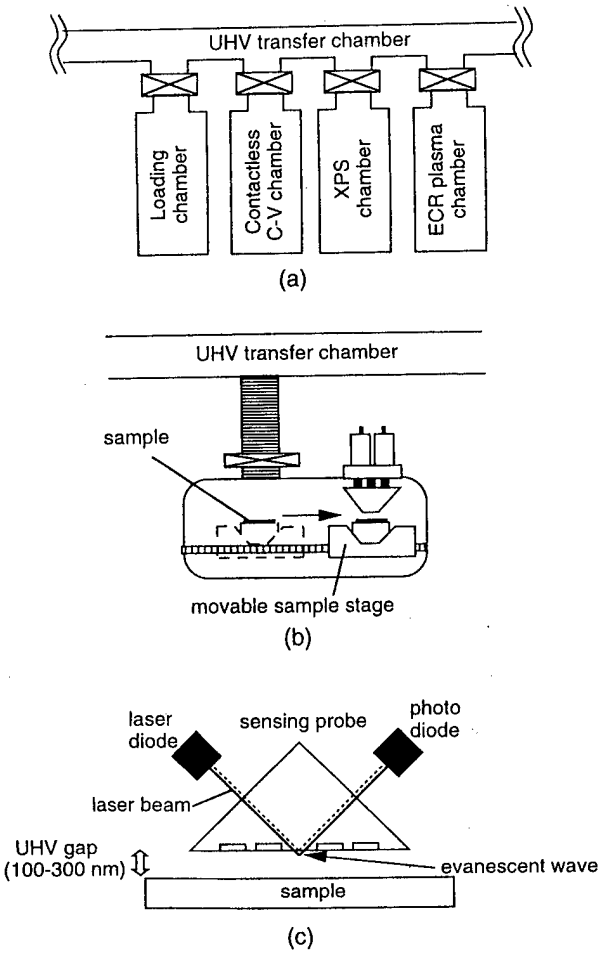


FIG. 1. (a) UHV-based integrated fabrication/characterization system; (b) schematic view of the UHV $C-V$ chamber; and (c) detailed structure of the field plate head and the sensing probe.

hydrogen termination treatment[11] in order to prepare well-defined and reproducible initial surfaces. As an additional treatment, some samples were subsequently treated with an UHV-ECR H_2 plasma after loading into the ECR plasma chamber. The treatment was made for 1 min at 200 °C with a gas flow rate of 10 sccm at a microwave power of 100 W ($f=2.45$ GHz). Then oxynitridation was carried out at 400 °C for 5 min by exposing the surface directly to the ECR N_2O plasma with the microwave power of 100 W with a gas flow rate of 5 sccm (gas pressure was $\sim 10^{-4}$ Torr).

B. Contactless *C–V* method

The principle of the contactless $C-V$ method is simple and is to carry out $C-V$ measurement from a field plate that is separated from the sample by a constant "UHV gap" of 100–300 nm. Thus a metal–insulator–semiconductor assessment of the "free" surface covered with an ultrathin insulator becomes possible in UHV environment by this technique without changing the surface properties and without running into leakage problems due to tunneling and electrical breakdown.

In the present study extensive efforts were made to modify an air-gap based commercial measurement system (CV-8000, Dainippon Screen Mfg. Co., Ltd.) into an UHV-compatible system with a suitable UHV sample transfer mechanism. A schematic view of the completed UHV $C-V$ chamber is shown in Fig. 1(b). Care was taken to make all parts "bakable" to achieve the base pressure of 8×10^{-10} Torr. The detailed structure of the field plate head is given in Fig. 1(c). The head maintains a parallelism and a constant distance between the field plate and sample surface as shown in Fig. 1(c). For this the head has three additional "parallelism electrodes" maintaining a constant distance of 100–300 nm via a piezomechanism with a capacitance feedback. The UHV gap was determined by the optical method utilizing the Goos–Haechen effect, i.e., considering change in reflectivity due to penetration of the evanescent wave.[12,13] Chromium was used as the field plate metal and the electrode area was 7.5×10^{-3} cm^{-2}.

The validity of this newly developed UHV $C-V$ system was checked by using a SiO_2/Si metal–oxide–semiconductor (MOS) system. A typical $C-V$ curve obtained from a standard MOS capacitor with a SiO_2 thickness of 100 nm prepared by high-temperature (1000 °C) thermal oxidization using dry oxygen gas is shown in Fig. 2(a). The obtained contactless $C-V$ curve is in agreement with the calculated ideal $C-V$ curve (indicated by broken line). For determination of the carrier density of the Si substrate, we used the $C-V$ data under deep depletion condition. Figure 2(b) shows the $C-V$ result on the MOS capacitor under pulsed bias mode, i.e., the gate voltage was supplied by pulse wave forms. In this condition, we can get the $C-V$ curve in deep depletion region because the generation rate of minority carrier cannot follow the sweep rate of the pulsed mode bias. A plot of $1/C^2$ vs gate bias voltage (V_G) as obtained from the $C-V$ curve in Fig. 2(b) is shown in Fig. 2(c). The carrier concentration of the n-type Si(100) wafer as estimated from the slope of the curve was 4.3×10^{14} cm^{-3}.

C. XPS technique

The chemical nature of the surface and interface was studied using a Perkin–Elmer PHI (model:1600C multitechnique system) under excitation by monochromated Al $K\alpha$ ($h\nu$ = 1486.7 eV). The angle-resolved XPS spectra of the samples was taken at different pass energies.

III. RESULTS AND DISCUSSION

A. Characterization of the chemically and H_2 plasma-treated Si(100) surfaces prior to oxynitridation

Figure 3 compares the Si $2p$ core-level XPS spectra of the Si(100) surface with and without the H_2 plasma treatment after the wet chemical treatment. Unlike the Si(111) surface where no native oxide component was found after the hydrogen termination treatment,[11] an appreciable amount of oxygen was found on the Si(100) surface even after the BHF- and NH_4F-based surface treatment as seen from Fig. 3. After the ECR H_2 plasma treatment with a power of 100 W for 1

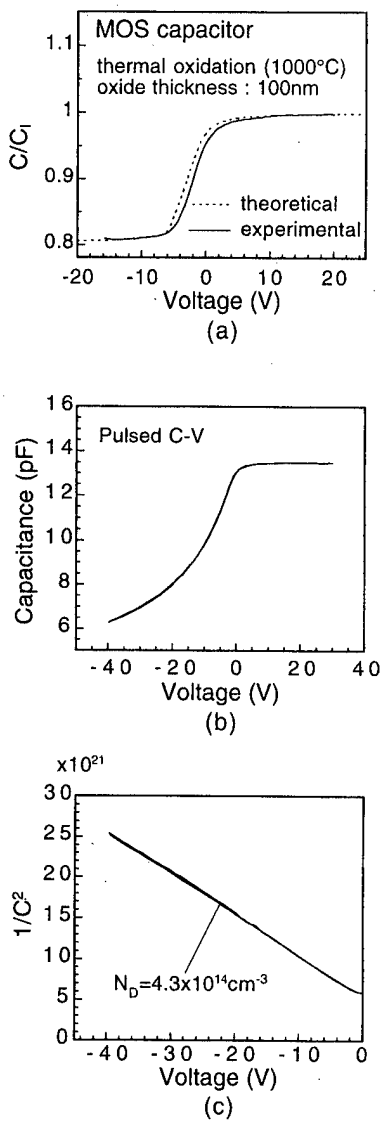


FIG. 2. Contactless $C-V$ curves obtained from standard MOS capacitor under (a) standard bias mode and (b) under pulsed bias mode. (c) $1/C^2$ vs gate bias voltage (V_G) as obtained from $C-V$ curves in (b).

min, the native oxide components on the Si(100) surface were completely removed. It is worthwhile to mention that the H_2 plasma had a very good cleaning effect of carbon and carbon-related contamination phases.

A further detailed analysis of the Si $2p$ XPS spectra was taken up to obtain a deeper insight to the chemical nature of the surface components at Si(100) surfaces. Figures 4(a) and 4(b) show high-resolution Si $2p$ core-level spectra taken at two take-off angles of 70° and 15° for surfaces with and without H_2 plasma treatment, respectively. For both cases, the spectra with a take-off angle of 70° could be successfully deconvulated into the Si $2p_{1/2}$ and Si $2p_{3/2}$ peaks by considering the ratio of the area under the Si $2p_{1/2}$ and Si $2p_{3/2}$ peaks respectively, to be equal to 0.5 and peak energy separation of 0.6 eV.[14] This indicated that the spectra with a take-off angle of 70° include almost all of the bulk component. As compared to this, it is clear that additional surface

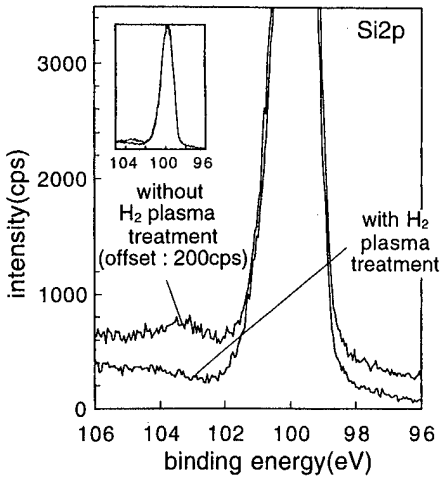


FIG. 3. Si $2p$ core-level XPS spectra of the bare silicon (100) surfaces with/without H_2 plasma treatment after BHF–NH_4F-based wet chemical treatment.

components appear in the spectra with a take-off angle of 15°.

Figure 4(c) compares the spectra obtained from the actual subtraction of the intensities of the x-ray photoelectrons from the 15° spectra to the 70° spectra. The abscissa indicates the energy separation from the Si $2p_{3/2}$ peak position of the 70° spectra. For the surface with the H_2 plasma treatment, the subtracted spectra also had two peaks which could be well fitted by the Si–H (chemical shift: 0.3 eV) and the Si–H_2 (chemical shift: 0.5 eV) related bonds[15–17] considering peak ratio of 0.5 and energy separation of about 0.6 eV, as shown by the broken lines in Fig. 4(c). This indicates that the surface exposed to ECR H_2 plasma was hydrided. Bjorkman *et al.*[17] and Yamamoto and Hasegawa[16] reported that both the monohydride and dihydride components exist at the Si(100) surfaces treated in HF solution and by atomic hydrogen in the UHV chamber. As seen from the upper spectrum in Fig. 4(c), the surface without H_2 plasma treatment seems to also have H-related components. However, it is clearly noticeable that both the intensity and FWHM of the higher-energy peak was remarkably higher than the theoretical ones. This obviously indicated that on this surface, an appreciable amount of native oxide components remained after the BHF- and NH_4F-based wet chemical treatment. Judging from the peak position relative to the Si $2p_{3/2}$ peak, the additional peak seems to originate from the Si^{1+} suboxide component. In fact, Aiba *et al.*[18] reported that the Si^{1+} is the dominant suboxide at the Si(100) surface after BHF treatment and a subsequent low-temperature (300 °C) oxidation in dry O_2. These components could thus be removed on an additional exposure of the Si(100) surface to ECR H_2 plasma.

The contactless $C-V$ measurements were performed on surfaces with and without H_2 plasma treatment. Figure 5(a) shows the results of $C-V$ measurements on the Si(100) surface just after BHF- and NH_4F-based wet chemical treatment. The insulator capacitance C_1 was determined from the optically determined "UHV gap" and the ideal $C-V$ curves (as shown by broken lines) are calculated from the doping

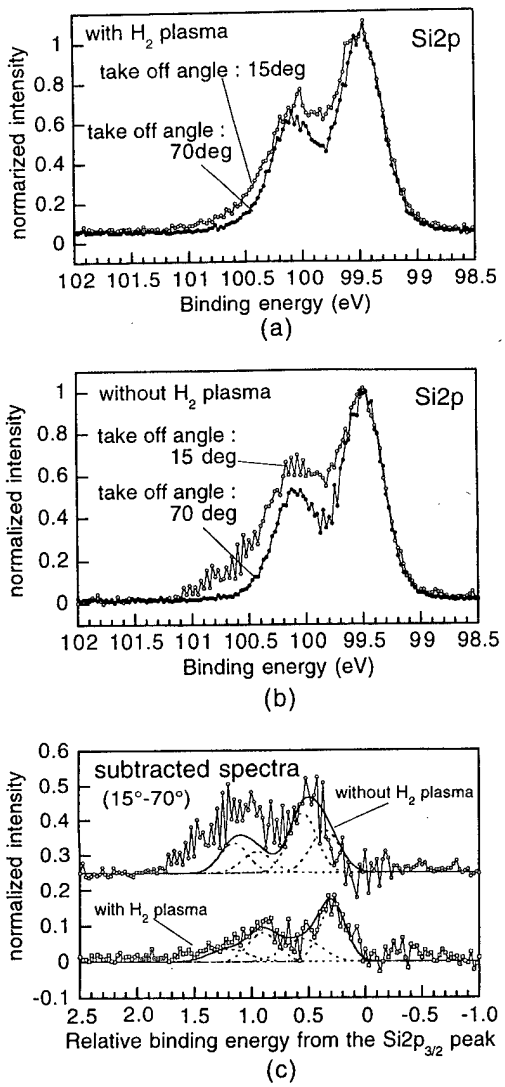


FIG. 4. High resolution Si 2*p* core-level XPS spectra for surfaces (a) with and (b) without H_2 plasma treatment with take-off angles of 15° and 70°. (c) Spectra obtained by subtracting the intensities of the photoelectrons from the 15° spectrum to the 70° spectrum in (a) and (b).

density and insulator capacitance, ignoring the effect of work function difference between the metal field plate and the semiconductor surface. A very large hysteresis was noticed under the negative bias sweep. Such *C–V* behavior is very different from that obtained on the high-temperature thermally oxidized surface as shown in Fig. 2(a). Similar presence of large hysteresis under negative voltage and flattening of capacitance under positive bias were reported in our *ex situ* contactless *C–V* measurements on the hydrogen terminated Si(111) surface[9] and Si(100) surface after BHF cleaning,[19] showing the presence of Fermi-level pinning. Such behavior could actually be explained by the presence of discrete surface states probably originating from the unsaturated Si dangling bonds.[9] In addition, for the present surface, i.e., the BHF- and NH_4F-treated surface, the residual suboxide components may be related to the formation of such surface states.

On exposing the surface to an ECR H_2 plasma, the hysteresis was remarkably reduced, as shown in Fig. 5(b). Such improvement in *C–V* behavior clearly indicates the reduction in the density of surface states, which may be probably due to the removal of residual native oxide components and saturation of Si dangling bonds after H_2 plasma treatment.

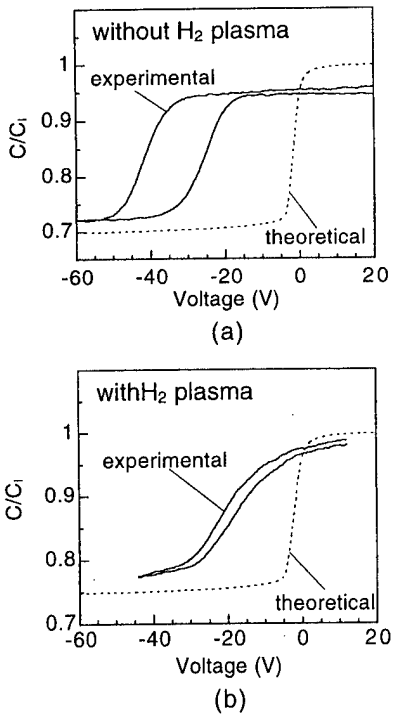


FIG. 5. Results of contactless *C–V* measurements on bare Si(100) surface (a) without any H_2 plasma treatment and (b) with H_2 plasma treatment after the BHF–NH_4F-based wet chemical treatment.

B. Characterization of the ultrathin oxide-Si interfaces after oxynitridation

Figure 6 shows the Si 2*p* core-level XPS spectra of the N_2O plasma oxynidridized Si(100) surface with H_2 plasma treatment. The chemical shift between the oxide bonding and Si–Si bonding was estimated to be 4.0 eV which was the same for surfaces with/without the H_2 plasma treatment. This is typical of Si–O bonds in the SiO_2 phase,[14,20] indicating

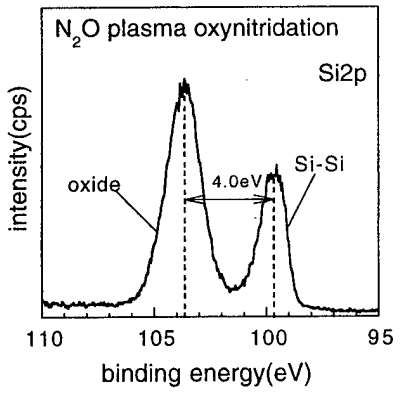


FIG. 6. Si 2*p* core-level spectrum for the N_2O plasma oxynitrided surface with a prior H_2 plasma treatment.

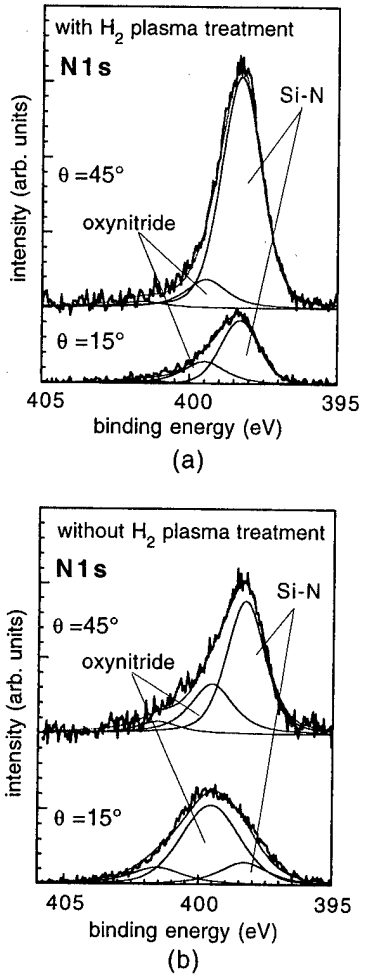


FIG. 7. Angle-resolved XPS spectra of the N 1*s* core levels for N_2O plasma oxynitrided surfaces (a) with and (b) without H_2 plasma treatment.

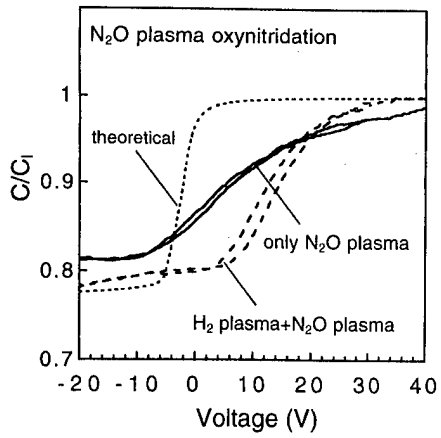


FIG. 8. Contactless $C-V$ curves taken on N_2O plasma oxynitrided surfaces with/without any H_2 plasma treatment.

that the composition of the ultrathin insulator is very close to SiO_2 with a possible inclusion of small amounts of N atoms. The oxide thickness was estimated to be 2.8 nm from the Si 2*p* spectra by assuming that the electron escape depths in Si and Si oxide are 2.6 and 3.4 nm, respectively.[21]

Figure 7 compares the angle-resolved XPS spectra of the N 1*s* core level taken with the photoelectron exit angle θ of 15° and 45° for either surfaces, i.e., with and without the H_2 plasma treatment prior to the oxynitridation process. For both surfaces, it was found that the concentration of the nitrogen atom was 2–3 at. % in the oxide films and the N atoms were concentrated at the interface region. This is consistent with the results at the SiO_2/Si interfaces formed by conventional furnace annealing (FA) process[22] and rapid thermal annealing (RTA) process[23] in N_2O ambient at much higher temperatures (900–1200 °C).

The spectra was deconvoluted into three components using binding energy data with bondings at 398.1, 399.8, and an unidentified bonding at 401.7 eV. The lowest energy peak at 398.1 eV can be assigned to the chemical environment of N similar to that in Si_3N_4,[22,24,25] i.e., one N atom bonded to three Si atoms and may be designated as Si–N in Fig. 7. The second peak has a similar energy position to those found at the interfaces formed by FA and RTA thermal oxidation in N_2O ambient.[22,25] Bhat *et al.*[25] suggested that the peak at 399.8 eV corresponds to the $Si_2{=}N{-}O$ bond, where one N atom is bonded to two Si and one O atom. On the other hand, Hussey *et al.*[22] reported that a similar peak was found at the SiO_2/Si interface formed by FA oxidation using N_2O gas and that it originates from N–Si–O bond where the N atom is partially replaced by the O atom. In fact, the possibility of such bonding was predicted by the analysis of the infrared absorption (IR) spectra obtained from the N_2O-grown oxide.[26]

As shown in Fig. 7(a), the angle-resolved XPS spectra clearly indicate the presence of nitride layer including a small amount of oxynitride phase, being localized at the interface. This indicates that ECR N_2O plasma oxynitridation after H_2 plasma treatment seems to form a well-defined interface configuration of $SiO_2/Si_3N_4/Si$. In comparison, for the surface which was not exposed to ECR H_2 plasma, a large amount of oxynitride components were included at the interface and distributed throughout the insulator film, as shown in Fig. 7(b). Thus, we speculate that complete removal of the native oxides from the surface prior to the oxynitridation process is imperative for formation of a phase separated $SiO_2/Si_3N_4/Si$ structure, otherwise, during oxynitridation they may play a significant role for formation of large amounts of oxynitrided components.

Figure 8 shows the contactless $C-V$ curves taken on the ECR N_2O plasma oxynitrided surfaces with and without H_2 plasma treatment as a part of our initial surface treatment. An enhancement in the $C-V$ variation on treatment of the surface with H_2 plasma prior to its oxynitridation is clearly observed, indicating the formation of a relatively well-behaved interface. The distributions in the density of the interface states (D_{it}), as calculated from the measured $C-V$ curves using the Terman method, are shown in Fig. 9. The interface treated with H_2 plasma gave a wider and lower U-shaped continuous D_{it} distribution with a minimum value of 1.0×10^{11} eV^{-1} cm^{-2}, even in the as-prepared phase with an ultrathin oxide thickness (2.8 nm) formed at low temperature (400 °C). At the ECR N_2O plasma oxynitrided surfaces

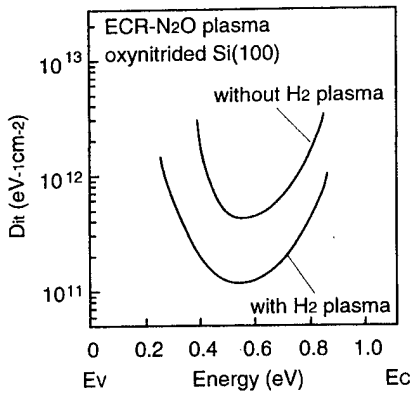


FIG. 9. Interface state density (D_{it}) distribution for the oxynitrided surfaces.

without H_2 plasma treatment, a large amount of oxynitride components was included both at the interface and in the film. Such complicated chemical status is likely to change the chemical environment of Si and O atoms and to leave disordered bonds at the interface, thereby resulting in the higher value of D_{it}. Thus, we think the phase-separated $SiO_2/Si_3N_4/Si$ structure is important to realize a low D_{it} value. In this regard, Lucovsky *et al.*[27] and Lee *et al.*[28] reported that such phase-separated structures can be realized by combined use of a N_2O remote-plasma-assisted oxidation of Si surface and plasma enhanced chemical vapor deposition (PECVD) of thin SiO_2 film and are effective for improvement in device performance and reliability. Lu *et al.*[29] recently reported that such interfacial phase is effective to reduce the interface strain due to a mismatch in a Si atomic density. Such an effect may be one of the possible reasons for improvement in the property of the interface formed by the ECR N_2O plasma process. However, further investigation of the effects of surface treatments as well as process conditions on the interface properties is still of prime necessity.

IV. CONCLUSION

Ultrathin insulator/Si(100) interface could be formed by a direct low-temperature ECR N_2O plasma process under optimized conditions. An UHV-based contactless $C-V$ measurement system allowed for the first time a direct *in situ* assessment of electronic properties of "free" Si surfaces and the surfaces covered with an ultrathin insulator layer. A systematic characterization by direct *in situ* techniques of the electronic and chemical properties of the surface and the interface using the *in situ* contactless $C-V$ and XPS methods, revealed that the Si(100) surface with an additional treatment with H_2 plasma (after the wet chemical treatment) is very attractive for understanding as well as control of ultrathin insulator films on an atomic scale because of its native-oxide-free clean nature, which is well reproducible. Formation of a phase-separated Si_3N_4/SiO_2 structure was found to be important to realize a well-behaved ultrathin insulator/Si interface at low temperatures.

[1]*The National Technology Roadmap for Semiconductors*, 1994 Edition (Semiconductor Industry Association, San Jose, 1994).
[2]H. Hasegawa, Extended Abstract of 1997 International Conference on Solid State Devices and Materials, Special Issue on Single Electron Devices, p. 302.
[3]T. Hori, H. Iwasaki, and K. Tsuji, IEEE Trans. Electron Devices **36**, 340 (1989).
[4]H. Fukuda, M. Yasuda, T. Iwabuchi, and S. Ohno, IEEE Electron Device Lett. **EDL-12**, 587 (1991).
[5]G. W. Yoon, A. B. Joshi, J. Kim, and D. L. Knong, IEEE Electron Device Lett. **EDL-14**, 179 (1993).
[6]H. Hwang, W. Ting, B. Maiti, D. L. Kwong, and J. Lee, Appl. Phys. Lett. **57**, 1010 (1990).
[7]H. Fukuda, T. Arakawa, and S. Ohno, Jpn. J. Appl. Phys., Part 2 **29**, L2333 (1990).
[8]S. Koyanagi, T. Hashizume, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **35**, 630 (1996).
[9]T. Hashizume, S. Koyanagi, and H. Hasegawa, J. Vac. Sci. Technol. B **14**, 2872 (1996).
[10]T. Yoshida, T. Hashizume, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **36**, 1453 (1997).
[11]G. S. Higashi, Y. S. Chabal, G. W. Trucks, and Raghavachari, Appl. Phys. Lett. **56**, 656 (1990).
[12]F. Goos and H. Haechen, Ann. Phys. (Leipzig) **6**, 344 (1947).
[13]T. Sakai, M. Kohno, S. Hirae, I. Nakatani, and T. Kusuda, Jpn. J. Appl. Phys., Part 1 **32**, 4005 (1993).
[14]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[15]C. J. Karlsson, F. Owman, E. Landemark, Y. C. Chao, P. Martensson, and R. I. G. Uhrberg, Phys. Rev. Lett. **27**, 4145 (1994).
[16]K. Yamamoto and M. Hasegawa, J. Vac. Sci. Technol. B **12**, 2493 (1994).
[17]C. H. Bjorkman, J. L. Alay, H. Nishimura, M. Fukuda, T. Yamazaki, and M. Hirose, Appl. Phys. Lett. **67**, 2049 (1995).
[18]T. Aiba, K. Yamauchi, K. Shimizu, N. Tate, M. Katayama, and T. Hattori, Jpn. J. Appl. Phys., Part 1 **34**, 707 (1995).
[19]T. Sakai, M. Kohno, H. Okada, H. Matsubara, and S. Hirae, Jpn. J. Appl. Phys., Part 1 **36**, 360 (1997).
[20]S. Iwata and A. Ishizaki, J. Appl. Phys. **79**, 6653 (1996).
[21]H. Ogawa and T. Hattori, Appl. Phys. Lett. **61**, 577 (1992).
[22]R. J. Hussey, T. L. Hoffman, Y. Tao, and M. J. Graham, J. Electrochem. Soc. **143**, 221 (1996).
[23]E. C. Carr and R. A. Buhrman, Appl. Phys. Lett. **63**, 54 (1993).
[24]J. A. Taylor, Appl. Surf. Sci. **7**, 168 (1981).
[25]M. Bhat, G. W. Yoon, D. L. Kwong, M. Arendt, and J. M. White, Appl. Phys. Lett. **64**, 2116 (1994).
[26]P. Lange, H. Bernt, E. Kartmannsgruber, and F. Naumann, J. Electrochem. Soc. **141**, 259 (1994).
[27]G. Lucovsky, D. R. Lee, S. V. Hattangady, H. Niimi, Z. Jing, C. Parker, and J. R. Hauser, Jpn. J. Appl. Phys., Part 1 **34**, 6827 (1995).
[28]D. R. Lee, G. Lucovsky, M. S. Denker, and C. Magee, J. Vac. Sci. Technol. A **13**, 1671 (1995).
[29]Z. H. Lu, R. J. Hussey, M. J. Graham, R. Cao, and S. P. Tay, J. Vac. Sci. Technol. B **14**, 2882 (1996).

Surface roughness of SiO_2 from a remote microwave plasma enhanced chemical vapor deposition process

M. J. Rack,[a] D. Vasileska, and D. K. Ferry
Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

M. Sidorov
Center for Solid State Science, Arizona State University, Tempe, Arizona 85287-1704

(Received 21 January 1998; accepted 28 May 1998)

We have investigated the roughness of the top surface of silicon dioxide deposited via a remote plasma enhanced chemical vapor deposition (RPECVD) process in a microwave reactor. We find a roughening transition at a deposition temperature of approximately 250 °C. Above this temperature, the surface is fairly smooth (root mean square roughness ~0.3 nm). Below this deposition temperature, the oxide surface becomes extremely rough. Rapid thermal annealing at 900 °C does not eliminate this roughness, which is very nonuniform at the nanometer scale. For very thin RPECVD oxide applications, oxide surface roughness could be a limitation. We have used our three-dimensional Poisson solver in order to investigate the effects of oxide surface roughness taken from actual atomic force microscopy measurements on the confining potential within the silicon inversion layer of a metal–oxide–semiconductor (MOS) field effect transistor. In order to assess the quality of our process and system, oxides are characterized electrically with MOS capacitors, and structurally with Fourier transform infrared spectroscopy, high-resolution cross-sectional transmission electron microscopy, and etch rates in HF containing solutions. © *1998 American Vacuum Society.* [S0734-211X(98)11804-8]

I. INTRODUCTION

Low temperature deposited oxide plays a critical role in novel nanostructure devices where either low thermal budget or the requirement to cover preexisting structures requires its use.[1,2] In some cases, it must have sufficient integrity to function as a gate oxide in silicon metal–oxide–semiconductor field effect transistor (MOSFET)-like structures. It has been demonstrated that remote plasma enhanced chemical vapor deposition (RPECVD) oxide is suitable for this task, particularly if a 900 °C rapid thermal anneal (RTA) can be tolerated.[3] With proper surface preparation, oxide deposited at 200–225 °C has demonstrated excellent electrical characteristics even without the high temperature anneal.[4,5] The advantage of this deposition temperature is that very little thermal oxidation at the silicon surface occurs so that the Si/SiO_2 interface remains at the original silicon surface, which can be initially 2×1 reconstructed in ultrahigh vacuum deposition systems. For the same reason, this temperature is attractive when depositing oxide on top of tiny metal gate structures where unintentional oxidation of the metal can effectively eliminate the structure.[6] The deposition temperature selected must be compatible with the previously fabricated device.

Various researchers have reported abrupt changes in etch rates of both direct and remote PECVD oxides in HF containing solutions occurring at approximately 180, 225, and 285 °C that were correlated with refractive index changes, and loosely with OH impurity content.[7–9] We have observed that oxides deposited below 270 °C tend to be less reliable than the oxides deposited above this temperature. In order to gain insight into the bulk oxide differences, we characterized our RPECVD oxides with etch rates, searching for a possible transition temperature that might coincide with these differences. In addition, we studied the top surface roughness with atomic force microscopy (AFM). The top surface may be expected to provide a glimpse into the bulk structure. The oxide deposited at 225 °C was found to be significantly rougher than that deposited at 350 °C, as shown in Fig. 1, even after a 900 °C, 1 min RTA.

In this study, we investigate the roughening of the RPECVD oxide as a function of deposition temperature and thickness. We also examine the effects of high temperature annealing. Finally, we assess the impact of this roughness on the MOSFET Si/SiO_2 interface for both inversion and depletion.

II. EXPERIMENT

RPECVD SiO_2 was deposited in a remote microwave system. Two processes were employed. Process I used 150 W, 750 mTorr, a helium flow of 1000 sccm, N_2O of 70 sccm, and dilute silane of 10 sccm at 5% silane in helium. Following various factorial experiments, process II was implemented with power increased to 250 W and N_2O to 75 sccm. There was no discernable change in the roughness between these two processes, but the etch rates were reduced slightly. Both *p*- and *n*-type Si(100) wafers were used. Wafers were cleaned initially with H_2O_2:H_2SO_4 (3:7), carefully rinsed, etched in buffered HF, and rinsed again in de-ionized water. A 120 nm sacrificial oxide was then grown at 1050 °C. Prior to deposition, these wafers were etched in 50:1 H_2O:HF for

[a]Electronic mail: jo.rack@asu.edu

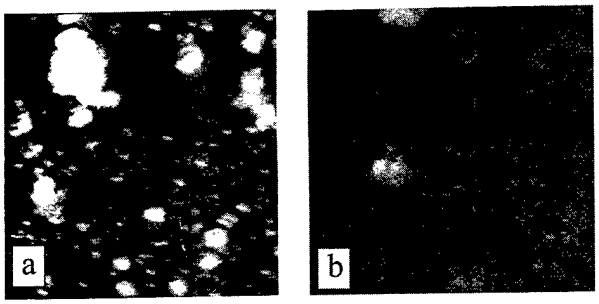


FIG. 1. AFM surface plots of 150 nm thick RPECVD SiO_2 deposited with process I at (a) 225 and (b) 350 °C following a 900 °C/1 min anneal in N_2. The grayscale range is 0–30 nm in height.

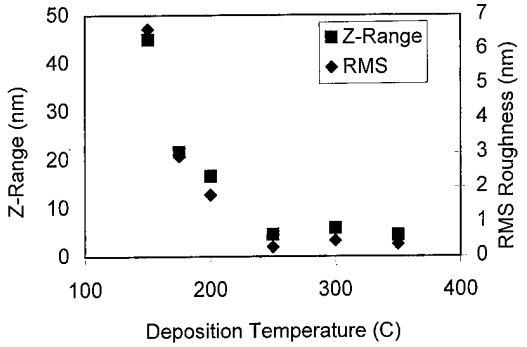


FIG. 2. rms (diamonds) and Z-range (squares) height deviations for 120 nm RPECVD SiO_2 deposited using process II at various temperatures followed by a 450 °C/15 min anneal in forming gas.

5–10 min beyond becoming hydrophobic. A 30 s exposure to plasma excited N_2O preceded the deposition. The rate of deposition was 2.8–3.0 nm/min. High temperature annealing for selected samples was done in a Tamarack model 180-M rapid thermal processor (RTA). Etch rate ratios (ERRs) were measured by comparing the RPECVD oxide etch rate to that of dry thermal SiO_2 grown at 1050 °C in both 20:1 buffered HF (BHF) and 100:1 H_2O:HF at 49%. The oxide thickness was measured using a Woolam M-44 spectroscopic ellipsometer. The noise in this measurement was reduced by frequently calibrating the bath with the thermal oxide etch rate, particularly for the 100:1 etch. The infrared Si–O stretching peak position and width were measured with a Nicolet 740 SX Fourier transform infrared (FTIR) spectroscopy system using 4 cm^{-1} resolution and a mercury cadmium telluride (MCT) detector.

AFM measurements were done on a DI scanning probe microscope. We used both etched silicon probes (2–10 nm tip radius) and silicon nitride tips (20–60 nm tip radius), and both contact and tapping mode. Data points are the mean of 5–25 scans with each scan containing 512×512 data points. Images are flattened on a line by line basis, and root mean square (rms) height deviations are then calculated. The Z-range values are the maximum height excursions (max–min) for the image after flattening. These points were also averaged over many images. Because roughness tends to scale with measured length, only 500 nm×500 nm areas were used. Electrical measurements were done with aluminum gates deposited using a shadow mask. All samples were given a postmetalization anneal (PMA), either at 420 °C for 20 min in nitrogen or, at 450 °C for 15 min in forming gas.

III. ROUGHNESS VERSUS DEPOSITION TEMPERATURE

A very strong dependence of surface roughness upon deposition temperature was found with a sharp transition occurring at about 250 °C. Figure 2 shows both the rms roughness and the average height range for a series of deposition temperatures. All samples were annealed at 450 °C for 15 min in forming gas after the deposition at the specified temperature. This series of wafers were *n*-type (100) double side polished 1–10 Ω cm which had a 120 nm sacrificial oxide removed prior to deposition. The roughness of these surfaces prior to deposition was below 0.15 nm rms. This temperature dependence is also found in the measured etch rate ratios shown in Fig. 3, which compares the RPECVD oxide with a 1050 °C thermal oxide. A transition in etch rate ratios is also observed near 250 °C. The ERRs in 100:1 H_2O:HF at 49% show greater sensitivity to the RPECVD oxide defects than the corresponding measurements in 20:1 BHF which are more frequently reported. The general behavior, though, is the same.

The FTIR results for these samples after the 450 °C anneal showed the Si–O stretching absorption peak to be at 1055±1 cm^{-1} for the 150–300 °C deposited samples and 1058 cm^{-1} for the 350 °C deposited sample. The full width at half maximum (FWHM) of this peak was 90±2 cm^{-1}. This may be compared to a 120 nm thick thermal oxide grown at 1050 °C having a peak at 1080 cm^{-1} and a FWHM of 75 cm^{-1}. There was no peak at 2250 or 870 cm^{-1} which would have indicated the presence of SiH, which has been associated with excess silicon in the oxide film.[10] The 450 °C anneal eliminates SiOH which is not detectable in the FTIR spectra. However, the refractive index, shown in Fig. 4, increases with deposition temperature, peaking near 300 °C.

If these samples are stoichiometric SiO_2, then the refractive index can be an indication of density. The fixed position of the Si–O stretching frequency at 1056 cm^{-1} would indicate that the SiO_2 tetrahedra were similarly constricted relative to relaxed high temperature thermal SiO_2 but, on a more

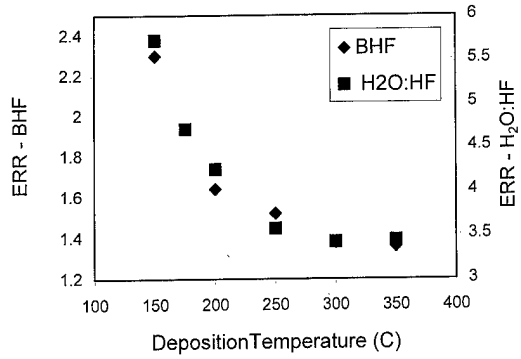


FIG. 3. Ratios of the etch rates of RPECVD oxide to 1050 °C thermally grown oxide in 10:1 BHF (diamonds) and 100:1 H_2O:HF (squares).

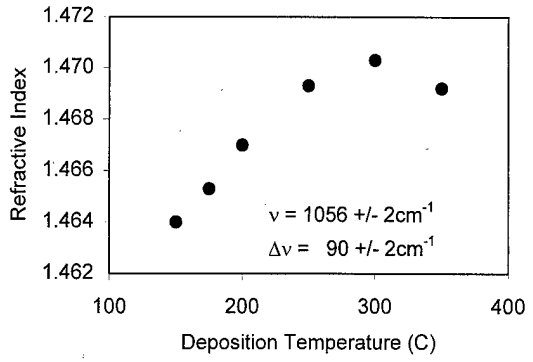


FIG. 4. Refractive index plotted vs RPECVD oxide deposition temperature. Oxides were deposited using process I at the given deposition temperature followed by a 450 °C/1 min anneal in forming gas.

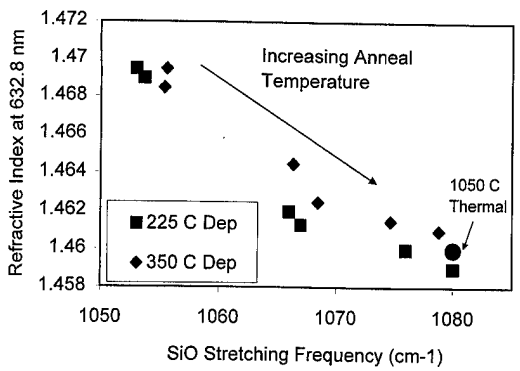


FIG. 5. Refractive index plotted vs Si–O stretching frequency for RPECVD SiO_2 deposited with process I. Included are the as deposited sample and those annealed at 420 °C for 20 min, at 700, 800, and 900 °C for 1 min, and at 1100 °C for 2 min. A thermal oxide reference is included.

macroscopic basis, the oxides deposited at lower temperatures were less dense. Roughness at the surface provides a glimpse into the bulk structure. This suggests that the reduced density of the lower temperature films might, in part, result from the incorporation of nanoscale clusters, whose occurrence is strongly correlated to deposition temperature. These clusters might create voids from not quite fitting together. For remote processes, the temperature range from 150 to 200 °C is characterized by decreasing incorporation of SiOH, which would also open the SiO_2 amorphous network structure.[4]

The MOS capacitors fabricated from oxide deposited at 225 and 350 °C show very little differences. Following a PMA, the fixed charge measured from the 225 °C sample was 1×10^{11} cm^{-2}, and from the 350 °C sample it was 9×10^{10} cm^{-2}. The difference most likely results from surface conditions prior to deposition. While the average breakdown field strength was about 8.5 MV/cm, lower temperature oxides suffered more premature breakdowns.

IV. EFFECTS OF HIGH TEMPERATURE ANNEALING

The effect of high temperature annealing in nitrogen on the RPECVD oxide surface roughness for deposition temperatures of 225 and 350 °C was studied. The samples in this series were annealed in nitrogen for 20 min at 420 °C prior to 1 min high temperature rapid thermal anneals in nitrogen. As shown Table I, there is apparently some reduction in the rms roughness between 400 and 700 °C but the variance in the 225 °C group is too large to conclude a trend has occurred. The majority of the difference between the relatively smooth oxide deposited at 350 °C and the much rougher one deposited at 225 °C is not eliminated by these 1 min anneals even at 900 °C. The IR Si–O stretching peak position is plotted against the refractive index in Fig. 5. There is very little difference between the 225 and the 350 °C deposited samples in this plot, except that the 225 °C sample appears to have a lower refractive index for similar stretching frequency peak position at corresponding annealing temperatures. Figure 5 includes the as deposited points as well as an 1100 °C/2 min anneal in nitrogen for the 225 and the 350 °C deposited oxides. If both oxides are stoichiometric SiO_2, this would indicate that the 225 °C deposited oxide might be less dense than the 350 °C deposited oxide. The peak position moves up with annealing temperature for both indicating a progressively increasing Si–O–Si bond angle[4,11] and a more relaxed film. With higher annealing temperature, the refractive index monotonically decreases becoming indistinguishable from thermal oxide only after the 2 min, 1100 °C anneals in nitrogen. The FWHM peak width (not shown) also decreased monotonically with anneal temperature from 95 to 75 cm^{-1} for both the 225 and the 350 °C deposited samples. There was no significant difference between the 225 and the 350 °C deposited linewidths at any of the annealing temperatures. The etch rate ratios are reduced by high temperature annealing also. Table II gives these results.

TABLE I. Roughness vs annealing temperature.

Anneal temperature (°C)	Deposition temperature 225 °C rms (nm)	225 °C Z-range (nm)	350 °C rms (nm)	350 °C Z-range (nm)
As deposited	1.61	11.7	0.40	4.19
420	2.09	14.9	0.41	4.45
700	1.44	10.7	0.32	4.1
800	1.92		0.43	5.1
900	1.38	13.6	0.41	4.51

TABLE II. Etch rate ratios vs annealing temperature.

Anneal temperature (°C)	Deposition temperature 225 °C ERR BHF	225 °C ERR 100:1	350 °C ERR BHF	350 °C ERR 100:1
420	1.72	4.53	1.51	3.62
700	1.32	2.65	1.25	2.06
800	1.25	2.16	1.15	1.82
900	1.13	1.39	1.06	1.33

The effect of high temperature annealing at 900 °C in nitrogen on MOS capacitors made from 50 nm thick RPECVD oxide deposited at 270 °C was to reduce the fixed and interface charges to below 2×10^{10} cm^{-2}. It had little effect upon mobile charge which remained at $3-5\times10^{10}$ cm^{-2}. In general, high temperature annealing in nitrogen did not improve the reliability of oxides grown at lower temperature. However, if nitrogen and 5% oxygen were used during the anneal, the capacitors showed significant improvement in reliability.

V. EVOLUTION OF ROUGHNESS WITH OXIDE THICKNESS

In order to understand how the surface evolves with growth, 5, 50, and 150 nm thick oxide surfaces from 225 °C depositions were studied. At 5 and 50 nm, the surfaces may be described as being rougher than thermal oxides but still fairly smooth, punctuated by bumps whose frequency increases with increasing oxide thickness. As the oxide becomes thicker, the background roughness increases as well. The height of these bumps was usually between 0 and 10 nm. The width, as seen from AFM, ranged from 15 to 150 nm. Only one larger bump was observed in over 100 scans on any of the wafers studied. This precludes optical detection, except as an intermediate thin film with a dielectric response between SiO_2 and air.[12] The minimum bump width detectable with AFM will be limited by the tip sharpness. Bumps that are more abrupt than the tip sharpness will tend to image the tip. We found that sharper tips tended to increase the measured rms roughness of the rougher surface (225 °C) and reduce the rms roughness of the smoother surface (350 °C). Therefore, the estimates of the rougher surfaces may be considered to be conservative. A 315 nm wide cross-sectional transmission electron micrograph of the 150 nm thick RPECVD oxide grown at 225 and 350 °C showed bump heights of 3–6 nm and widths of 7–64 nm, which is very much within the range of bumps imaged with AFM with the low length cut off by finite tip resolution. The Z-range values measured in these cross-sectional transmission electron microscopy (XTEM) images closely matched those of the AFM measurements, AFM images were only accepted if a wide variety of feature sizes appeared. Because the bumps were of similar height, it is clear that they represent variations in the film thickness that are large for thin oxides, and become less important for thick oxides. As they randomly accumulate, the range of image heights increases.

VI. NUMERICAL MODELING

We assessed the impact of these surface fluctuations on the potential energy at the $Si-SiO_2$ interface of a MOSFET under both inversion and depletion, using Poisson simulations with the AFM data taken from the 5 and 50 nm thick oxides.

A. Theoretical model

A three-dimensional (3D) plot of the top surface of the MOS structure under investigation is shown in Fig. 6(a). In

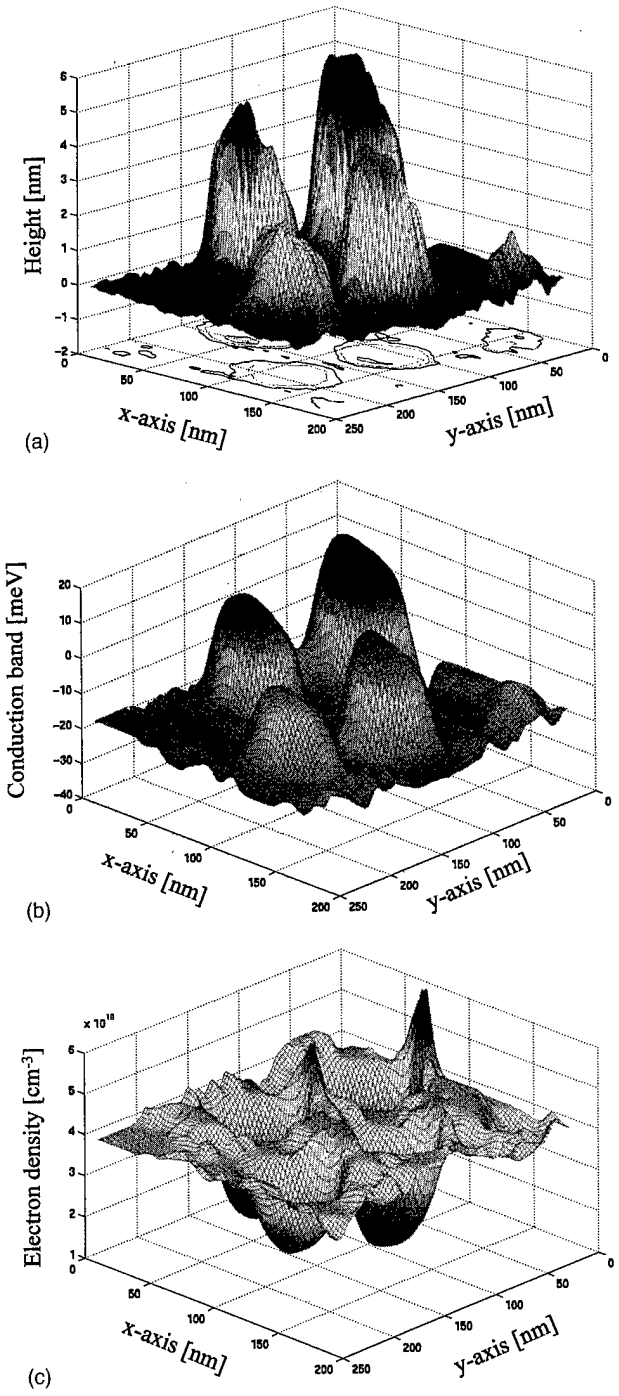


FIG. 6. (a) AFM topograph of a 5 nm RPECVD 200 nm×200 nm surface deposited at 225 °C. The height scale is 6 nm so the peaks are exaggerated. (b) The resulting simulated conduction band energy surface in a MOSFET inversion layer showing fluctuations that mirror the oxide thickness variations. (c) Inversion layer electron density fluctuations from simulation that result from surface roughness.

Fig. 6(a), the $\bar{0}$ reference height corresponds to the top surface of a 5 nm oxide that serves as the gate oxide. The bumps, which are exaggerated by the scale factor, are taken from the actual AFM scans of a 5 nm oxide deposited at 225 °C. Their presence actually doubles the oxide thickness in small areas. The substrate doping density is N_a

$=10^{16}$ cm^{-3}. The conduction band profile and the electron density at the semiconductor/oxide interface is obtained from the solution of the 3D Poisson equation

$$\nabla(\epsilon\nabla\varphi)=-q(C-n+p), \tag{1}$$

where φ is the electrostatic potential, ϵ is the spatially dependent dielectric constant, $C=N_d-N_a$ is the net ionized impurity concentration, and n and p are the electron and hole concentrations for which we use the general Fermi–Dirac statistics. This allows us to consider MOS structures with high substrate doping. The Fermi potential is taken to be the zero-reference potential. Under this condition, the electron and hole concentrations are evaluated from

$$n=N_C F_{1/2}[(q\varphi-\Delta E_C)/k_B T] \tag{2a}$$

and

$$p=N_V F_{1/2}[(\Delta E_C-E_g-q\varphi)/k_B T], \tag{2b}$$

respectively. In Eqs. (2), N_C and N_V are the effective densities of states of the conduction and valence bands, $F_{1/2}(x)$ is the Fermi–Dirac integral of order 1/2, E_g is the silicon band gap, k_B is the Boltzmann constant, T is the temperature, and ΔE_c, which describes the position of the conduction band with respect to the intrinsic Fermi level, is obtained as a solution of the implicit equation

$$N_C F_{1/2}(-\Delta E_C/k_B T)=N_V F_{1/2}[(\Delta E_C-E_g)/k_B T]. \tag{3}$$

Since the Fermi–Dirac integral $F_{1/2}(x)$ does not have a closed-form solution, for its calculation we use the analytic approximation of Bednarczyk and Bednarczyk[13]

$$F_{1/2}(x)=(e^{-x}+3\sqrt{\pi}/4\nu^{3/8})^{-1}, \tag{4a}$$

where

$$\nu(x)=x^4+50+33.6x\{1-0.68\exp[-0.17(x+1)^2]\}. \tag{4b}$$

B. Numerical details

It is well known that the application of a finite-difference technique to the 3D Poisson equation leads to algebraic equations having a well-defined structure, represented by a matrix equation $\mathbf{Ax}=\mathbf{b}$. We used Stone's strongly implicit procedure[14] for the numerical solution of this matrix equation to estimate the influence that the fluctuations in the oxide thickness have on the electron density in the channel. The details of this are given elsewhere.[15] The nonuniform finite-difference tensor-product mesh used in these simulations has $120\times90\times65$ grid points along the x axis (length), y axis (width), and z axis (height). This discretization scheme allowed us to resolve fluctuations in the top surface larger than 0.3 nm. The predetermined threshold for self-consistency for the electrostatic potential is 10^{-5} thermal voltages.

C. Simulation results

In Figs. 6(b) and 6(c) we present the simulation results for the conduction band profile E_C and the corresponding electron density at the Si/SiO_2 interface under inversion conditions (applied gate bias $V_G=1$ V on the metal gates). It is obvious that the roughness of the top interface affects the magnitude of the conduction band edge E_C which, in turn, leads to significant fluctuations of the electron density in the channel because of the exponential dependence of n on the electrostatic potential φ, and therefore E_C ($E_C=\Delta E_C-q\varphi$). This observation led to a concern for possible problems when trying to deplete the carriers underneath the gates. However, our subsequent numerical calculations showed that a gate bias of -0.8 to -1 V is sufficient for complete depletion and that the roughness of the top interface does not pose any problems. The situation was even better for the sample with 50 nm gate oxide for both inversion and depletion conditions. For example, for the MOS structure with 5 nm oxide thickness and applied bias of 1 V, for which the mean value of the roughness is $\bar{h}=0.5$ nm and the standard deviation of the surface fluctuations is $\sigma_h=1.33$ nm, we calculated that the mean conduction band offset equals to $\overline{E_C}=-15.3$ meV and the standard deviation is $\sigma_{Ec}=10.3$ meV. The corresponding mean electron density was $\bar{n}=3.4\times10^{19}$ cm^{-3}, whereas $\sigma_n=0.83\times10^{19}$ cm^{-3}. For the MOS structure with 50 nm grown oxide and applied bias of 12 V, for which $\bar{h}=0.8$ nm and $\sigma_h=2.9$ nm, we found that $\overline{E_C}=-43.61$ meV, $\sigma_{Ec}=2.242$ meV, $\bar{n}=6.7\times10^{19}$ cm^{-3}, and $\sigma_n=0.32\times10^{19}$ cm^{-3}. It is obvious that, even though the sample with thicker oxide had a much rougher surface and larger gate bias that further deteriorates the statistics, the fluctuations in the conduction band edge and the corresponding electron density for this sample were significantly smeared out.

VII. CONCLUSIONS

We have investigated the outer surface roughness of RPECVD oxides that, in terms of other bulk properties, compare very well with others previously reported. We find that at temperatures below about 250 °C, the surface becomes rough and gets rougher as the temperature is further reduced. The evolution of the roughness suggests a gradual incorporation of clusters that accumulate more quickly at low temperature. The top surface provides a glimpse into the bulk structure, suggesting that the excess porosity attributed to PECVD oxides in general may, in part, result from these clusters in the bulk creating voids for lack of fitting together.

The most obvious explanation for this is gas phase nucleation producing particulate SiO_2 that are of very small dimension (tens of nanometers). However, the dependence upon wafer temperature and the constancy of the roughness regardless of the processing parameters suggest that it might be a temperature-controlled surface or very near surface clustering phenomenon.

The effect of roughness on a silicon MOSFET was numerically simulated using the AFM data taken from 5 and 50 nm RPECVD oxide grown at 225 °C. These simulations showed that it created significant fluctuations in the conduction band potential energy when very thin (5 nm) oxides were used in inversion. The interface could still be depleted,

however, if slightly higher voltage was applied. For thicker oxides, no serious disruptions occur.

[1]J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62**, 583 (1989).
[2]A. H. Kamal, J. Lutzen, B. A. Sanborn, M. V. Siderov, M. N. Kozicki, D. J. Smith, and D. K. Ferry, Semicond. Sci. Technol. (submitted).
[3]G. Lucovsky, H. Niimi, K. Koh, D. R. Lee, and Z. Jing, Proc. Electrochem. Soc. **96-1**, 441 (1996).
[4]G. Lucovsky, S. S. Kim, and J. T. Fitch, J. Vac. Sci. Technol. B **8**, 822 (1990).
[5]B. K. Ip, K. C. Kao, and D. J. Thomson, Solid-State Electron. **34**, 123 (1991).
[6]M. J. Rack, A. D. Gunther, M. Khoury, D. Vasileska, M. Sidorov, and D. K. Ferry, Semicond. Sci. Technol. (to be published).
[7]M. F. Ceiler, Jr., P. A. Kohl, and S. A. Bidstrup, J. Electrochem. Soc. **142**, 2067 (1995).
[8]R. L. Jackson, J. E. Spencer, J. L. McGuire, and A. M. Hoff, Solid State Technol. **30**, 107 (1987).
[9]J. D. Chapple-Sokol, W. A. Pliskin, R. A. Conti, E. Tierney, and J. Batey, J. Electrochem. Soc. **138**, 3723 (1991).
[10]G. Lucovsky and D. V. Tsu, J. Cryst. Growth **86**, 804 (1988).
[11]G. Lucovsky, J. T. Fitch, E. Kobeda, and E. A. Irene, *Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface* (Plenum, New York, 1988), p. 139.
[12]D. E. Aspnes, Phys. Rev. B **41**, 10 334 (1990).
[13]D. Bednarczyk and J. Bednarczyk, Phys. Lett. **64A**, 409 (1978).
[14]H. L. Stone, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **5**, 530 (1968).
[15]D. Vasileska, W. J. Gross, V. Kafedziski, and D. K. Ferry, VLSI Design (in press).

Investigation of postoxidation thermal treatments of Si/SiO_2 interface in relationship to the kinetics of amorphous Si suboxide decomposition

B. J. Hinds,[a] F. Wang, D. M. Wolfe, C. L. Hinkle, and G. Lucovsky
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 21 January 1998; accepted 21 May 1998)

Interfacial Si suboxides (SiO_x, $x<2$) are detrimental to transistor performance and are typically minimized during postoxidation anneals. To study the kinetics of SiO_x decomposition, thick films (~2000 Å) of amorphous Si:O:H alloys ($0.7<x<1.4$) were deposited by remote plasma enhanced chemical vapor deposition and subjected to rapid thermal anneals. Films were characterized by Fourier transform infrared spectroscopy, Raman spectroscopy, ellipsometry, photoluminescence (PL), and transmission electron microscopy. At temperatures >500 °C initially there is a rapid segregation into amorphous Si (a-Si) surrounded by a SiO_2 shell which acts as a diffusion barrier decelerating the reaction. Phenomenological modeling of kinetics with a one-dimensional Avrami–Erofe've treatment gives an upper limit for a-Si lateral growth rates of 1.2 Å/s at 900 °C with an activation energy of 120 kJ/mol. PL, Raman, transmission electron microscopy and ellipsometry confirm this segregation model in the amorphous state. Due to the rapid initial decomposition and relatively large diffusion coefficients, a simple kinetic hindrance explanation for the 4–8 Å of SiO_x at the SiO_2/Si interface is unlikely. © *1998 American Vacuum Society.* [S0734-211X(98)09104-5]

I. INTRODUCTION

Due to its dominating effects in metal–oxide–semiconductor field-effect transistor (MOSFET) performance, the Si/SiO_2 interface is the subject of extensive studies. This interface is well known to be stoichiometric SiO_2 within 4–6 Å of the interface where suboxides, SiO_x ($x<2$) are seen.[1,2] The amount and nature of the suboxides at the interface is highly dependent on surface preparation, annealing temperatures, and growth temperatures.[1] Solid $SiO_{1.0}$ is thermodynamically unstable below 1173 °C,[3] and decomposes as in Eq. (1). Typically, suboxides in a MOSFET are minimized during a postoxidation anneal. However the quantification of suboxide by x-ray photoelectron spectroscopy (XPS) after this process is extremely difficult due to attenuation of signal from buried interface. If ultrathin oxides are used, the volatilization of SiO gas during anneals becomes significant. It is thus of interest to study the kinetics of Si suboxide disproportionation in bulk samples which can be related to the interface:

$$SiO_x \rightarrow \frac{x}{2} SiO_2 + \left(1 - \frac{x}{2}\right) Si. \quad (1)$$

Solid $SiO_{1.0}$ grown from evaporation is highly pyrophoric and difficult to study, while H termination stabilized Si suboxides can be synthesized by chemical vapor deposition of semi-insulating polysilicon (SIPOS) or by remote plasma enhanced chemical vapor deposition (RPECVD). Early reports show that SIPOS films decompose into Si_{nc} and amorphous SiO_2 matrix[4] as expected in Eq. (1) and have been studied further,[5–8] however none of these studies closely examined the temperature region, where ultrathin MOSFET gate dielectric growth is important (700–900 °C). There is one report of SiO_x decomposition kinetics in SIPOS films which is derived from measurement of Si nanocrystal size.[9] However, that study measured a crystallization process and is insensitive to reactions in the amorphous state. RPECVD derived films of SiO_x have been shown to similarly form Si_{nc} in SiO_2 matrix after anneals of 900 °C.[10] Therefore it is of interest to study the SiO_x decomposition kinetics with a method insensitive to crystallization processes at temperatures less than 1000 °C.

II. EXPERIMENT

A RPECVD reactor previously described[11] was employed to grow thin films SiO_x ($0.7<x<1.4$). 100 sccm He with 0.05 to 0.2% O_2 was passed through the rf plasma coil tube with 33 W power at 13.56 MHz. 10 sccm of 10% SiH_4/He was introduced through a showerhead ring downstream from the plasma source. Total pressure was maintained at 300 mTorr as well as a substrate temperature of 160 °C. Film thicknesses were typically 2000 Å (16 Å/min). For H_2 plasma treatments, a plasma was generated with 100 sccm of He passing through the plasma coil (50 W power) and 30 sccm of H_2 was introduced downstream with a total pressure of 100 mTorr. The sample was heated to 300 °C and treated for 2 h.

The suboxide films were then annealed *ex situ* in an AG Associates rapid thermal annealing (RTA) furnace under 1 atm of Ar for times of 5 s–3 min (ramp rate 60 °C/s). Fourier transform infrared (FTIR) measurements were performed with a Nicolet 750 spectrometer in transmission mode with a resolution of 4 cm^{-1}. Spectra were subtracted from a reference sample cut from the same Si(100) single side chemically roughened wafer to prevent internal reflection errors. To assure low experimental scatter (from film to film variation), films for each series were annealed sequentially, unless stated otherwise. Photoluminescence (PL) data measurements utilized a capacitance–voltage integrated (CVI) model

[a]Electronic mail: bruce_hinds@ncsu.edu

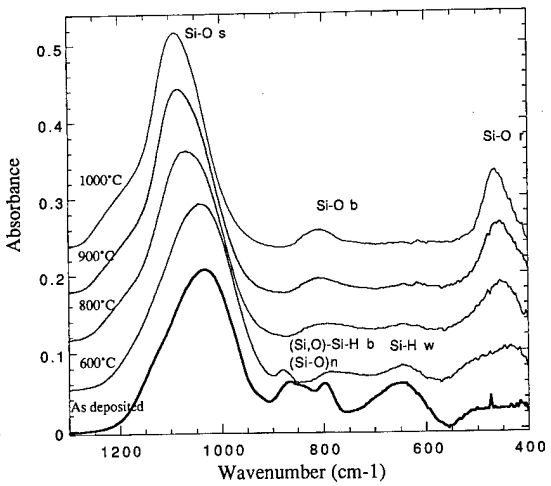


FIG. 1. FTIR absorbance for $SiO_{1.0}$ film after various sequential anneals for 6 min at each temperature, showing blue shift in ν_{Si-O} as well as loss of Si–H bonds. Spectra are offset by 0.06 abs units for clarity. Integrated area under Si–O stretch remains constant for all anneals.

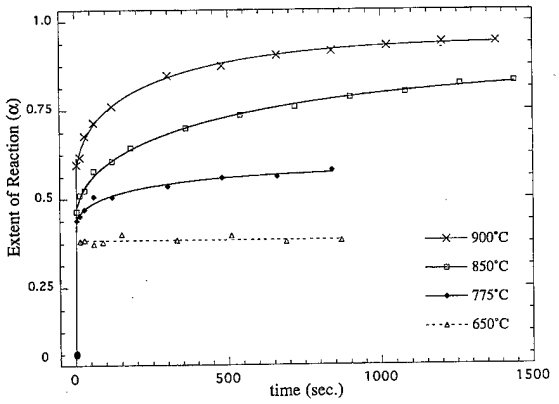


FIG. 2. Extent of reaction (α) for initial $SiO_{0.93}$ thin film as a function of time at various annealing temperatures, showing kinetic metastability of Si suboxides. Curve shown is guide to eye. For each temperature the same film is annealed sequentially to reduce experimental scatter.

DK480 single monochromator with a R928 photomultiplier. A Spectra Physics Ar^+ laser at 488 nm (2.54 eV) provided 100 mW incident excitation energy. Spectral response was calibrated to a standard tungsten–halogen lamp. Conductivity measurements were made at 100 V (resolution to 0.1 pA) between parallel evaporated Al strips with 0.7 mm gap.

III. RESULTS AND DISCUSSION

A. General mechanism of disproportionation

FTIR can be effectively used to monitor the degree of oxygen content (x) in homogeneous SiO_x films, as the Si–O stretch frequency increases linearly from ~965 to ~1065 cm^{-1} with oxygen content[12] as in Eq. (2). All reported stoichiometries for as-deposited films are calculated using Eq. (2) with observed ν_{Si-O}:

$$x = 0.020(\nu_{Si-O}) - 19.3. \quad (2)$$

Figure 1 shows the IR spectra and assigned vibrations for an as-deposited film and the effects of annealing treatments. The most important features due to annealing are the loss of Si–H modes, the increase in ν_{Si-O}, and increase in intensity of Si–O bend and rock modes. The gradual increase in ν_{Si-O} indicates that Eq. (3) describes the reaction chemistry, that is SiO_x disproportionates into a more oxygen rich suboxide $(x+y)$ and Si. The single Si–O stretch peak shape would be expected for either a homogeneous $SiO_{(x+y)}$ alloy or one with concentration gradients. Two distinct Si–O stretch peaks would form if only SiO_2+Si were formed, which is not observed:

$$SiO_x \rightarrow [x/(x+y)]SiO_{(x+y)} + [1-x/(x+y)]Si, \quad (3)$$

$$\alpha = (\nu_m - \nu_i)/(\nu_f - \nu_i) \approx y/(2-x). \quad (4)$$

To quantify the kinetics of reaction (3), the extent of reaction (α) is given in Eq. (4), where reaction completion ($\alpha = 1$) is when $x+y=2$. ν_m is the measured ν_{Si-O}, ν_i is the as-deposited ν_{Si-O}, and ν_f is the measured ν_{Si-O} at completion (after anneals at 1000 °C with $\nu_{Si-O} \approx 1075$ cm^{-1}). The integrated area of the Si–O stretch peak remains constant indicating no significant parasitic oxidation. With successive anneals the Si–O stretch peak becomes sharper and higher in frequency as expected in forming SiO_2.

Figure 2 shows the extent of reaction (α) of initial $SiO_{0.93}$ films as a function of time at various annealing temperatures. At all temperatures there is a rapid (<5 s resolution of RTA technique) initial reaction followed by a rate which is highly deceleratory. On this time scale, the reaction proceeds asymptotically to an extent of reaction less than completion. Si suboxide stability can be found through the kinetic hindrance of the disproportionation reaction. We propose a mechanism which is consistent with the observed kinetics. That is, an initial rapid segregation of SiO_x into a-Si which would by necessity form an SiO_2 shell. This shell acts as a diffusion barrier for continued reaction and explains the observed deceleratory behavior. The higher initial α with increasing anneal temperature is likely the result of a combination of increased nucleation density and increased growth rates.

To examine the temperature dependence of the kinetics, SiO_x films of various composition were annealed at successively higher temperatures. Each point of Fig. 3 is for an anneal of 6 min (where the reaction has decelerated appreciably) at the given temperature. For all compositions, the reaction is less hindered with increasing temperature. This suggests that higher temperatures allow for higher diffusion coefficients to negate the deceleratory diffusion barrier. The reaction kinetics are similar for all initial compositions. The notable exception is that the extent of initial reaction (anneal at lowest temperature, 600 °C) is greater for samples with lower oxygen content. This is consistent with an initial segregation into Si regions with the SiO_2 shell. By a simple stoichiometric argument, less oxygen content will by necessity require more SiO_x to decompose to form an SiO_2 shell at a given nucleation density. Figure 4 shows a cross-sectional transmission electron microscope (TEM) micrograph of $SiO_{1.0}$ after an anneal at 900 °C. Si nanocrystals are formed

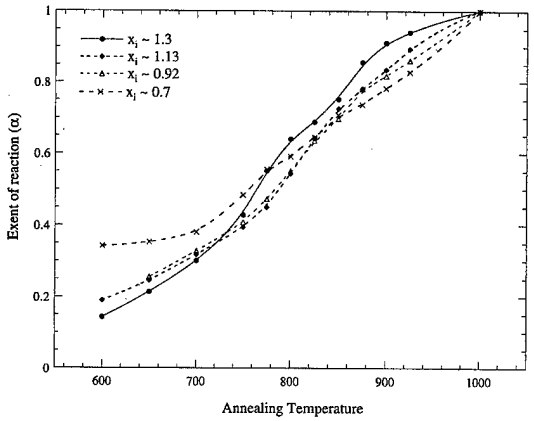


FIG. 3. Extent of reaction (α) for SiO_x (initial $0.7<x_i<1.3$) thin films as a function of annealing temperature showing nearly linear progression of α. Each composition is annealed sequentially on the same film with annealing times of 6 min for each temperature.

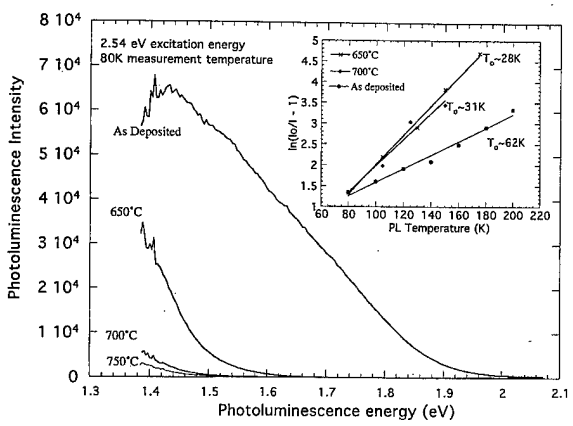


FIG. 5. Photoluminescence of $SiO_{0.93}$ after sequential annealing at various temperatures for 6 min, followed by low temperature H_2 plasma treatment. No PL is observed in annealed samples without H_2 plasma treatment. Inset shows $\ln(I_0/I-1)$ vs T. Annealed samples (1/slope) give T_0 of 28–31 K which is near the characteristic of T_0 of a-Si (25 K).

in an amorphous SiO_2 matrix, consistent with a mechanism of Si segregation with the SiO_2 shell. TEM micrographs of films annealed at 750 °C are completely amorphous; thus the disproportionation reaction proceeds in the amorphous state, not requiring a crystalline growth interface.

SiO_x and amorphous silicon (a-Si) have well known PL properties and can be readily applied to this system to demonstrate the phase segregation in the amorphous state. Figure 5 shows the characteristic visible PL of as-deposited $SiO_{1.3}$. With modest annealing $>\sim$500 °C all PL is lost, due to nonradiative carrier recombination with Si dangling bond defects. Only after treatment of annealed films in H_2 plasma, to passivate Si dangling bonds, is PL restored, similar to results with a-Si after moderate anneals.[13] This PL coincides with the characteristic energy of a-Si (1.3 eV). The inset of Fig. 5 shows the dependence of PL on measurement temperature. The treatment of $\ln(I_0/I-1)$ vs T is commonly applied to amorphous Si with the slope being a characteristic temperature (T_0) for carriers to hop from localized states in band tails to conduction band.[14] With higher T_0, PL is more stable to increasing measurement temperature. As-deposited $SiO_{1.3}$ has a T_0 near 63 K while after annealing the temperature stability is markedly reduced to T_0 of 29 K, which is nearly identical to a-Si ($T_0\sim$25 K). Optical absorption measured band gap is also consistent with the segregation forming a-Si, with the $E04$ decreasing from 2.2 to 1.8 eV after annealing (1.8 eV is characteristic of a-Si).

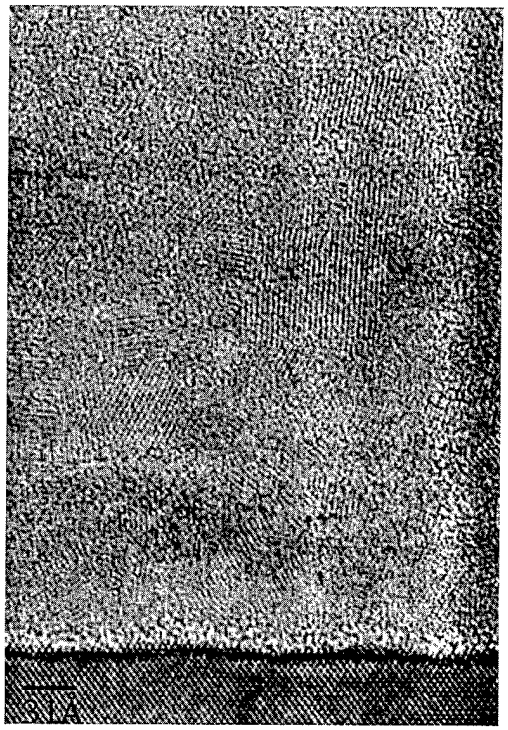


FIG. 4. Cross-sectional TEM of initial $SiO_{1.0}$ after annealing at 900 °C for 30 s clearly showing $SiO_x \rightarrow Si_{nc}+SiO_2$. At lower temperatures $<\sim$800 °C films are completely amorphous.

Observations in conductivity measurements also support an amorphous state segregation mechanism. Figure 6 shows both the dark- and photocurrent of films of various composition before and after annealing at 650 °C. For as-deposited films, the dark current is very low indicating a low level of Si dangling bond density. With increasing O content, the photocurrent decreases, which is expected for decreased absorption coefficient. After annealing, the dark current increases markedly as expected from the formation of Si dangling bonds. However at high O content, the photocurrent after annealing actually increases because of the increase in absorbance from a-Si regions. Both dark current and photocurrent decrease exponentially with increasing O content, which is consistent with a segregation model in which conduction paths between a-Si regions is reduced. Both PL and conductivity measurements show that Si dangling bond defect density is significant. Therefore it is of importance to determine whether the kinetics of SiO_x disproportionation are accelerated by defects. Figure 7 shows the extent of reaction as a function of annealing temperature for films of different initial H content. A factor of two difference in initial Si–H

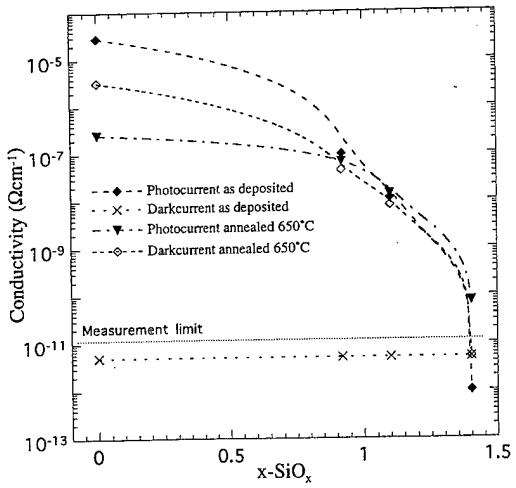


FIG. 6. Dark and photoconductivity measurements of various SiO_x films before and after anneals at 650 °C for 6 min. Changes in conductivity are consistent with a phase segregation in the amorphous state.

concentration does not effect observed kinetics, thus is not likely to be involved in the dominate mechanism.

B. Kinetic modeling

The previous experimental observations support the general mechanism of an initial rapid segregation of SiO_x into a-Si with an SiO_2 diffusion barrier which decelerates further reaction. This mechanism would also form a concentration gradient resulting in an oxygen rich $SiO_{(x+y)}$ matrix. To model the kinetics, we can apply the three-dimensional (3D) Avrami–Erofe'ye approach [Eqs. (5)–(7)] in which there is an instantaneous nucleation followed by a diffusion limited growth process.[15] α is the extent of reaction, N_c is nucleation density (cm^{-3}), D is diffusion coefficient (cm^2/s), and t is time of reaction in seconds. In this treatment an alloy of matrix composition (C_m, SiO_x) decomposes into nucleation sites of β phase (C_β, a-Si) and α phase (C_α, SiO_2) which acts as a diffusion barrier for the growth of β phase. This geometry is schematically shown in the inset of Fig. 8. With

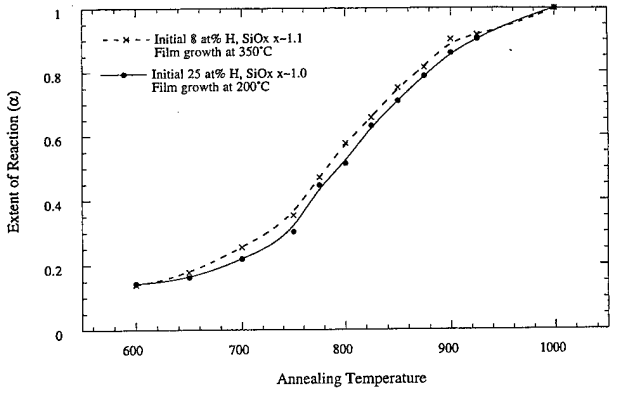


FIG. 7. Extent of reaction (α) for SiO_x thin films with different initial H content, as a function of annealing temperature showing nearly identical kinetics. Each series is annealed sequentially on the same film with annealing times of 6 min for each temperature.

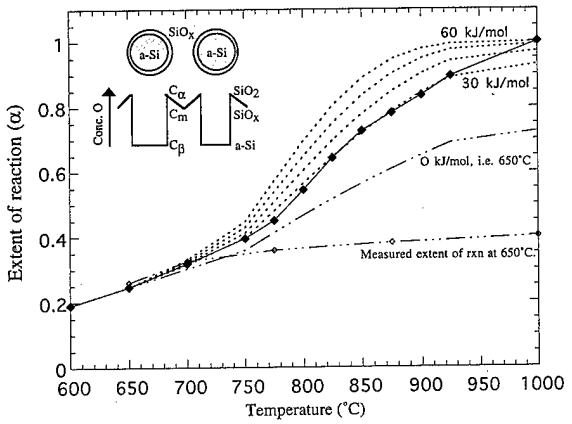


FIG. 8. Application of Avrami–Erofe'ye 3D treatment [Eqs. (5)–(7)] to observed α of $SiO_{1.1}$ decomposition (solid line). Inset figure shows geometry of transformation and oxygen gradient involved. Each solid point is a successive anneal at given temperature for 6 min. The dashed lines are calculated using activation energy (E_a) shown (30–60 kJ/mol). The open diamond points are isothermal data at 650 °C with times corresponding to cumulative times of temperature ramped data, which do not agree with predicted α (E_a=0 kJ).

O as the diffusion species, $C_x=(2.27\times10^{22}\ cm^{-3})\cdot x$, $C_\alpha=4.55\times10^{22}\ cm^{-3}$, $C_\beta=0$, R_f is the final Si_{nc} radius, and V_f is the final Si volume fraction calculated from stoichiometry (x). N_c is $1\times10^{18}/cm^3$ (from TEM size measurements at $\alpha=1$):

$$\ln(1-\alpha)=\frac{8\sqrt{2}\pi}{3}N_0D^{3/2}\left(\frac{(C_m-C_\alpha)}{(C_\beta-C_\alpha)}\right)^{1/2}t^{3/2}, \tag{5}$$

$$D=D_0\exp(-E_a/kT), \tag{6}$$

$$N_c=\frac{4/3\pi R_f^3}{V_f},\quad V_f=\frac{1-x/2}{1+0.6x}. \tag{7}$$

In Fig. 8, the solid points are the extent of reaction for initial $SiO_{1.1}$ after successive annealing for 6 min at each temperature. The dashed lines are calculated from Eqs. (5)–(7) using the activation energy (E_a) shown and the initial reaction rate at 650 °C. In other words, the rate of reaction at 650 °C (i.e., $\Delta\alpha$ 600–650 °C) is used to calculate D_0 for a given E_a. This D_0 and E_a are used to calculate α at successively higher temperatures for the 6 min periods, resulting in the dashed line in Fig. 8. As can be seen in Fig. 8, rather low activation energies of 30 kcal/mol are found to fit experimental data. This treatment predicts significant reaction at 650 °C while isothermal experiments (Fig. 2) show significant deceleration well before reaction completion. This is best seen in Fig. 8, where the calculated extent of reaction at 650 °C (E_a=0) is compared to experimental data (open diamonds) for anneals at 650 °C (in this case the x axis acts as time instead of temperature). The measured kinetics are significantly decelerated compared to those predicted. There is a fundamental inadequacy of the 3D Avrami–Erofe'ye treatment. That is while diffusion barrier decelerates the reaction by a factor of $\sqrt{2}$, the volume of a-Si phase grows as the cube of radius. This results in an overall $t^{3/2}$ dependence, which is accelerator. Significant deceleration would occur only near comple-

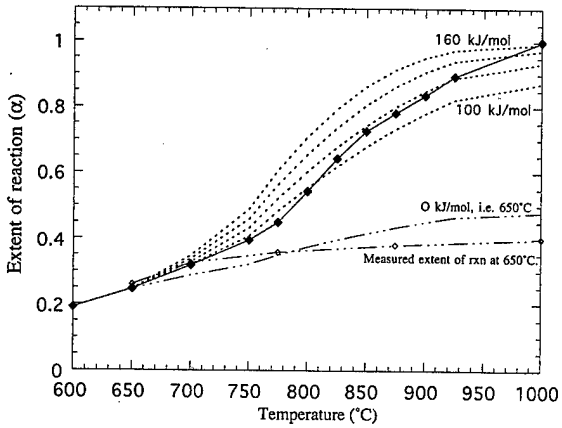


FIG. 9. Application of Avrami–Erofe've 1D treatment [Eqs. (8)–(9)] to observed α of $SiO_{1.1}$ decomposition (solid line). Each solid point is a successive anneal at given temperature for 6 min. The dashed lines calculated using shown activation energy (E_a) (100–160 kJ/mol). The open diamonds are isothermal data at 650 °C with times corresponding to cumulative times of temperature ramped data, which approach predicted α (E_a=0 kJ).

tion (α>0.8) from impingement, which is not experimentally observed. Thus the 3D Avrami–Erofe've treatment is not directly applicable to this system.

To account for the highly deceleratory nature of the reaction we can look at the dimensionality of phase growth. Typically the dimensionality of a reaction is found by microstructure of the resultant products, i.e., spherical particles imply 3D growth, platelets 2D, and needles 1D. The observed spherical Si nanocrystals imply 3D growth, however this may not necessarily reflect reactions in the amorphous state. The Si nanocrystals may be nucleated and grown in existing a-Si regions, representing a different reaction, which is 3D. Crystallization of a-Si regions in SIPOS has been similarly proposed from Raman experiments.[16] Since the initial segregation in amorphous Si is extensive, it is possible to have volume growth without significantly expanding surface area. Impingement upon itself and regions of decreasing reaction surface area will reduce the dimensionality of the growth reaction. Equation (8) describes the Avrami–Erofe've for the 1D limit:

$$\ln(1-\alpha)=\frac{A_i(C_m-C_\alpha)}{(C_\beta-C_\alpha)^{1/2}(C_\beta-C_m)^{1/2}}D^{1/2}t^{1/2}, \tag{8}$$

$$A_i=\frac{V_f}{d_f}, \tag{9}$$

where A_i is the initial surface area of the amorphous Si region after initial segregation reaction which is approximated by Eq. (9) and d_f is the average Si nanocrystal diameter as seen by TEM. Figure 9 shows the fit of Eq. (8) using the same method used in Fig. 8. The reduced dimensionality increases activation energy to 120 kJ/mol. Table I shows calculated parameters for 1D Avrami–Erofe've treatment. The extent of reaction (α) calculated for 650 °C (E_a=0) approaches that observed isothermal data shown in Fig. 9. Thus 1D Avrami–Erofe've is a reasonable limit for the observed kinetics. If we apply the calculated D to the initial

TABLE I. Parameters for kinetic data of $SiO_{1.1}$ disproportionation shown in Fig. 9 which are fit to 1D Avrami–Erofe've treatment [Eqs. (8)–(9)].

x_i=1.1	End thickness=50 Å
E_a=120 kJ/mol	$D_0=5.5\times10^{-9}$ cm²/s
D(650 °C)=8.9×10^{-16} cm²/s	
D(900 °C)=2.4×10^{-14} cm²/s	
Linear growth rate (α_i=0.19, 650 °C)=0.04 Å/s	
Linear growth rate (α_i=0.09, 900 °C)=1.2 Å/s	

barrier (α=0.19) we find a rapid linear growth rate near 1.2 Å/s at 900 °C. This rate is significantly faster and with lower E_a than the observed by Nesbit's treatment of Si_{nc} growth (E_a~180 kJ/mol), suggesting that those kinetic values are for the crystallization of a-Si with O impurities. This interpretation is further supported by recent work which has shown that as-deposited SIPOS films have significant segregation into a-Si regions (~50%).[17]

C. Uncertainty in kinetic parameters

It is important to quantify the uncertainties in the kinetic parameters found in the 1D Avrami–Erofe've treatment. Most significant is the uncertainty of IR measurements. As-deposited films contain significant amounts of Si–H which broaden the Si–O stretch peak and adds uncertainty to ν_i in Eq. (4). Uncertainty in ν_i will shift the position of all measured α. The most extreme case is if the observed shift in ν_{Si-O} after the 600 °C anneal was due solely to H loss and not to the reaction in Eq. (4). Using the kinetic treatment as in Fig. 9 on this limiting case would result in an E_a of 170 kJ/mol. However photoluminescence, ultraviolet visible (UV-vis) absorption, conductivity, and Raman data all show the presence of a-Si after low temperature anneals, thus there must be a disproportionation reaction at these low temperatures with simultaneous H loss. The data in Fig. 7 also strongly suggest the H content does significantly shift the Si–O stretch position. A factor of two reduction in the initial H content did not change α observed at 600 °C. A 50% error in α after 600 °C anneal would result in an E_a of 135 kJ/mol. Viscoelastic relaxation can play a role in shifting ν_f. Using α at 925 °C as a the lower limit where viscoelastic effects are known to be minimal, the maximum error in the derived E_a is +5 kJ/mol. Thus the E_a value of 120 kJ/mol for the 1D Avrami–Erofe've treatment is a lower limit with an uncertainty of +20 kJ/mol.

Any significant nucleation rate after initial rapid segregation would result in an accelatory reaction rate, which is not observed. A possibility is that there is no growth of existing nuclei and the measured extent of reaction is only due to forming new nuclei. The observed deceleration could be explained as function of a reduction in nucleation rate. An increase in temperature may allow otherwise hindered sites to nucleate. It is difficult to separate this effect from the growth of existing nucleation centers as in the Avrami–Erofe've treatment. This effect can be discounted for several reasons. Generally, impingement effects which decelerate nucleation typically are seen at α~0.8. Yet this system of deceleration

occurs around $\alpha \sim 0.2-0.4$. If only a nucleation mechanism were present, an SiO_2 shell would be formed and the IR measurements would show two peaks corresponding to SiO_2 and SiO_x, not the observed gradient. Thus there must be an a-Si phase growth mechanism. Any effects of nucleation at higher temperature would increase observed transformation rates, thus the E_a derived from the 1D Avrami–Erofe'ye treatment is a lower limit.

IV. CONCLUSIONS

SiO_x metastability in thick films can be found through kinetic hindrance. The disproportionation reaction proceeds by (1) rapid (<5 s) initial segregation forming a-Si and SiO_2 diffusion barrier, (2) highly deceleratory growth of a-Si regions form $SiO(x+y)$ diffusion gradient, (3) crystallization of a-Si regions into Si nanocrystals. FTIR, PL, TEM, Raman, and conductivity measurements all support a model with significant disproportionation in the amorphous state. Analysis of the highly deceleratory kinetics shows that the initial a-Si regions is extensive and growth is of low dimensionality. The 1D limit of Avrami–Erofe'ye treatment gives lateral growth rates of 1.2 Å/s (900 °C) and 0.04 Å/s (650 °C) with an activation energy of 120 kJ/mol, after initial segregation reaction. Due to the rapid initial segregation and relatively large lateral growth rate, a simple kinetic hindrance explanation for the 4–8 Å of interfacial SiO_x at the SiO_2/Si interface is unlikely.

ACKNOWLEDGMENTS

The authors would like to thank Meimei Xu and Professor Denis Maher for the collection of TEM data. This research has been funded by the National Science Foundation and Office of Naval Research.

[1]F. J. Grunthaner and P. J. Grunthaner, Mater. Sci. Rep. **1**, 65 (1986).
[2]T. Hatttori and T. Suzuki, Appl. Phys. Lett. **43**, 470 (1983).
[3]L. Brewer and R. K. Edwards, J. Electrochem. Soc. **58**, 351 (1954).
[4]M. Hamasaki, T. Adachi, S. Wakayama, and M. Kikuchí, J. Appl. Phys. **49**, 3987 (1978).
[5]J. Liday, S. Tomek, and J. Breza, Appl. Surf. Sci. **99**, 9 (1996).
[6]B. Greenberg and T. Marshall, J. Electrochem. Soc. **135**, 2295 (1979).
[7]G. Kragler, H. Bender, G. Willeke, E. Bucher, and J. Vanhellemont, Appl. Phys. A: Solids Surf. **58**, 77 (1994).
[8]M. Catalano, M. J. Kim, R. W. Carpenter, K. Das Chowdhury, and J. Wong, J. Mater. Res. **8**, 2893 (1993).
[9]D. J. Nesbit, Appl. Phys. Lett. **46**, 38 (1986).
[10]A. Banerjee and G. Lucovsky, in *Amorphous Silicon Technology-1996*, edited by M. Hack, R. Schropp, E. A. Schiff, A. Matsuda, and S. Wagner, Mater. Res. Soc. Proc. (MRS, Pittsburgh, PA, in press), Vol. 420.
[11]G. Lucovsky, D. V. Tsu, R. A. Rudder, and R. J. Markunas, in *Thin Film Processes II*, edited by J. L. Vossen and W. Kern (Academic, San Diego, CA, 1991), pp. 565–619.
[12]D. V. Tsu, G. Lucovsky, and B. N. Davidson, Phys. Rev. B **40**, 1795 (1989).
[13]J. I. Pankove, Appl. Phys. Lett. **32**, 812 (1978).
[14]R. A. Street, Adv. Phys. **30**, 593 (1981).
[15]M. E. Brown, D. Dollimore, and A. K. Galwey, in *Comprehensive Chemical Kinetics, Volume 22 Reactions in the Solid State*, edited by C. H. Bamford and C. F. H. Tipper (Elsevier, Oxford, 1980), pp. 68–78.
[16]D. J. Olego and H. Baumgart, J. Appl. Phys. **63**, 2669 (1988).
[17]M. Trchova, J. Zemek, and K. Jurek, J. Appl. Phys. **82**, 3519 (1997).

Cathodoluminescence measurements of suboxide band-tail and Si dangling bond states at ultrathin Si–SiO_2 interfaces

A. P. Young[a)]
Department of Electrical Engineering, The Ohio State University, Columbus, Ohio 43210-1272

J. Schäfer
Center for Materials Research, The Ohio State University, Columbus, Ohio 43210-1272

G. H. Jessen
Department of Electrical Engineering, The Ohio State University, Columbus, Ohio 43210-1272

R. Bandhu
Department of Physics, The Ohio State University, Columbus, Ohio 43210-1272

L. J. Brillson
Department of Electrical Engineering, Center for Materials Research, and Department of Physics, The Ohio State University, Columbus, Ohio 43210-1272

G. Lucovsky and H. Niimi
Department of Material Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 21 January 1998; accepted 28 May 1998)

We have directly observed cathodoluminescence (CL) in ultrahigh vacuum over a broad spectral range (0.7–4.0 eV) from ultrathin 5 nm layers of remote plasma enhanced chemical vapor deposition grown a-SiO_2:H deposited on silicon substrates. In the infrared regime, luminescence is observed at 0.8 eV, consistent with the presence in the as-deposited film of Si dangling bond localized states, as well as at 1.1 eV due to band edge emission. In the optical regime, three peaks (1.9, 2.7, and 3.4 eV) are observed showing evidence for band tail state emission from an amorphous silicon–oxygen bonded suboxide region in the film, with smaller contributions from either substrate related c-Si or defect containing, stoichiometric SiO_2. CL spectra with a range of beam voltages indicate that the stoichiometric SiO_2 is very close to the free surface of the film, possibly due to oxidation of the air-exposed wafer or due to a nonuniformity in the film. When the films are annealed *in situ* in stages up to 500 °C, we observe no change in the shape of the a-SiO_x:H peak at 1.9 eV, showing the stability of this suboxide to such temperatures. These observations are consistent with CL measurements of thicker films of a-SiO_2:H and a-SiO_x, and demonstrate the utility of CL spectroscopy for the study of ultrathin dielectric studies. © *1998 American Vacuum Society.* [S0734-211X(98)11904-2]

I. INTRODUCTION

Si–SiO_2 interfaces are a subject of both fundamental interest as well as being extremely important to the performance of electronic devices, notably, metal–oxide–semiconductor (MOS) type structures. Recently gate linewidths have shrunk deep into the submicron regime and, if the scaling laws of electronic devices continue to hold, oxide thicknesses must shrink as well. For amorphous SiO_2, the current oxide used in MOS structures to remain as a viable insulating material, the thickness of this layer must decrease to 3 nm or less in future generation devices.

At the Si–SiO_2 interface, it is well known that a silicon-rich suboxide 0.4–0.6 nm thick exists between the stoichiometric amorphous SiO_2 (a-SiO_2) and the crystalline Si (c-Si) substrate.[1] As the thickness required of the oxide shrinks to the point where a significant fraction of the suboxide is part of the interface, detailed knowledge correlating the chemical and electrical properties of the interface becomes of increasing importance to the performance of the device.

Because the thickness of the transition layer is so thin, it is difficult to easily deconvolute the electronic properties of the suboxide from the rest of the "bulk" stoichiometric oxide. Photoinduced luminescence is difficult with conventional sources, given the large band gap of SiO_2 (9.1 eV) and of other dielectrics.[2] Cathodoluminescence (CL) and photoluminescence (PL) spectroscopies have been used to study the electronic properties of bulk (~1 μm) SiO_2 films,[3,4] thick (~1 μm) SiO_x films,[5–7] as well as oxidized porous silicon structures.[8] While CL has been extensively used for the study of bulk SiO_2, to our knowledge, we know of no positive CL results for extremely thin, homogeneous, silicon dioxide films. In this work, we present what we believe is the first spectroscopic CL measurements of ultrathin (5 nm) SiO_2 films on Si substrates.

a)Electronic mail: apyoung@ee.eng.ohio-state.edu

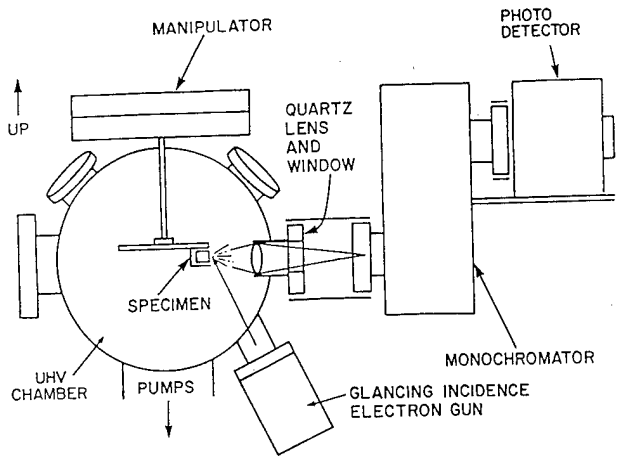


FIG. 1. Schematic diagram of the experimental apparatus showing the relative orientation of the electron gun, the *in situ* lens, and the external optical bench (from Ref. 13).

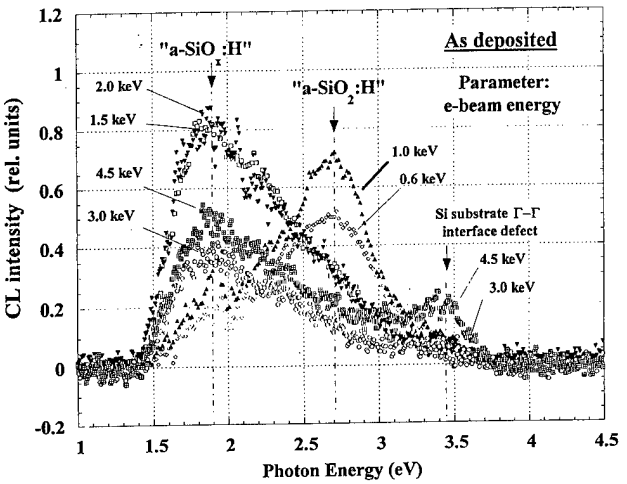


FIG. 2. Voltage-dependent CL spectroscopy of the as-deposited a-SiO_2:H film in the visible and near UV spectrum using an S-20 PMT. The spectra were taken with the substrate at room temperature.

II. EXPERIMENT

The SiO_2 films were initially grown with a ~0.5 nm oxide layer prepared by *in situ* plasma-assisted oxidation onto *p*-type silicon wafers in a multichamber UHV system for both oxides and dielectrics. On top of this layer, the 5-nm-thick, hydrogenated SiO_2 (a-SiO_2:H) films were deposited by remote plasma enhanced chemical vapor deposition (RPECVD). Additional details of the vacuum chamber apparatus and the growth process have been described elsewhere.[9–11] All the films were then transported through air to a separate UHV chamber with a base pressure of $\sim 1-2 \times 10^{-10}$ Torr for CL analysis.

The Si–SiO_2 films were measured by CL at room temperature and $T \sim 90$ K. As shown schematically in Fig. 1, an electron gun employing voltages and currents in the range 0.6–4.5 keV and ~1.0–4.0 μA, respectively, was used for the generation of the luminescence. The electron beam is incident on the sample at an angle of 45° with a spot diameter of ~200–500 μm. A liquid nitrogen cooled North Coast *p-i-n* Ge detector measured the luminescence in the spectral range 0.7–2.0 eV, while a thermoelectrically cooled S-20 photomultiplier tube (PMT) was used to measure the luminescence in the range 1.4–4.0 eV.

The Leiss monochromator employed contains a flint glass prism with nominally 50 meV resolution over the measurement range from the near infrared through the near ultraviolet (UV) part of the spectrum. For the *in situ* annealing experiments, the specimens were rapidly brought up to temperature (~1 min), held for 5 min at 400 °C or 500 °C, then brought back to room temperature for analysis. During the annealing process, the specimen temperature was measured using an infrared pyrometer. The CL spectra presented here have not been corrected for the spectral responsivity dependence of the detector or the transmission of the optical train.

III. RESULTS

Even though the a-SiO_2:H film was only 5 nm thick, the luminescence was strong enough to be faintly visible to the naked eye at room temperature. The light appeared white, with blue and red components mixed in. Besides providing quantitative luminescence spectra for the wide band gap materials, CL has another distinct advantage over optical excitation techniques, namely, a large variable excitation depth controllable via the incident electron beam energy. Figure 2 shows spectra from the as-deposited specimen for various excitation voltages (0.6–4.5 keV) with discrete peaks apparent in the visible and near UV region of the electromagnetic spectrum. There is an obvious change in spectral features as a function of beam energy. The low energy spectra at 0.6 and 1.0 keV show broad luminescence peak features centered at 2.7 and 1.9 eV. When the beam intensity is increased from 1.5 up to 4.5 keV, the features change dramatically, with a much broader band centered at 1.9 eV and a weak peak developing at 3.4 eV, which increases with increasing beam voltages.

The temperature dependence of the luminescence can help to clarify the stoichiometry and bonding arrangements of the various forms of Si and SiO_x in the film, since the luminescence from a-Si is known to be strongly quenched, while a-SiO_x:H does not show this behavior. As can be seen in Fig. 3, after lowering the temperature of the as-deposited film to 90 K, the overall spectrum shape did not change significantly; there was only a minor intensity increase.

In addition to the spectra in the visible and the UV, the infrared spectrum is shown in Fig. 4 for the as-deposited film at an excitation voltage of $E=2$ keV. Two peaks were clearly resolved at ~0.8 and ~1.0 eV. The infrared spectrum at 2 keV was typical. For voltages between 0.6 and 3.0 keV, the only significant change observed was an increase in the relative intensity of the 1.0 eV band compared to the 0.8 eV band as the excitation voltage increased. Also shown in Fig.

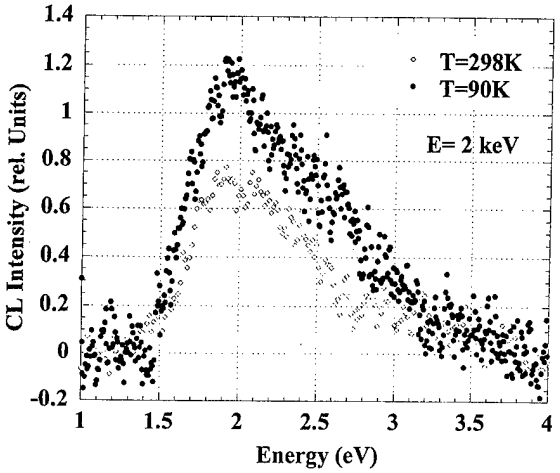


FIG. 3. Temperature-dependent CL spectra of the as-deposited a-SiO_2:H. Notice the lack of change in the spectra intensity, consistent with the picture of an amorphous suboxide film.

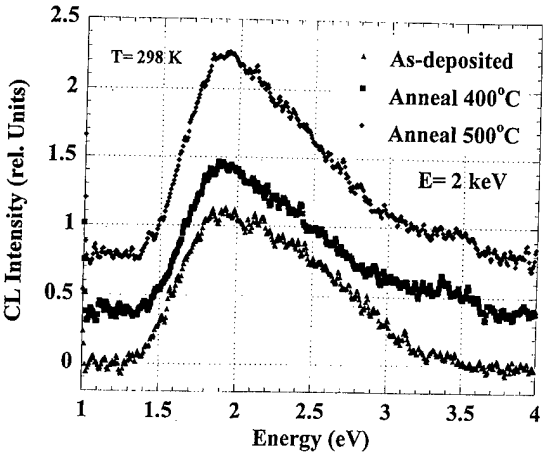


FIG. 5. Room temperature CL spectroscopy of the a-SiO_2:H film after 5 min consecutive anneals between room temperature and 500 °C using the S-20 PMT detector. All spectra were obtained with an electron-beam voltage of 2 keV. There is no change in the overall spectra with annealing, indicating no change in the stoichiometry of the film up to 500 °C. The spectra are offset for clarity.

4 is the CL spectrum after the specimen was annealed in 5 min stages *in situ* up to 500 °C. In Fig. 4, it is clear that there is dramatically enhanced intensity at 0.8 eV after the annealing compared to the as-deposited case.

Figure 5 shows the corresponding visible/UV spectrum as a function of annealing temperature for the same excitation voltage. In contrast with the strong increase in luminescence in the infrared, there is relatively little change in intensity observed in the visible region of the spectrum. The only significant change that we observed is the slight strengthening of the 3.4 eV band as the annealing progresses.

We also addressed the possibility of electron-beam damage of the film. During our experiments, the incident beam delivers ~milliwatt over $\sim 10^{-3}$ cm^2. Given the relatively high thermal conductivity of the p-doped substrate and the minimal thickness of the SiO_2 overlayer, this input power should not produce any significant thermal damage, especially in comparison to bulk SiO_2 studies. In addition, for a 2 keV beam energy with a current of 1 μA, the total exposure is on the order of 10^{17} electrons/cm^2, two orders of magnitude below the reported damage threshold for a thicker SiO_2 film.[12] For our experiment for possible damage, a typical CL spectrum was measured, then the beam was allowed to remain on one spot for 2 h, and the spectra was then remeasured. As can be seen in Fig. 6, there is only a minor change in the intensity of the 1.9 eV peak. These spectral changes are within the limits of reproducibility of our experiment. Typical CL spectra are obtained with over an order of magnitude lower exposure (flux times time), so that damage should not be a significant problem for this system with these short (~12 min) exposure times and currents.

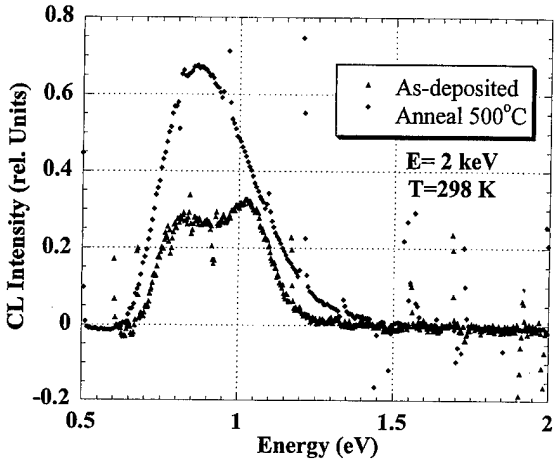


FIG. 4. Room temperature infrared CL spectra of the a-SiO_2:H film before and after a 5 min anneal at 500 °C. The spectra were obtained with a Ge detector and with an electron-beam voltage of 2 keV. The 0.8 eV peak increases significantly with annealing, indicating an increase in luminescence from Si dangling bonds, and correlating with the removal of H from the film.

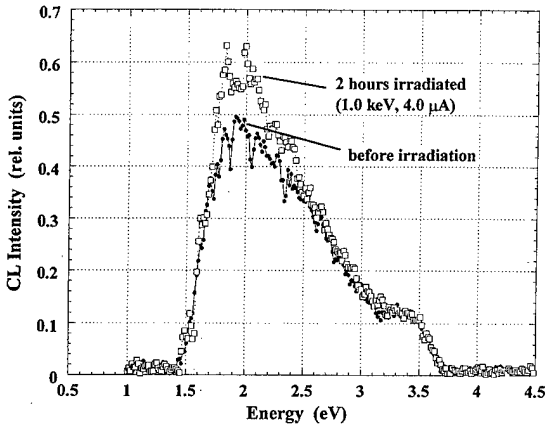


FIG. 6. Test for the possibility of electron-beam damage during the measurement of a single CL spectrum. Only a small increase in the 1.9 eV peak is observed after 2 h of continuous irradiation on a single spot.

IV. DISCUSSION

The high energy of the incident electron beam relative to the insulator band gap, i.e., 1–2 keV vs 9.1 eV, is a distinct advantage over traditional optical excitation for the generation of large densities of electron-hole pair recombination in wide band gap materials such as SiO_2. With such a high energy excitation source, the initial electron generates a cascade of electrons and holes, multiplying its efficiency. Simultaneously, compared to the 15–20 keV normally used by a scanning electron microscope for CL, the relatively low 0.5–5 keV beam we use has a significantly shallower penetration depth compared to typical scanning electron microscopy based CL and PL spectroscopies.[13,14] This shallow penetration depth is particularly true in wide band gap materials. For calculations based on Ref. 13, the range for a 1 keV electron into a material with the density of SiO_2 is ~10 nm with the peak of the deposited energy at 1/3 of this range. Therefore low energy CL is particularly well suited for probing thin, wide band gap layers and can explain why CL has been successful in producing detectable luminescence from the SiO_2/Si interface, in contrast to more penetrating laser excitation.

From the data presented in Fig. 2, we can immediately identify the presence of stoichiometric SiO_2 at the very top surface layer. The 2.7, 1.9 eV combination is very similar to that observed in bulk a-SiO_2.[3] At the same time, for the higher excitation energies (1.5–4.5 keV), the 2.7 eV peak diminishes significantly, and in its place is a single, broad peak at 1.9 eV [full width at half maximum (FWHM)=0.9 eV], very suggestive of band-tail transitions in an amorphous, hydrogenated, Si-rich suboxide.[15–17] Additional evidence supporting the presence of this local silicon–oxygen bonding is provided by the data in the infrared. The 0.80 eV peak has been observed in thick, a-Si:H films corresponding to an optical transition from an amorphous Si band tail to midgap dangling bond states.[18] In addition to the 0.80 eV peak, the luminescence band at 0.95–1.05 eV can be assigned to the crystalline silicon substrate, since the band edge (E_g=1.12 eV) emission increased as the electron beam penetrated deep into the substrate. Therefore, both the infrared and visible spectra support the presence of different silicon–oxygen and silicon–silicon bonding arrangements in the film.

The final peak feature observed is a weak ~3.4 eV shoulder at higher incident beam energies in the as-deposited specimen. A clue to its origin is the increasing intensity of the band as the electron beam penetrates deeper into the film. From this observation and the lack of a strong 2.7 eV peak, the luminescence band might be due to the Γ–Γ transition (~3.5 eV) of bulk crystalline silicon, most likely coming from the substrate not the SiO_2 film itself. On the other hand, this band could also be due to a defect in SiO_2 where peaks have been observed before by PL;[19] however, most luminescence measurements of SiO_2 do not have this defect feature and, when present, it is only in combination with luminescence from other defect bands between 2.2 and 2.7 eV, which are not resolved in these spectra. Therefore, at this time we cannot clearly correlate this feature with the properties of bulk films.

From the temperature dependence of the as-deposited film, we see only a minor change in the 1.9 eV band intensity down to T=90 K. This observation further supports the conclusion that the luminescence band is associated with SiO_2 or a suboxide rather than a-Si:H, whose PL spectra is known to be strongly quenched at room temperature.[20]

The *in situ* annealing experiment points out the stability of the main bonding arrangement, and it emphasizes the role of H in passivating Si dangling bonds in the film. While the visible and UV range spectra in Fig. 5 show little change, the infrared spectrum shows a dramatic increase in the intensity of the 0.8 eV band, consistent with the evolution of H from the film and the consequent increase of Si dangling bond states. Annealing to higher temperatures such as 900 °C, where it is known any residual silicon would completely separate into c-Si and SiO_2,[20] would help to clarify the role of H and the excess Si.

V. CONCLUSION

In summary, we have directly observed cathodoluminescence from ultrathin 5.0 nm SiO_2 films. From our measurements, we can reducibly differentiate between different bonding arrangements in these ultrathin films using a relatively simple, yet surface sensitive, characterization technique. While the layers are thin, they are thick enough to show a characteristic luminescence spectrum similar to what is observed in thicker films. The 5 nm SiO_2 grown by RPECVD is shown to initially have silicon-rich regions in addition to stoichiometric SiO_2 stable up to at least 500 °C. Simultaneously, the passivating role of H on Si dangling bonds was also clearly observed in the luminescence spectra. From these measurements, it is clear CL is a powerful tool for probing the electronic structure at these ultrathin Si–SiO_2 interfaces and for seeing how they change with thermal annealing.

ACKNOWLEDGMENTS

This work is supported in part by the NSF Division of Materials Research, the Office of Naval Research, and the NSF ERC for Advanced Electronic Materials Processing.

[1]F. J. Grunthaler and P. J. Grunthaler, Mater. Sci. Rep. **1**, 65 (1986).
[2]W. C. Choi, M. S. Lee, C. K. Kim, S. K. Min, C. Y. Park, and J. Y. Lee, Appl. Phys. Lett. **69**, 3402 (1996).
[3]M. A. Kalceff and M. R. Phillips, Phys. Rev. B **52**, 3122 (1995).
[4]S. W. McKnight and E. D. Palik, J. Non-Cryst. Solids **40**, 595 (1980).
[5]B. J. Hinds, F. Wang, D. M. Wolfe, C. L. Hinkle, and G. Lucovsky, J. Non-Cryst. Solids (to be published).
[6]F. Koch and V. Petrova-Koch, J. Non-Cryst. Solids **198–200**, 840 (1996).
[7]Y. Nakayama, M. Uecha, and T. Ikeda, J. Non-Cryst. Solids **198–200**, 915 (1996).
[8]A. G. Cullis, L. T. Canham, and P. D. J. Calcott, J. Appl. Phys. **82**, 909 (1997), and references therein.
[9]G. Lucovsky, A. Banerjee, B. Hinds, B. Claflin, K. Koh, and H. Yang, J. Vac. Sci. Technol. B **15**, 1074 (1997).

[10]D. R. Lee, G. Lucovsky, M. S. Denker, and C. Magee, J. Vac. Sci. Technol. A **13**, 1671 (1995).
[11]T. Yasuda, Y. Ma, S. Habermehl, and G. Lucovsky, Appl. Phys. Lett. **60**, 434 (1992).
[12]J. S. Johannessen, W. E. Spicer, and Y. E. Strausser, J. Appl. Phys. **47**, 3028 (1976).
[13]B. G. Yacobi and D. B. Holt, *Cathodoluminescence Microscopy of Inorganic Solids* (Plenum, New York, 1990), p. 151.
[14]L. J. Brillson and R. E. Viturro, Scanning Microsc. **2**, 789 (1988).
[15] M. A. Paesler, D. A. Anderson, E. C. Freeman, G. Moddel, and W. Paul, Phys. Rev. Lett. **41**, 1492 (1978).
[16]J. C. Knights, R. A. Street, and G. Lucovsky, J. Non-Cryst. Solids **35–36**, 279 (1980).
[17]R. Carius, R. Fischer, E. Holzenkampfer, and J. Stuke, J. Appl. Phys. **52**, 4241 (1981).
[18]R. A. Street, Adv. Phys. **30**, 593 (1981).
[19]K. Kim, M. S. Suh, T. S. Kim, C. J. Youn, E. K. Suh, Y. J. Shin, K. B. Lee, H. J. Lee, M. H. An, H. J. Lee, and H. Ryu, Appl. Phys. Lett. **69**, 3908 (1996).
[20]B. Hinds, A. Banerjee, R. S. Johnson, and G. Lucovsky, Mater. Res. Soc. Symp. Proc. (in press).

Effects of interface roughness and conducting filaments in metal–oxide–semiconductor tunnel structures

D. Z.-Y. Ting
Thomas J. Watson, Sr. Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125 and Department of Physics, National Tsing Hua University, Hsinchu 30043, Taiwan

T. C. McGill
Thomas J. Watson, Sr. Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125

(Received 21 January 1998; accepted 21 May 1998)

The current–voltage characteristics of n^+ poly-Si/SiO_2/p-Si tunnel structures containing nonuniform ultrathin oxide layers are studied using three-dimensional quantum mechanical scattering calculations. We find that, in general, roughness at the Si/SiO_2 interface renders the oxide layer more permeable. In the direct-tunneling regime, interface roughness induces lateral localization of wave functions, which leads to preferential current paths. But in the Fowler–Nordheim tunneling regime it affects transport primarily through scattering. These two distinct mechanisms lead to opposite current density dependencies on island size. We have also examined oxide-embedded conducting filaments, and found that they act as highly efficient localized conduction paths and lead to dramatic increases in current densities. Depending on the filament length, our model can mimic experimental current voltage for ultrathin oxides having undergone either quasibreakdown or breakdown. We also found that the lower bias current densities in the structure with long filaments are greatly enhanced by resonant tunneling through states identified as quantum dots, and that this current enhancement is highly temperature dependent. We also report on the dependence of current–voltage characteristics on filament diameter size and filament density.
© *1998 American Vacuum Society.* [S0734-211X(98)07804-X]

I. INTRODUCTION

The continued scaling of metal–oxide–semiconductor (MOS) device structures has brought much attention to ultrathin oxides. Normal operation of MOS field-effect transistor with 1.5 nm direct-tunneling gate oxide has been reported.[1] Tunneling through oxide barriers, as a mechanism for leakage currents, is of particular interest. Typical theoretical analysis models the oxide layer as a one-dimensional barrier with an effective barrier height and an effective mass. The barrier height may be obtained experimentally or treated as a fitting parameter, while the effective mass is normally used as a parameter for fitting measured current–voltage (I–V) characteristics. Tunneling coefficients can be calculated using the well-known WKB approximation. Approximate integration of tunneling coefficient curves, with the appropriate Fermi factors describing carrier statistics, then yields an analytical I–V curve formula for the direct-[2] and Fowler–Nordheim[3] tunneling regimes, which can be used conveniently for comparison with experimental data. A somewhat similar treatment uses multiple scattering theory instead of the WKB approximation to compute tunneling coefficients to provide clarification of mechanisms for leakage currents through ultrathin oxides.[4] A still more advanced treatment solves Poisson and Schrödinger equations self-consistently for accumulated layers in MOS devices to calculate tunneling currents.[5]

Nonuniformity in oxides are, typically, not treated theoretically due to the much increased complexity and computational demands. Yet, they can have dramatic effects on the current–voltage characteristics of MOS tunnel structures with ultrathin oxide barriers. One example is interface roughness. If we view interface roughness as local fluctuations in oxide thickness, then this fluctuation as a percentage of total oxide thickness can be quite large in ultrathin oxides. In addition, recently Cundiff and co-workers[6] found experimental evidence that, in typical industrial oxides, roughness at the Si/SiO_2 interface increases with decreasing oxide layer thickness; this further enhances the importance of interface roughness in ultrathin oxides. Another type of nonuniformity is conducting filaments embedded in oxides. It has been shown that constant current stressing of MOS structures in the Fowler–Nordheim tunneling regime can lead to quasibreakdown or breakdown in ultrathin oxides, which are characterized by dramatic increases in leakage currents. Based on experimental observations, several groups have proposed the formation of oxide-embedded conducting filaments as a model for breakdown. Hirose and co-workers[7] proposed that the onset of dielectric degradation takes place rather homogeneously close to the SiO_2/Si interface where the Si–O bonds are heavily strained. Based on their data, they postulated the existence of localized conducting filaments approximately 50 nm in diameter, and extending for no more than 3 nm from the SiO_2/Si interface into the oxide layer. Apte and Sarawat[8] proposed a physical-damage model of dielectric breakdown where the damages in the form of broken bonds in the strained SiO_2 layer near the anode links up with islands of bulk damages to create filamentary paths, which enables excessive conduction. Halimaoui and co-workers[9]

envisioned that the oxide layer contains narrow paths (defects) running from anodes to cathodes. Under currents stressing, they merge to form larger conducting paths, resulting in quasibreakdown; further stressing leads to the merging of quasibreakdown paths and causes breakdown. In this article we use a three-dimensional (3D) model which allows us to analyze the current–voltage characteristics of MOS tunnel structures containing nonuniform oxide layers. Specifically, we examine the cases of interface roughness and conducting filaments.

II. METHOD

Standard treatment uses a one-dimensional potential to describe the oxide barrier. With interfacial nonuniformity, we need to use a three-dimensional description. In principle, variations in the nonuniform potential extend indefinitely in the directions along the interface. In practice, we do not perform computation on an infinite domain, but use instead a quasi-3D supercell geometry to approximate the physical structure. We treat the problem of tunneling through a nonuniform barrier using the open-boundary planar supercell stack method (OPSSM).[10] The device structure treated by OPSSM consists of an active layer sandwiched between two semi-infinite flatband electrode regions. Let the z axis be the direction perpendicular to the interfaces. Then, the active region is composed of a stack of N_z layers perpendicular to the z direction, with each layer containing a periodic array of rectangular planar supercells of $N_x \times N_y$ sites. A one-band nearest-neighbor tight-binding Hamiltonian is used to describe the potential and effective-mass variations over this volume of interest. Our model is formally equivalent to the one-band effective-mass equation[11]

$$-\frac{\hbar^2}{2}\nabla\cdot\frac{1}{m^*(\mathbf{x})}\nabla\psi+V(\mathbf{x})\psi=E\psi, \tag{1}$$

discretized over a Cartesian grid, and subject to periodic boundary conditions (with supercell periodicity) in the x and y directions, and open-boundary conditions in the z direction. Since we are free to choose the values of $V(\mathbf{x})$ and $m^*(\mathbf{x})$ at each of the $N_x \times N_y \times N_z$ sites in our computational domain, we have tremendous flexibility in dictating the geometry of the device structure we simulate. OPSSM solves the quantum mechanical scattering problem exactly for the 3D geometry described by the planar supercell stack, and allows us to compute transmission coefficients with a high degree of numerical accuracy and efficiency. Note that even though the supercell geometry imposes an artificial periodicity to make computations tractable, the use of sufficiently large supercells can minimize supercell artifacts and yield excellent descriptions of the physical problem. Once transmission coefficients are obtained, current densities can be obtained using the standard formula[12]

$$J=\frac{em^*kT}{2\pi^2\hbar^3}\int_0^\infty T(E,V)\ln\left[\frac{1+e^{(E_F-E)/kT}}{1+e^{(E_F-E-eV)/kT}}\right]dE. \tag{2}$$

III. RESULTS AND DISCUSSION

Our model of the MOS tunnel structure consists of an n^+ poly-Si electrode, followed by the oxide layer, and finally, a p-type silicon region. The conduction-band edge is chosen to be at $E_c^M=0$, and the poly-Si Fermi level at $E_F^M=0.1$ eV. The tunneling barrier height at the n^+ poly-Si/SiO_2 interface is taken to be $\Phi^B=3.25$ eV,[7] and the SiO_2/p-Si conduction-band offset is taken as 3.29 eV.[13] The p-Si Fermi level is chosen to be 0.88 eV below the Si conduction-band edge. Thus, at zero gate bias, the p-Si conduction-band edge is 0.98 eV higher than the n^+ poly-Si conduction-band edge, and a bias of $V_{FB}=-0.92$ V is required to bring the oxide into flatband (FB) condition. The effective masses of the n^+ poly-Si, SiO_2, and p-Si are taken to be 1.0, 0.35, and 0.9 m_0, respectively. For convenience, we also assume flatband conditions in the electrodes, and let all the voltage drop occur in the oxide. This should be valid for the high doping levels considered for this structure. We use a cubic mesh with discretization distance of 0.135 75 nm, and 32×32 or 64×64 planar supercells in our simulations. We will consider the characteristics of these structures under negative gate biases (i.e., p-Si lowered relative to poly-Si).

A. Interface roughness

We consider a MOS tunnel structure with a 0.27 nm rough interfacial layer sandwiched in between a 1.36 nm pure oxide layer and the p-Si region. We assume that the rough interfacial layer consists of a 50%–50% mixture of oxide and Si in random configurations. The Si sites, and the oxide sites, for that matter, may aggregate and form patches. We will call the silicon patches islands, and characterize them by their lateral extent (average island size, λ) and the thickness of the interfacial layer (island height, h).

Figure 1 shows the calculated J–V curves for three MOS tunnel structures with rough Si/SiO_2 interfaces characterized by average island sizes of $\lambda=0.33$, 0.97, and 2.95 nm. For comparison, we also construct a reference structure with a smooth interface by replacing the rough interfacial layer with a pure oxide layer of the same thickness (resulting in a total oxide thickness of 1.63 nm). We note that in the direct tunneling regime ($|V_G|<4$ V), current density increases with island size. For instance, at $|V_G|=2$ V, the $\lambda=0.33$, 0.97, and 2.95 nm structures show current densities at 2.7, 3.4, and 4.6 times higher than the reference structure, respectively. This is the result of lateral localization of tunneling electrons, and can be understood by analyzing transmitting state wave functions in the rough interfacial layer. Let the silicon island transmission fraction be defined as the ratio of the sum of silicon site probability densities in the interfacial layer divided by the total probability density in the same. Since in this case the interfacial layer consists of 50% silicon sites and 50% oxide sites, a fraction greater than 0.5 would indicate a preference for transmission through the silicon islands. Figure 2 shows the silicon island transmission fractions as functions of electron incident energy for the three structures at $|V_G|=2$ V and $|V_G|=5$ V, representing direct and Fowler–Nordheim tunneling cases, respectively. At $|V_G|$

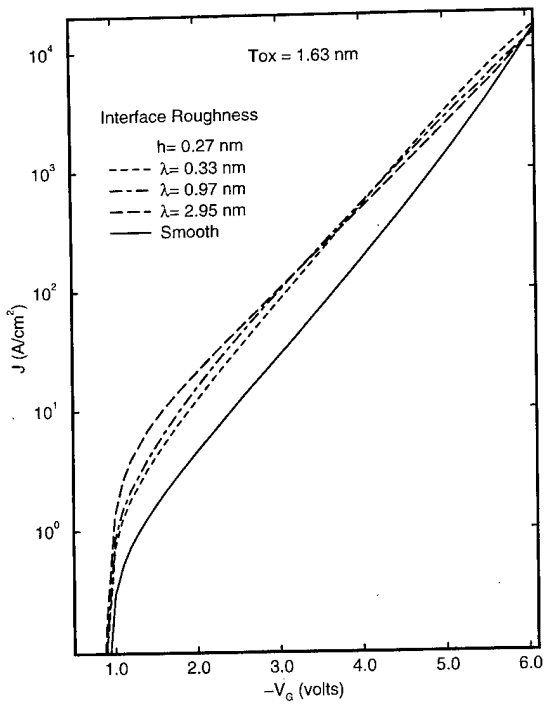


FIG. 1. Supercell calculation of current density–voltage curves for a set of n^+ poly-Si/SiO_2/p-Si tunnel structures with varying degrees of interface roughness. The oxide thickness is 1.63 nm.

$=2$ V, the electron deBroglie wavelength in the Si portion of the rough interfacial layer is approximately $\lambda_e \approx 1$ nm. In this instance the Si island transmission fractions are fairly constant in the energy range shown, yielding values of 0.54, 0.65, and 0.87 for the $\lambda = 0.33$, 0.97, and 2.95 nm structures, respectively. Evidently, in the structure with average island size λ much smaller than the electron deBroglie wavelength λ_e, there is only a very slight preference for transmission through the silicon sites. But in the structure with λ considerably larger than λ_e, an electron can readily distinguish the oxide energy barriers form the silicon open pathways, and preferentially traverses the silicon sites to which it is laterally localized. The localization in the more conducting portion of the rough interfacial layer leads to higher current densities for structures with larger islands.

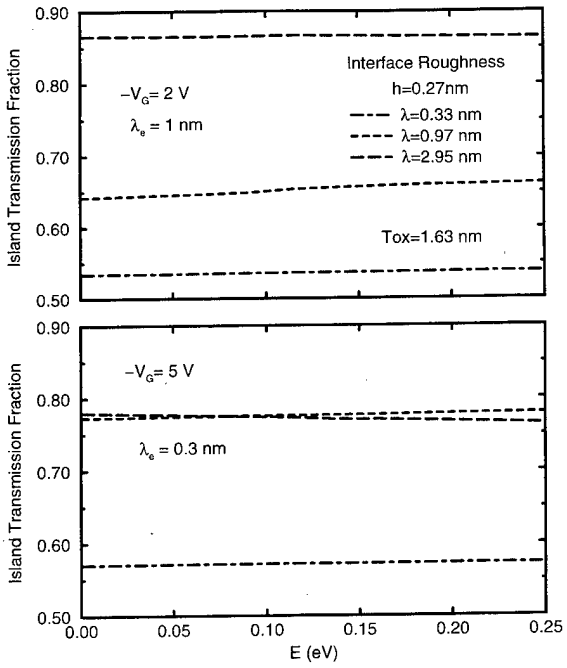


FIG. 2. Silicon island transmission fractions as functions of incident electron energy at two different gate biases for a set of MOS tunnel structures with varying degrees of interface roughness.

The bottom portion of Fig. 2 shows the silicon island transmission fractions for $|V_G| = 5$ V. Comparing to the $|V_G| = 2$ V case, the trailing interface is biased lower with respect to the poly-Si electrode. Thus, upon reaching the rough interface, an electron would be relatively more energetic (we assume it does not suffer inelastic scattering), with correspondingly shorter deBroglie wavelength of $\lambda_e \approx 0.3$ nm. For the structures with $\lambda = 0.97$ and 2.95 nm (both considerably larger than λ_e), the silicon island transmission fractions have comparable values of ≈ 0.77. And for the smaller island structure ($\lambda = 0.33$ mm), the fraction is only ≈ 0.57, which still does not much exceed 50%. Judging by these results, at $|V_G| = 5$ V, we might expect comparable current densities for the $\lambda = 0.97$ and 2.95 nm structures, and a smaller current density for the $\lambda = 0.33$ nm structure. But this is not the case. Figure 1 shows that in the Fowler–Nordheim tunneling regime ($|V_G| > 4$ V), current density decreases with island size; this trend is the opposite of that for the direct tunneling case. Clearly, a physical mechanism other than localization must be invoked in order to explain this behavior.

In the Fowler–Nordheim tunneling regime ($|V_G| > 4$ V), the conduction-band edge at the trailing interface of the barrier (in our case, the rough interfacial layer) is biased below incoming electron energies. Unlike in the direct-tunneling regime, where an electron traverses the oxide portions of the interface with evanescent characteristics, in the Fowler–Nordheim tunneling regime, an electron transmits through both the oxide and silicon portions of the interfacial layer with propagating characteristics. Therefore, the rough interface affects transport primarily through scattering, rather than lateral localization. Here, the larger islands produce more scattering, and thereby reduce the transmission amplitude in the forward direction. Figure 1 shows that, indeed, contrary to the direct-tunneling case, in the Fowler–Nordheim tunneling current densities decrease with increasing island size. Note also that at higher biases the size of the tunneling current is primarily determined by the leading edge of the tunnel barrier. Since the four structures differ only at the trailing edge, the differences in their current densities tend to be less pronounced than in the direct-tunneling regime.

B. Conducting filaments

It has been conjectured that quasibreakdown and breakdown in MOS tunnel structures are the results of current-stressing-induced conducting filaments in the oxide layer.[7–9]

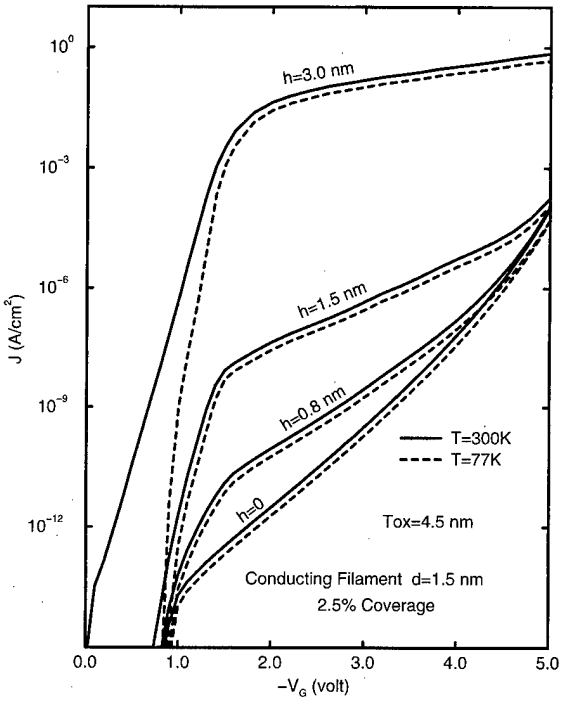


FIG. 3. Calculated current density–voltage curves at 300 and 77 K for a set of n^+ poly-Si/SiO_2/p-Si tunnel structures with oxide-embedded cylindrical conducting filaments with various cylinder heights. Oxide thickness is 4.5 nm. Conducting filaments have a diameter of 1.55 nm, and cover approximately 2.5% of the cross-sectional area.

In our simulation we consider a set of structures with 4.5 nm thick oxides, embedded with cylindrical conducting filaments 1.55 nm in diameter. 64×64 planar supercells were used in our simulations. The filaments account for approximately 2.5% of our computational domain in cross-sectional area, and extend from the SiO_2/Si interface into the oxide layer with cylinder heights of $h=0.8$, 1.5, and 3.0 nm. A fourth, "undamaged" ($h=0$) structure is also included for comparison. Because the nature of the filamentary conducting material is not well known, we choose to fill the cylinders with silicon for simplicity. Figure 3 shows the J–V curves for these structures calculated at 300 and 77 K. We note that, in general, current densities increase dramatically with filament length h. This again is due to lateral localization of transmitting state wave functions. We similarly define the filament transmission fraction for a transmitting state as the sum of probability densities over the filament sites, divided by the total probability densities in the filament-containing layers. Table I shows filament transmission fractions for tunneling states with incoming energy equal to E_F^M at gate biases of $|V_G|=2$ V (direct tunneling) and $|V_G|=5$ V (Fowler–Nordheim tunneling). In all cases (with perhaps the exception of $h=0.8$ nm at $|V_G|=5$ V), the filament transmission fraction greatly exceeds 2.5%, the fraction of cross-sectional area occupied by the filaments. This clearly indicates that conduction is strongly localized to the filaments. Table I also shows that the confinement is weaker in the Fowler–Nordheim regime, especially for the shorter filaments. This is because the trailing edge of the oxide does not act as confining barriers when its band edge is biased below the incoming electron energy. Again, since tunneling properties in this regime are primarily determined by the leading edge of the oxide barrier, the $h=0.8$ and 1.5 nm structures show current densities which converge with that of the undamaged ($h=0$) structure at high bias. The filament in the $h=3.0$ nm structure extends sufficiently close to the leading edge of the tunnel barrier so that the large current density increases persist even at higher biases. Our $h=1.5$ nm and $h=3.0$ nm curves bear strong qualitative resemblance to experimental I–V curves for the quasibreakdown[7,9] and breakdown[9] cases, respectively. It is worth noting that within our model we have reproduced both the quasibreakdown and the breakdown behaviors using the same mechanism, with the only difference being whether the filaments extend from the trailing interface far enough into the oxide layer towards the leading interface to have a substantial impact on the Fowler–Nordheim tunneling characteristics. We note that, in particular, we can reproduce the breakdown behavior without using oxides which run through the entire length of the oxide layer; we have kept the filament height to under 3 nm, as suggested by Hirose and co-workers.[7]

TABLE I. Filament transmission fraction calculated using transmitting state wave functions for electrons with incident energy $E=E_F^M$.

$\|V_G\|$ (V)	$h=0.8$ nm	$h=1.5$ nm	$h=3.0$ nm
2.0	0.43	0.58	0.66
5.0	0.06	0.16	0.64

The $h=3.0$ nm curve differs from the others in its low-bias temperature dependence. In the $h=0$, 0.8, and 1.5 nm cases, the 77 and 300 K results appear essentially the same. In the $h=3.0$ nm curve, however, current densities at low biases ($|V_G|<|V_{FB}|$) increase significantly with temperature. This turns out to be due to resonant tunneling through quantum dots states. We describe their properties in detail elsewhere.[14] Suffice it to say here that the quantum dots are laterally localized in the cylinders, and electrostatically confined along the third direction. When the bias is low ($|V_G|<|V_{FB}|$), the p-Si band edge is actually higher than that of the incoming electrode. Therefore, current contributing resonance levels must be above the p-Si band edge to allow electrons to scatter elastically into the p-Si conduction band through resonant tunneling. This, typically, places these resonance levels far above the Fermi level or the incoming electrode, therefore, their contributions to current densities are highly sensitive to temperature.

In Fig. 4 we examine the dependence of current densities on the fraction of gate area occupied by the filaments. We have performed calculations where we kept the filament diameter constant, but reduced the supercell size from 64×64 to 32×32, thereby increasing filament density by a factor of 4. We find that the corresponding current density increase is almost exactly fourfold in the direct-tunneling regime. The fact that the current density scales linearly with filament density indicates strongly that filament conduction

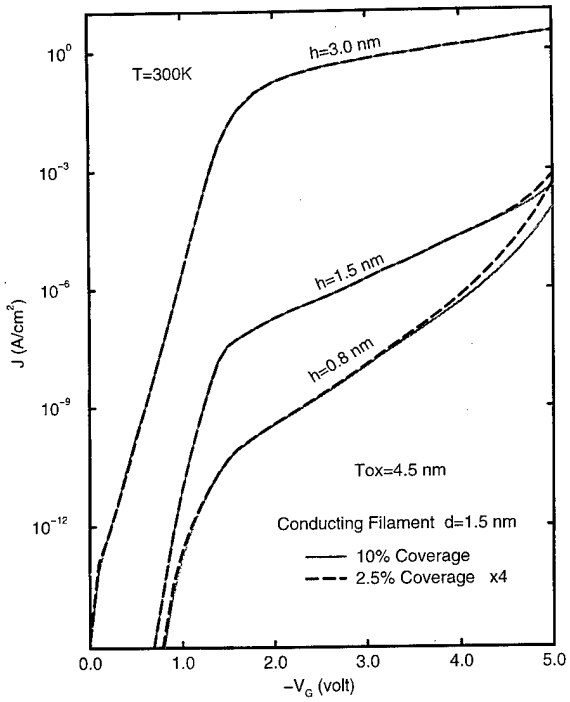


FIG. 4. Calculated current density–voltage curves at 300 K for two sets of MOS tunnel structures with oxide-embedded cylindrical conducting filaments. Both sets of structures have an oxide thickness of 4.5 nm, filament diameter of 1.55 nm, and cylinder heights of 0.8, 1.5, and 3.0 nm. The filaments in the first and second sets, respectively, cover 10% and 2.5% of the cross-sectional area. We have multiplied the second set of current densities by a factor of 4 for ease of comparison.

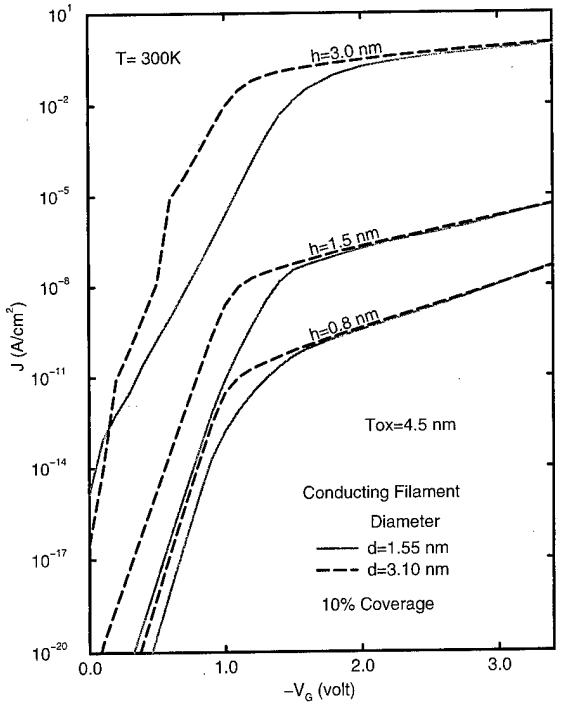


FIG. 5. Calculated current density–voltage curves at 300 K for two sets of MOS tunnel structures with oxide-embedded cylindrical conducting filaments. Oxide thickness is 4.5 nm, and cylinder heights are 0.8, 1.5, and 3.0 nm. In all cases, the cylinders cover 10% of the cross-sectional area, but have different diameters: 1.55 nm for the first set, and 3.10 nm for the second set.

provides the dominant conducting mechanism. In the Fowler–Nordheim regime, the scaling is dependent on filament length. In the case of long filaments ($h=3.0$ nm), the current density scales linearly with filament density well into the Fowler–Nordheim regime. But, in the cases of shorter filaments, the current densities do not scale up as rapidly. In fact, for the $h=0.8$ nm structure, at $|V_G|=5$ V, current densities are essentially independent of filament density, since, again, high-bias currents are primarily controlled by the leading edge of the barrier, which is essentially unaffected by short filaments.

In Fig. 5 we examine the dependence of current densities on filament diameter. We simulated structures with a filament diameter of $d=1.55$ nm using 32×32 planar supercells, and structures with $d=3.10$ nm using 64×64 supercells. In both cases, the filaments cover 10% of the cross-sectional area. In general, the larger diameter filaments result in higher current densities at low bias. But the difference between the current densities in the larger and smaller diameter structures diminishes with increasing gate bias. This is best understood by considering the electron deBroglie wavelength in the filaments. As the gate bias ($|V_G|$) increases, the energy of a typical injected electron increases relative to the band edge of the filament, and the deBroglie wavelength λ_e of the electron decreases correspondingly. At lower biases, λ_e might be comparable to the filament diameters. The transmission properties of the filaments then depend on the number of available modes in the cylindrical filaments, and this generally favors the larger diameter filaments. But at higher biases, λ_e is much shorter than the diameter sizes considered here, and the transmission properties then scale linearly with the cross-sectional area of the filaments.

IV. SUMMARY

We performed 3D quantum mechanical calculations to analyze the current–voltage characteristics of n^+ poly-$Si/SiO_2/p$-Si tunnel structures containing nonuniform ultrathin oxide layers. We find that, in general, roughness at the Si/SiO_2 interface renders the oxide layer more permeable. In the direct-tunneling regime interface roughness induces lateral localization of wave functions, which leads to preferential current paths, and is characterized by current densities which increase with island size. In the Fowler–Nordheim tunneling regime, however, interface roughness affects transport primarily through scattering, which increases with island size, manifesting in current densities which decrease with island size. We have also examined oxide-embedded conducting filaments, and found that they act as localized conduction paths and lead to dramatic increases in current densities. Depending on the filament length, our model can produce current–voltage characteristics reminiscent of those observed experimentally for ultrathin oxides having undergone either quasibreakdown or breakdown. We also found that the lower bias ($|V_G|<|V_{FB}|$) current densities in structures with long filaments are greatly enhanced by resonant tunneling through states identified as quantum dots, and that

this current enhancement is highly temperature dependent. We also report on the dependence of current–voltage characteristics on filament diameter size and filament density.

ACKNOWLEDGMENTS

The authors would like to thank O. J. Marsh, E. S. Daniel, and Z. Q. Zhang for helpful discussions, and M. A. Barton for technical assistance. This work was supported by the U.S. Office of Naval Research (ONR) under Grant No. N00014-89-J-1141, the U.S. Air Force Office of Scientific Research (AFOSR) under Grant No. F49620-96-1-0021, and by the ROC National Science Council under Grant No. NSC 87-2112-M-007-005.

[1]H. S. Momose, M. Ono, T. Yoshotomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, IEEE Trans. Electron Devices **43**, 1233 (1996).
[2]See J. G. Simmons, J. Appl. Phys. **34**, 1793 (1963), and references therein.
[3]See M. Lenzlinger and E. H. Snow, J. Appl. Phys. **40**, 278 (1969), and references therein.
[4]S. Nagano, M. Tsukiji, K. Ando, E. Hasegawa, and A. Ishitani, J. Appl. Phys. **75**, 3530 (1994).
[5]F. Rana, S. Tiwari, and D. A. Buchanan, Appl. Phys. Lett. **69**, 1104 (1996).
[6]S. T. Cundiff, W. H. Knox, F. H. Baumann, K. W. Evans-Lutterodt, M.-T. Tang, M. L. Green, and H. M. van Driel, Appl. Phys. Lett. **70**, 1414 (1997).
[7]M. Hirose, J. L. Alay, T. Yoshida, and S. Miyazaki, in *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface—3,* edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (The Electrochemical Society, Pennington, NJ, 1996), Vol. 96-1, p. 485.
[8]P. P. Apte and K. C. Saraswat, IEEE Trans. Electron Devices **49**, 1595 (1994).
[9]A. Halimaoui, O. Brière, and G. Ghibaudo, Microelectron. Eng. **36**, 157 (1997).
[10]D. Z.-Y. Ting, S. K. Kirby, and T. C. McGill, J. Vac. Sci. Technol. B **11**, 1738 (1993).
[11]D. J. Ben Daniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).
[12]R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).
[13]J. L. Alay and M. Hirose, J. Appl. Phys. **81**, 1606 (1997).
[14]D. Z.-Y. Ting and T. C. McGill (unpublished).

Equilibrium shape of $CoSi_2$ hut clusters on Si(100)

S. H. Brongersma[a)]
Department of Physics and Astronomy, University of Western Ontario, London, Ontario N6A 3K7, Canada

M. R. Castell and D. D. Perovic
Department of Metallurgy and Materials Science, University of Toronto, Toronto, Ontario M5S 3E4, Canada

M. Zinke-Allmang
Department of Physics and Astronomy, University of Western Ontario, London, Ontario N6A 3K7, Canada

(Received 21 January 1998; accepted 26 May 1998)

$CoSi_2$ clusters of constant height on a Si(100) surface grow in a square shape at first, but at a critical size a shape transition to clusters with large aspect ratios occurs. With each cluster connected to an implanted layer of cobalt by a thermally induced defect that serves as a diffusion channel, clusters can grow independently with a continuous supply of cobalt. When the cobalt supply is limited, clusters grow up to a specific volume and then have time to adjust their height, assuming their minimum-energy shape. Although calculations indicate that this should be a square pyramid, experiments indicate that a more elongated cluster corresponds to equilibrium. © *1998 American Vacuum Society.* [S0734-211X(98)10404-3]

I. INTRODUCTION

Mo *et al.*[1] were the first to observe rectangular Ge hut clusters on Si during the initial stages of growth. Their results could be explained in terms of a strained layer that releases tension by creating a sawtooth surface with very flat structures (aspect ratio 30 Å height; 1500 Å length). Similar elongated structures were observed with Ag on Si(100) by Hembree and Venables[2] and with Au_4Si on Si by Sekar *et al.*,[3] and three dimensional $CoSi_2$ hut clusters were recently imaged by Scheuch *et al.*[4] The latter formed when codepositing thin layers of cobalt and silicon. Tersoff and Tromp[5] argued that these elongated structures are the result of a strain induced shape transition, due to a trade-off between surface and interfacial energies on the one hand, and stress relaxation in the three dimensional islands due to an elastic distortion of the substrate on the other. Clusters of constant height h are predicted to be square up to a certain critical size, but when the cluster volume V increases further, there is a shape transition from a square to a rectangular shape with width s and length t. Interestingly, the aspect ratio $a=t/s$ increases not merely due to an increase in t, but also a decrease in s, indicating that the cluster is actively minimizing its energy. Recently, we were able to confirm this prediction quantitatively for $CoSi_2$ clusters on Si(100).[6] Using a novel growth technique we obtained a large number of clusters of all ages, while at the same time maintaining a dilute morphology of noninteracting clusters. The data showed that all the clusters follow an identical growth path, well described by the theory of Tersoff and Tromp.[5,6] However, the calculations depend critically on the condition of constant h during elongation. Tersoff and LeGoues[7] as well as Khor and Das Sarma[8] have indicated that, when h becomes a variable (and the cluster energy is minimized with respect to V instead of the substrate–cluster interfacial area A), the theoretical optimum shape is always a square pyramid. Nevertheless, experiments show a slightly elongated cluster to be the minimum-energy configuration.[1–6] The present analysis of the clusters is meant to increase our understanding of this persistent disagreement with the calculations.

II. EXPERIMENT

A cobalt implant of 9×10^{15} atoms/cm^2 at 450 keV into commercially available n-doped Si(100) wafers ($T_{Si}=375$ °C) was performed in the implant stage of a General Ionex 1.7 MV tandem accelerator. This resulted in a Gaussian implant profile centered around 330 nm. Subsequent annealing of the implanted Si wafers was done in the controlled environment of a molecular beam epitaxy chamber (base pressure 3×10^{-10} Torr). During an anneal of 150 min at 950 °C the pressure did not exceed 10^{-9} Torr.

As discussed previously,[6] the anneal results in the forming of thermal etching pits (TEPs)[9] that provide a highly efficient diffusion channel for the implanted Co back to the surface, resulting in the formation $CoSi_2$ clusters. This new technique helps to overcome a problem that is encountered when using continuous deposition from a vapor phase or molecular beam when attempting to increase cluster volume. These latter techniques not only induce an increase of the size of the clusters in a very dilute initial cluster morphology, but also cause new clusters to be created. As a result intercluster distances become small and elongation is interfered with by cluster collisions and cluster interaction through adatoms on the surface. In the present case clusters only form at TEP sites and obtain Co from the implanted subsurface reservoir independently.

While previous experiments used high dose implants (6×10^{16} atoms/cm^2) to ensure a continuous Co supply throughout the anneal, here the dose is reduced considerably. This causes the implanted layer to be depleted within the

a)Electronic mail: sbronger@julian.uwo.ca

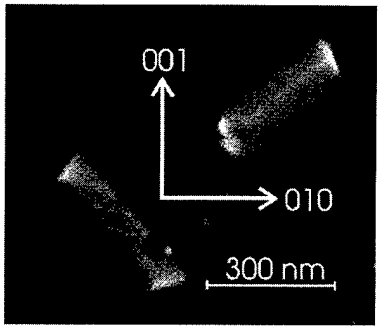


FIG. 1. Two $CoSi_2$ hut clusters on Si(100) as imaged using a FE-SEM. All clusters on the surface grow along the two distinct directions shown here.

anneal time (increased by factor of 2.5), leaving ample room for the cluster to assume its equilibrium shape.

A Hitachi S-4500 field emission-scanning electron microscope (FE-SEM) with a resolution of about 3 nm at 1 kV is used to obtain images of both the surface and the cluster's cross section.

In Fig. 1 two $CoSi_2$ hut clusters on Si(100) can be seen to be aligned along two different crystallographic orientations. In fact, all clusters on Si(100) grow along these two directions. Contrary to theoretical calculations[7,8] that predict a square pyramid when the cluster height is a variable, we observe a triangular cross section in one direction only. As consecutive annealing does not change the average cluster aspect ratio, we must accept the rectangular shape to be the equilibrium shape of the clusters. It corresponds to a width $s=109\pm10$ and length $t=285\pm62$. Thus we obtain an aspect ratio of $a=2.6$ and the predicted square pyramid is apparently not the minimum-energy shape, as also observed for several other systems.[1–4] The variance in s and t is most likely due to variations in cluster volume as the amount of extracted Co may not be identical for each cluster.

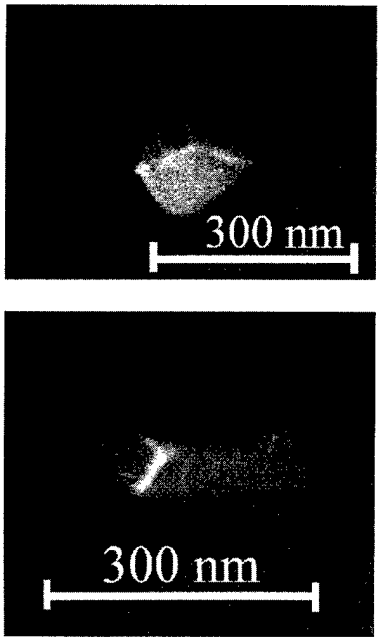


FIG. 2. Two cross sections of a $CoSi_2$ cluster. The top image shows a V-like shape with interfaces along the (111), (100), and (311) planes. The bottom image shows (111) interfaces coming down to the horizontal lower end of the V shape.

Figure 2 shows the cross section of a similar cluster along both its width (top) and its length (bottom). Thus we note that the cluster not only increases its height to obtain the optimum triangular shape, but also grows into the substrate causing its cross-sectional area to increase even more. Both this, and the orientational preference, are related to the $CoSi_2$–Si interfacial energy along various planes. Adams *et al.*[10] who analyzed the equilibrium shapes of isolated silicide precipitates buried by a Si capping layer, found the (111) $CoSi_2$–Si interface to have the lowest energy with $\gamma(100)/\gamma(111)=1.43\pm0.07$. Other interfaces, such as the (211) and (311), were also found to be stable during high temperature anneals, but their occurrence was lower. Thus, instead of lying on top of the substrate [along a (100) plane], the cluster of volume V can lower its energy by reducing its length and increasing its cross section by forming (111) planes at 53° with the substrate surface. The cluster shape underneath the surface then becomes an upsidedown rooftop, while above the surface the previously observed[6] normal rooftop is formed but with different limiting planes.

Additionally we note that a V shape of $(n11)$ planes (where n is an integer and preferably unity) can only be formed for growth along the (011) or $(0\bar{1}1)$ direction. This clarifies the preference for growth along these two crystallographically equivalent directions.

III. DISCUSSION

We now revisit the model of Tersoff and Tromp[5] for the energy of a strained epitaxial island, which so neatly described the observed cluster growth and shape transition in Ref. 6. It balances the surface/interfacial energies (E_s) due to cluster formation with an energy change due to cluster relaxation (E_r) obtained by an elastic distortion of the substrate. Now that the equilibrium shape of the clusters is known we can recalculate the change in surface/interfacial energy E_s. By setting $h=(s/2)-\tan\theta$ we find

$$E_s=st\left(\frac{\gamma_{\text{int}}}{\cos\phi}+\frac{\gamma_{\text{hut}}}{\cos\theta}-\gamma_s\right), \tag{1}$$

where γ_{int}, γ_{hut}, and γ_s are the energies per unit area of the $CoSi_2$–Si, the cluster–vacuum, and the substrate–vacuum interfaces, respectively. We take the cluster's contact angle θ and $CoSi_2$–Si interface angle ϕ to be constants. The simplicity of this result is a direct consequence of the surprisingly simple expression for the surface area of the cluster top given by $A=s-t/\cos\theta$. The volume of the cluster is given by

$$V=\frac{s^2}{4}\left(t-\frac{s}{3}\right)(\tan\theta+\tan\phi). \tag{2}$$

From simple geometric considerations the minimum E_s per unit volume is obtained for an aspect ratio $a=1$, repre-

senting a *square* pyramid. In this part of the total energy equation there clearly is no incentive for a rectangular shape to be at equilibrium.

The balancing energy gain term, which lowers the energy of the island at the cost of some strain in the substrate, was given as[5]

$$E_t = -2ch^2\left[s\ \ln\left(\frac{te^{3/2}}{h\ \cot\ \theta}\right) + t\ \ln\left(\frac{se^{3/2}}{h\ \cot\ \theta}\right)\right]. \tag{3}$$

The constant c depends on the cluster bulk stress, the Poisson ratio, and the shear modulus of the substrate. Tersoff and LeGoues[7] then showed that a minimization of the total energy leads to $s=t=h\ \cos\ \theta$, which is no surprise when considering the symmetry of Eq. (3). This result does not change when using, in our case the more accurate, Eq. (2) for the cluster volume. However, in the derivation of this equation it was assumed that the strain ϵ within the island does not vary along the z direction. This is an excellent approximation when both $s \gg h$ and $t \gg h$, and can still be quite accurate for small angles of θ and ϕ. In a square pyramid extending into the substrate, however, the stress release will be much smaller close to the tip, as the elastic distortion of the substrate cannot be maintained. Therefore it becomes favorable to change the aspect ratio so that stress release can be maintained in at least one direction.

A more mathematical description of the stress in a sharp groove was given by Yang and Srolovitz[11] who indicate that stress and strain fields indeed change dramatically near the groove tip. When approximating the groove as an ellipse with depth d and half-width b, they describe the stress field by $\sigma^{\infty}\cdot(1+2d/b)$ where σ^{∞} is the applied tensile stress parallel to the surface.

IV. SUMMARY

In conclusion, when given time after a constant cluster volume is reached, clusters assume their equilibrium rooftop shape. In addition, the $CoSi_2$ hut clusters are not lying on the surface, but extend into the substrate as well. The energy minimum does not correspond to the previously predicted square pyramid, but rather to a slightly elongated shape. This results from a reduction of stress release near the sharp tip of a square pyramid extending into the substrate.

ACKNOWLEDGMENTS

The authors acknowledge financial support by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Ontario Centre for Materials Research (OCMR).

[1]Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Phys. Rev. Lett. **65**, 1020 (1990).
[2]G. G. Hembree and J. A. Venables, Ultramicroscopy **47**, 109 (1992).
[3]K. Sekar, G. Kuri, P. V. Satyam, B. Sundaravel, D. P. Mahapatra, and B. N. Dev, Phys. Rev. B **51**, 14 330 (1995).
[4]V. Scheuch, B. Voiguänder, and H. P. Bonzel, Surf. Sci. **372**, 71 (1997).
[5]J. Tersoff and R. M. Tromp, Phys. Rev. Lett. **70**, 2782 (1993).
[6]S. H. Brongersma, M. R. Castell, D. D. Perovic, and M. Zinke-Allmang, Phys. Rev. Lett. **80**, 3795 (1998).
[7]J. Tersoff and F. K. LeGoucs, Phys. Rev. Lett. **72**, 3570 (1994).
[8]K. E. Khor and S. Das Sarma, Phys. Rev. B **49**, 13657 (1994).
[9]F. J. Gómez, M. L. Rodriguez-Méndez, J. Piqueras, J. Jiménez, and J. A. De Saja, Inst. Phys. Conf. Ser. **149**, 67 (1995).
[10]D. P. Adams, S. M. Yalisove, and D. J. Eaglesham, J. Appl. Phys. **76**, 5190 (1994).
[11]W. H. Yang and D. J. Srolovitz, Phys. Rev. Lett. **71**, 1593 (1993).

Heterointerface dipoles: Applications to (a) $Si-SiO_2$, (b) nitrided $Si-N-SiO_2$, and (c) $SiC-SiO_2$ interfaces

G. Lucovsky[a)]
Department of Physics, Department of Materials Science and Engineering, and Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

H. Yang
Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695-8204

H. Z. Massoud
Department of Electrical Engineering, Duke University, Durham, North Carolina 27708-0291

(Received 21 January 1998; accepted 21 May 1998)

This article identifies procedures to calculate charge-transfer dipoles at semiconductor–dielectric interfaces, focusing primarily on the $Si-SiO_2$ system. Since SiO_2 is more polar than Si, there is a transfer of electrons from Si to SiO_2 to balance the difference in chemical potentials creating a dipole localized at the semiconductor–dielectric interface. This dipole increases the conduction-band offset energy difference between Si and SiO_2, and therefore, has important effects on interface electrical and optical properties. Empirical chemistry and *ab initio* methods have been applied to molecular clusters which emulate the interface bonding, and have been used to calculate interfacial charge transfer at (i) abrupt $Si-SiO_2$ interfaces and (ii) nitrided $Si-N-SiO_2$ interfaces. Additional calculations have applied to determine the average dipoles at $Si-SiO_2$ interfaces with suboxide bonding in excess of the monolayer level required to form an ideal interface. The calculations support experimental data that indicate that the effective conduction-band offset energies at nitrided $Si-N-SiO_2$ interfaces and at interfaces with minimized suboxide bonding are essentially the same. Finally, the calculations have been extended to $SiC-SiO_2$ interfaces to illustrate the effects of changing from a homopolar to a more ionic or heteropolar semiconductor.
© *1998 American Vacuum Society.* [S0734-211X(98)07604-5]

I. INTRODUCTION

The energy difference (or offset) between the conduction bands of crystalline Si and the SiO_2 gate dielectric in metal–oxide–semiconductor (MOS) field-effect transistors (FETs) plays a significant role in determining many optical and electrical properties including the barrier heights for internal photoemission and electrical tunneling, and the threshold voltages which define the onset of current flow in *n*-channel FETs. Similar considerations also apply for the valence-band offset energies. As FET device dimensions shrink in the deep submicron (e.g., channel lengths <0.15 μm), the oxide equivalent thickness of gate dielectrics must be reduced proportionally to less than 3 nm, a thickness regime in which direct tunneling can play a significant role in the off-state leakage current.[1] Benefits of nitridation of ultrathin oxides have been identified, and include (i) decreased defect generation under accelerated stress bias testing, (ii) decreased Fowler–Nordheim and direct tunneling currents, and (iii) containment of boron penetration out of heavily doped p^+ polycrystalline gate electrodes.[2–4] The first two improvements are associated with interface nitridation at the monolayer level,[2,3] whereas the third is optimized by plasma-assisted top surface nitridation,[3] or the use of stacked oxide-nitride (ON) dual-layer dielectrics.[4] This article addresses one aspect of interface nitridation by comparing charge-transfer dipoles at nitrided and non-nitrided $Si-SiO_2$ interfaces.

The driving force for the charge-transfer process underlying the results of this article was identified more than ten years ago.[5–7] At that time, it was shown that equalization of the chemical potential across the interface layer was responsible for interfacial charge transfer. The chemical-potential equalization process was quantified using an empirical approach by applying the principal of electronegativity equalization as formulated by Sanderson.[8–10] The motivation for the analysis in Ref. 7 was to explain differences in the sign of the valence-band offset energies for $Si-SiO_2$ interfaces that had been hydrogenated or cesiated. The motivation in Refs. 8 and 9 was to explain differences between flatband voltages at $Si(111)-SiO_2$ and $Si(100)-SiO_2$ interfaces. The major contribution of this article is to go significantly beyond the *zeroth-order* model calculations of Refs. 7–9, and develop a more quantitative approach that could be applied to (i) nitrided $Si-N-SiO_2$ interfaces, (ii) $Si-SiO_2$ interfaces with suboxide transition regions, as well as (iii) other semiconductor dielectric interfaces such as $SiC-SiO_2$. In addition, the success of this model in addressing nitrided $Si-N-SiO_2$ and $Si-SiO_2$ interfaces with suboxide transition regions has provided an important perspective with respect to the different roles that interface nitridation and suboxide transitions regions play in strain relief at $Si-SiO_2$ interfaces.

[a)]Electronic mail: gerry_lucovsky@ncsu.edu

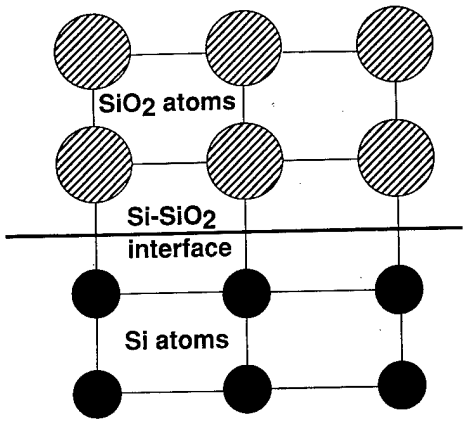


FIG. 1. Molecular model for abrupt interfaces between Si(111) and SiO_2 "atoms."

The approach of Refs. 8–10 is reviewed to provide a background for development of the calculations of this article. The Sanderson electronegativity approach uses an empirically based method for determining charge transfer between near-neighbor atoms with different electronegativities.[8–10] Figure 1 indicates the simplified structural model of Refs. 6 and 7 as it was first applied to abrupt Si(111)–SiO_2 interfaces. The calculation of the charge-transfer dipole was performed at a *molecular level* where the two *molecular constituents* were the Si atoms of the crystalline substrate and SiO_2 *pseudoatoms* of the oxide layer. The driving force for interfacial charge transfer is the electronegativity equalization principle of Refs. 8–10. If S_A is the electronegativity of one atom of a diatomic molecule AB, and S_B the electronegativity of the other atom, then charge is transferred between A and B until they reach the same electronegativity or chemical potential. A geometric mean is used to define electronegativity equalization, so that if S_A^* and S_B^* are the equalized atomic electronegativities after charge transfer has occurred, then $S_A^*=S_B^*=(S_A\times S_B)^{0.5}$. This approach has been extended to larger molecules in which the atoms have additional nearest, and/or more distant neighbors.[11] Once the values of S_A^* and S_B^* have been determined, partial charges on A and B, e_{pA}, and e_{pB}, respectively, can be calculated using an empirical relationship developed in Refs. 8–10,

$$e_{pA}=(S_A^*-S_A)/0.8\times S_A^{0.5}$$

and

$$e_{pB}=(S_B^*-S_B)/20.8\times S_B^{0.5}, \quad (1)$$

where the normalization constant in the denominator is based on an assumed ionicity of 75% for the NaF molecule. If A has lower electronegativity, the partial charge of A is positive and the partial charge of B is negative, indicating a direction of charge transfer consistent with the definition of electronegativity as first proposed by Pauling.[12]

Applying the diatomic molecule model to the Si–SiO_2 interface in Fig. 1, the renormalized electronegativities after chemical-potential equalization are 3.85 for both Si and SiO_2, yielding partial charges of $+0.18\ e$ for Si and $-0.18\ e$ for SiO_2. If a less "electronegative" dielectric such as Si_3N_4 is substituted for SiO_2, then the renormalized electronegativities are smaller, leading to reduced partial charges of $+0.11$ for Si and -0.11 for Si_3N_4. Applied to Si–SiO_2 interfaces in MOS structures, the molecular model of Refs. 6 and 7 has yielded charge-transfer dipoles that account for differences in flatband voltages reported for MOS devices on Si(100) and Si(111) substrates.[6,7] In Refs. 6 and 7 a *road map* was presented for extending the electronegativity equalization approach to more complicated interfaces such as those that include interfacial nitridation, and/or interfacial suboxide bonding. This article follows that *road map* and focuses on interfacial bonding issues that have emerged as being important at Si–SiO_2 interfaces with ultrathin oxide dielectrics and nitrided interfaces.

TABLE I. Calculated charges, dipole moments, and dipole potential steps for Si(111) interfaces.

Interface bonding	Partial charge on Si ±0.01 (e)	Dipole moment ±0.01 ($e\times$Å)	Potential step ±0.01 (eV)
(a) Abrupt Si–SiO_2			
Two-layer	0.18	0.29	0.54
Empirical cluster	0.23	0.37	0.67
Ab initio cluster	0.22	0.36	0.68
(b) Nitrided Si–N–SiO_2			
Two-layer	0.15	0.26	0.48
Empirical cluster	0.18	0.31	0.57
Ab initio cluster	0.18	0.36	0.65
(c) Si–SiO_x–SiO_2			
Two-layer	0.16	0.25	0.47
Empirical cluster	0.17	0.28	0.52
Ab initio cluster	0.18	0.29	0.54
(d) Abrupt Si–Si_3O_4			
Two-layer	0.11	0.20	0.30
Empirical cluster	0.14	0.24	0.36

II. CALCULATIONS OF CHARGE-TRANSFER INTERFACE DIPOLES

Charge-transfer dipoles are calculated in two different ways in this article: (i) by applying the empirical chemical method of Refs. 5–7 to larger clusters, which include the effects of more distant neighbors in a manner consistent with the approaches previously developed for large molecules[11,13] and disordered networks;[14,15] and (ii) by using hydrogen-terminated molecular clusters as discussed in Ref. 16. These results are then compared in Table I with the results calculated from the "pseudomolecule" approach of Refs. 6 and 7. In order to make these comparisons, it was necessary to extend the approach of Refs. 6 and 7 to include monolayer interface nitridation and suboxide transition regions. For the nitrided interface, the "pseudomolecule" that was used to replace SiO_2 in Fig. 1 was characterized by the average electronegativity of SiO_2 and Si_3N_4, 3.63. This gives a Si-atom partial charge of 0.15 e, as compared to 0.18 e for the Si–SiO_2 interface. Similarly, for the interface with a subox-

ide transition region, the "pseudomolecule" that was used to replace SiO_2 in Fig. 1 was characterized by the average electronegativity of SiO_2 and SiO, 3.85, yielding a Si-atom partial charge of 0.16 *e*. Since the Si–N bond length is about 7% larger than the Si–O, this calculation gives essentially equal interfacial dipole moments of about 0.25 *e* Å.

Consider next the empirical chemical approach. As noted above, the computational basis for including near-nearest- and more-remote-neighbor chemical induction effects in molecules was first developed by Carver and Gray.[11] This was subsequently extended by Lucovsky and co-workers to network amorphous solids.[14,15] Using the method of Ref. 14, the interface between crystalline Si oriented in a (111) direction, and noncrystalline SiO_2 is represented by the molecular cluster in Fig. 2(a). The Si side of the cluster consists of a three atomic shells (two are shown in the diagram): (i) the Si atom at the *interface* in shell (1), (ii) the three nearest-neighbor Si atoms in shell (2), and (iii) nine pseudoatoms in shell (3), which are characterized by the average electronegativity of the rest of the network. A similar construction is used for the SiO_2 network on the other side of the interface. In this case, the cluster contains atomic shells of alternating Si and O atoms, as well as a final shell that is comprised of pseudoatoms with the average electronegativity of the SiO_2 network ($S_{average}=[(S_{Si})(Si_O)^2]^{1/3}=4.26$). A similar cluster has been used for the *ab initio* calculations, except that the terminating species are hydrogen atoms rather than *pseudoatoms* with average properties. The cluster used for empirical calculations as applied to abrupt nitrided Si–N–SiO_2 interfaces is shown in Fig. 2(b). It is not possible to characterize suboxide bonding by a single representative cluster. In order to illustrate the effects of suboxide bonding it is, therefore, reasonable to average over two clusters that emulate an average SiO composition. This is done by combining the cluster in Fig. 2(a) for the idealized Si–SiO_2 interface with the cluster in Fig. 2(c), which effectively extends the Si substrate into oxide and increases the coordination of the interface Si atom of the substrate with oxygen atoms from one to three.

Total energy *ab initio* calculations were performed using a many-electron embedding theory that permits accurate computation of molecule–solid-surface interactions. Calculations were carried out at an *ab initio* configuration interaction (CI) level, in which all electron–electron interactions are explicitly calculated and there are no exchange approximations or empirical parameters. The details of the method are extensively discussed in Refs. 17–19. The *ab initio* calculations are performed by first obtaining self-consistent-field (SCF) solutions for the H-atom-terminated dielectric and Si substrate clusters. The occupied and virtual orbitals of the SCF solution are then transformed separately to obtain orbitals spatially localized within the bonds of the dielectric cluster and the Si substrate. This unitary transformation of orbitals, which is based upon exchange maximization with atomic valence orbitals, enhances convergence of the configuration interaction expansion.[17–19]

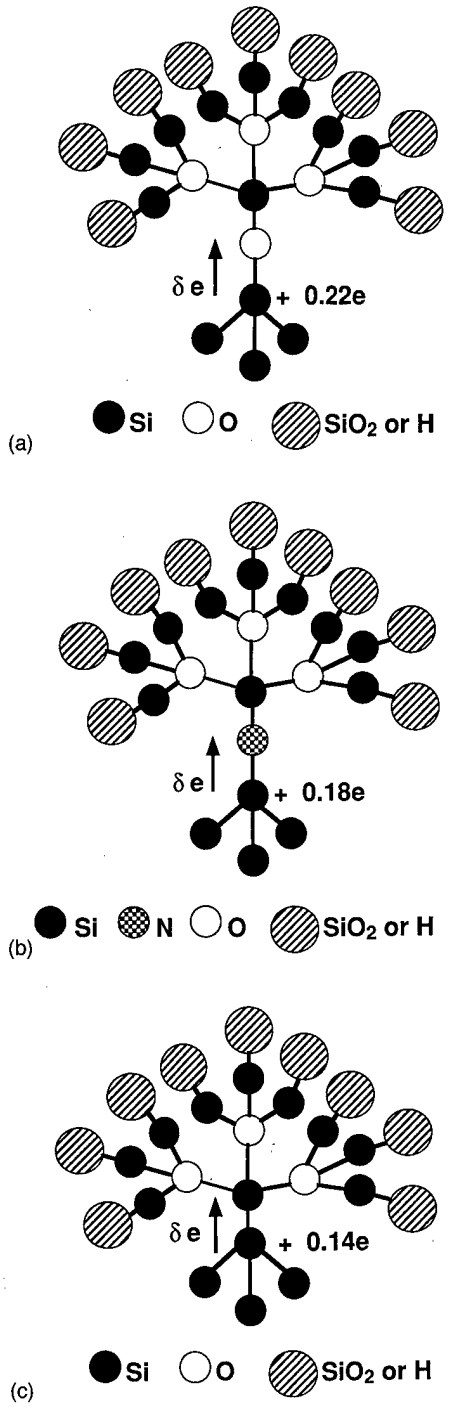


FIG. 2. Cluster model for Si–dielectric interfaces: (a) abrupt Si(111)–SiO_2, (b) abrupt nitrided Si(111)–N–SiO_2, and (c) Si(111)–SiO_2 with interfacial transition region with suboxide transition region. The dipole charges in this figure are from the *ab initio* calculations.

III. RESULTS

Table I includes the results of the three calculations described above. Presented in Table I are: (i) the partial charge on the substrate Si atom, (ii) the charge-transfer dipole, and (iii) the calculated potential step at the interface. The potential step ΔV was obtained using the approach of Refs. 6 and 7, i.e.,

$$\Delta V=[e_{p\mathrm{Si}}\times l_d\times N_{\mathrm{Si}}(111)\times\sin\Phi/(\epsilon_{\mathrm{AV}})], \quad (2)$$

where l_d is the atomic separation between the Si atom of the substrate and the first atomic shell of the dielectric, $N_{\mathrm{Si}}(111)$ is the areal density of Si atoms on a Si(111) surface, Φ the angle between the substrate Si-atom bond and the interface, and e_{AV} is an average dielectric constant, e.g., $e_{\mathrm{AV}}=(e_{\mathrm{Si}}+e_{\mathrm{SiO_2}})/2=7.75$.

Consider first the abrupt Si(111)–SiO_2 interface. The empirical two-layer molecular model of Refs. 6 and 7 gives values of $e_{p\mathrm{Si}}$ and ΔV that are smaller, respectively, than those obtained by the cluster calculations approaches. In addition, the value of ΔV is less than that given in Refs. 6 and 7 due to the assumption of different charge separation dipole distances. In the calculations of this article, a nominal Si–O bond length of 0.162 nm (1.62 A) was used, whereas the calculations of Refs. 6 and 7 employed a larger value that was estimated from the covalent bonding radius of Si and the "size" of a SiO_2 "molecule." More importantly, the empirical calculations and the *ab initio* calculations based on the cluster structures of this article [see Fig. 2(a)] give essentially the same results for the partial charge and the interface potential step at abrupt Si–SiO_2 interfaces. For the empirical calculation, the nominal dipole length of 0.162 nm was used in calculating ΔV, whereas for the *ab initio* approach, the ~3% larger relaxed Si–O bond length of the energy optimization calculation was used. Based on an electron energy representation of the interface band structure, the direction of the dipole increases the potential step at the conduction-band discontinuity between the Si(111) substrate and the SiO_2 conduction band. This would lead to negative shifts of the flatband and threshold voltages in MOS devices.[6,7]

Similar results were obtained using the cluster in Fig. 2(b) for an abrupt N-atom-terminated interface Si–N–SiO_2. As displayed in Table I, the *diatomic molecular model* of Refs. 6 and 7 gives values of $e_{p\mathrm{Si}}$ and ΔV that are again smaller than those obtained by the cluster calculations of this article. However, as the case of the Si–SiO_2 calculations, the empirical and *ab initio* calculations based on the cluster model of this article give essentially the same results for the partial charge and the interface potential steps. The difference between the nominal Si–N bond length of 0.174 nm, and the one obtained in the *ab initio* calculation accounts for the differences in ΔV between these calculations. The significant result of these calculations on nitrided Si–SiO_2 interfaces is that the partial charge on the Si atom of the substrate is reduced by the interposition of a N atoms at Si–SiO_2 interfaces, and that this leads to a small reduction of the interface potential step of approximately 0.1 eV with respect to non-nitrided abrupt Si–SiO_2 interfaces. To confirm the trends in $e_{p\mathrm{Si}}$ and ΔV for interface nitridation, calculations based on the *molecular model* of Refs. 6 and 7, and the empirical cluster model of this article were performed for a Si(111)–Si_3N_4 interface, and these have also been included in Table I. These calculations support the differences obtained for the abrupt Si(111)–SiO_2 and nitrided Si(111)–N–SiO_2 interfaces, since increased nitridation leads to smaller interfacial dipoles with correspondingly smaller contributions to the conduction-band offset energies.

Finally, Table I includes the results of calculations for Si(111)–SiO_x–SiO_2 interfaces with suboxide transitions regions. Different degrees of suboxide bonding have been approximated by placing a Si atom in the first shell of the dielectric and then increasing the number of Si atoms in the second shell bonded to that atom [see Fig. 2(c)]. As the number of Si atoms in the second shell increases, the partial charge on the Si atom of the substrate decreases, and the potential step contribution ΔV also decreases proportionally. However, the *ab initio* calculation for this type of interface was performed for the cluster with only one layer of suboxide bonding with no additional Si atoms attached to the Si atom of the first shell of the transition region. For purposes of comparison, the entries into Table I for the interface with suboxide bonding, Si(111)–SiO_x–SiO_2, are the average of the results obtained for the clusters in Figs. 2(a) and 2(c).

The charge transfer interface dipole has been calculated as a function of the interatomic distance of the interface atoms, i.e., the Si–O distance as in Fig. 2(a), Si–N distance as in Fig. 2(b), and Si–Si distance as in Fig. 2(c), and these results are presented in Figs. 3(a), 3(b), and 3(c). The results in Figs. 3(a) and 3(b) for the clusters of Figs. 2(a) and 2(b), respectively, demonstrate that the charge-transfer dipole decreases approximately linearly with increasing interatomic distance for the abrupt Si–SiO_2 and nitrided Si–N–SiO_2 interfaces. This dependence of the charge-transfer dipoles requires that the partial charge on the Si atom decrease approximately quadradically with increasing interatomic spacing for the range of the interfacial bond lengths explored. In contrast and as shown in Fig. 3(c), the variation of the charge-transfer dipole with interface bond length is reversed for the cluster that is used to emulate suboxide bonding so that the dipole increases with increasing Si–Si separation. Analysis of this dependence shows that it is a direct consequence of competition between two different contributions to the interface dipole. One contribution comes from a dipole between the Si substrate atom and the Si atom of the first shell of the dielectric that is directed toward the Si substrate, and the second contribution that is oppositely directed comes from the atoms of the second and third shells of the dielectric. The net effect of these two contributions is to give an effective interfacial dipole that is significantly smaller than that of an abrupt Si(111)–SiO_2 interface. This explanation is supported in Fig. 3(c) by an additional observation that shows that as the Si–Si distance is increased out of a linear regime the value of the interfacial charge-transfer dipole approaches the value of the dipole expected for a Si–SiO_2 bonding arrangement.

The calculations of this article have been extended to Si(100)–SiO_2 interfaces for the empirical cluster calculation method, and the results of the these calculations are presented in Table II. In all instances the partial charge on the Si(100)-substrate atom is increased, but the dipole, and hence, the potential step are decreased. The increase in partial charge is due to the increase in the number of bonding connections per Si interface atom (from one to two) to the

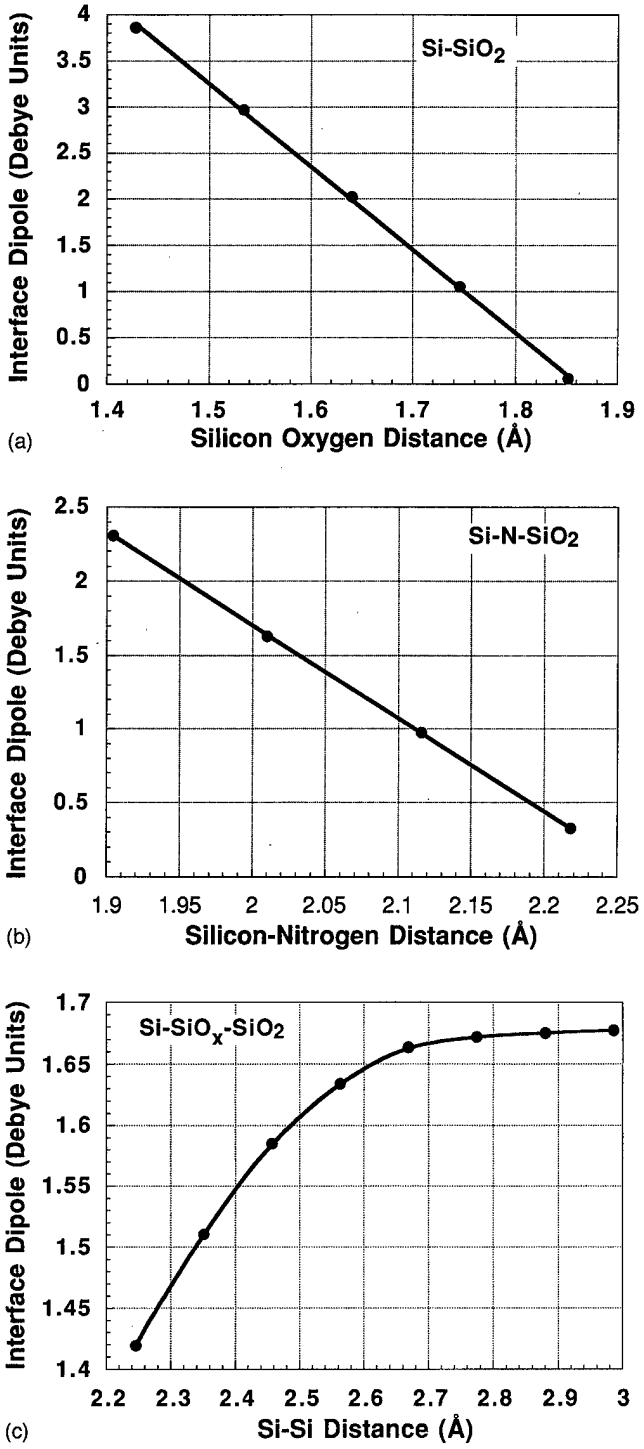


FIG. 3. Interface dipole as a function of interface bonding distance from the abate calculations: (a) abrupt Si(111)–SiO_2, (b) abrupt nitrided Si(111)–N–SiO_2, and (c) Si(111)–SiO_2, with interfacial transition region with suboxide transition. The solid lines in (a) and (b) are a linear fit to the calculated data points; the solid line in (c) is a smooth fit to the calculated data points.

dielectric, and the decrease in the potential step is due to the bonding geometry. In particular, the sin Φ factor is reduced from 1 to 0.577 as the projection angle of the dipole onto a normal to the plane of the interface decreases from 90° to 35.25°.

TABLE II. Calculated charges, dipole moments, and dipole potential steps for Si(100) interfaces.

Interface bonding	Partial charge on Si ±0.01 (e)	Dipole moment ±0.01 ($e\times$Å)	Potential step ±0.03 (eV)
(a) Abrupt Si–SiO_2			
Empirical cluster	0.32	0.28	0.46
(b) Nitrided Si–N–SiO_2			
Empirical cluster	0.26	0.26	0.41
(c) Si–SiO_x–SiO_2			
Empirical cluster	0.21	0.24	0.38

Table III presents the results of empirical cluster calculations for SiC(0001)–SiO_2 interfaces. The major difference between the calculated partial charges on the Si atom on the SiC(0001) substrate and on the Si(111) substrate are driven by the heteropolar character of the SiC semiconductor. The increased polar character of the SiC substrate reduces the partial charge on the Si surface atom relative that of the Si(111) substrate.

IV. DISCUSSION

This section addresses several different aspects of the calculations of Sec. III including (i) development of a band model for the Si–SiO_2 interface that takes into account differences due to interface nitridation, and interfacial suboxide regions; (ii) the comparison of these calculations with experimental results for Si–SiO_2 interfaces; and (iii) a perspective for semiconductor dielectric interfaces that parallels a framework established for lattice-mismatched semiconductor–semiconductor heterointerfaces.

A. Band model for Si–SiO_2 interfaces

Figure 4 presents a band model for the conduction- and valence-band offset energies between Si and SiO_2. As shown in the diagram, there are two contributions to the conduction-band (and therefore, also the valence-band) potential steps: (i) the first comes from differences between the potential steps of the respective interface components relative to

TABLE III. Comparisons of partial charges on Si interface atoms between Si(111) and Si-atom terminated SiC(0001) interfaces.

	Partial charge on Si	
Interface bonding	Si(111) ±0.01 (e)	SiC(0001) ±0.01 (e)
(a) Abrupt Si–SiO_2		
Two-layer	0.18	0.13
Empirical cluster	0.23	0.17
(b) Nitrided Si–N–SiO_2		
Two-layer	0.15	0.10
Empirical cluster	0.18	0.12
(c) Si–Si_3N_4		
Two-layer	0.11	0.06
Empirical cluster	0.14	0.08

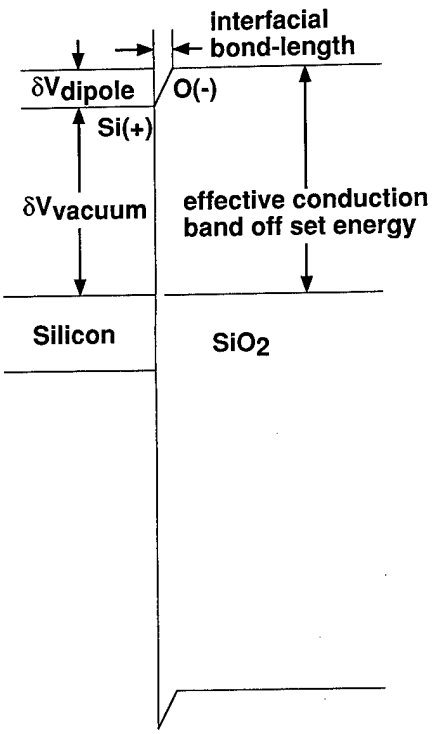


FIG. 4. Band model for Si–SiO_2 interfaces show interface potential due to charge-transfer dipoles.

vacuum, and (ii) the second comes from the interfacial dipoles of this article. DiMaria and co-workers performed internal photoemission measurements to determine the conduction-band potential steps between Si and SiO_2 (Ref. 20) and Si and Si_3N_4,[21] and these values have been used extensively in characterizing interface properties, as for example, in calculations of direct tunneling currents.[22] The results of this article indicate that care must be taken in applying these nominal conduction-band offset potential-energy steps. For example, the calculations of this article have demonstrated significant differences in effective conduction-band offset energies that depend on the local atomic bonding at the interface, including both interface nitridation, and transition regions with suboxide bonding.

B. Comparison with experiments

The calculations of this article are now compared with results of several different experimental studies. First, a word of caution regarding such comparisons. There are two factors involved in interface formation that contribute to differences in interfacial bonding arrangements of the types discussed in this article: (i) the process, or process steps by which the interface is formed, as for example, thermal oxidation, plasma-assisted oxidation, plasma-assisted oxidation combined with post-oxidation nitridation, etc., and (ii) the chemical and structural interfacial relaxations associated with post-formation rapid thermal annealing (RTA), or downstream thermal exposures. The combination of these two interface fabrication effects is illustrated by results obtained from optical second-harmonic-generation (SHG) experiments on vicinal Si(111)–SiO_2 interfaces,[23–25] which are presented in Table IV.

TABLE IV. Summary of results from studies of vicinal Si(111) surfaces off-cut ~5° in the 112 bar direction.

(a) Surface treatment—plasma processing[a]	Phase: $\Delta\Phi_{13}$ ±2°	A_1/A_3 ±0.03
O_2, 15 s, 300 °C	68	0.20
O_2, 15 s, 30 s 900 °C RTA ($0.5\%O_2$/Ar)[b]	23	0.35
N_2O, 15 s, 300 °C	67	0.21
N_2O, 15 s, 30 s 900 °C RTA (Ar)	11	0.37
N_2O, 30 s, 300 °C	65	0.17
N_2O, 30 s, 30 s 900 °C RTA (Ar)[b]	11	0.35
(b) Surface treatment—thermally grown interfaces		
Furnace oxidation at 850 °C	72	0.19
Postoxidation anneal (30 s at 950 °C ($0.5\%O_2$/Ar)	23	0.33

[a]Predeposition remote plasma-assisted oxidation step.
[b]Processing conditions for optimum electrical properties.

The experimental approach and the data reduction procedures for the optical SHG studies are described in detail in Refs. 23 and 24. The columns in Table IV labeled A_1/A_3 and $\Delta\Phi_{13}$, which are obtained from a harmonic analysis of angular anisotropy measurements in vicinal Si(111) surfaces and reflect differences in interface bonding as a function of the two processing factors identified above. The SHG response is dominated by the Si–SiO_2 interface for the polarizations of the incident and reflected optical waves at ω and 2ω, respectively. Combined with the results presented in Ref. 26, which demonstrate significant reductions in interfacial transition regions following a 30 s 900 °C RTA in a chemically inert ambient such as Ar or He, the SHG results in Table IV demonstrate that interfaces formed by thermal oxidation at temperatures below 850 °C, and remote plasma-assisted oxidation/nitridation at 300 °C, display the same values of A_1/A_3 and $\Delta\Phi_{13}$, and are essentially the same with suboxide transition regions being the determinant factor in determining interface properties. After the anneal, the non-nitrided interfaces formed by either thermal or remote plasma-assisted oxidation are essentially the same, but are different from the interfaces that contain approximately 1 monolayer (ML) of interface nitridation. X-ray photoelectron spectroscopy (XPS) studies on non-nitrided interfaces have established that (i) there is a reduction of suboxide bonding following the 30 s 900 °C RTA for non-nitrided interfaces formed by both thermal and plasma-assisted oxidation, and (ii) following the RTA, there is still measurable suboxide bonding ~1.5 excess monolayer equivalents of suboxide bonded Si atoms, as opposed to approximately 3 ML equivalents before RTA.[27]

In the examples that are discussed below, the Si–SiO_2 interfaces were subjected to a 30 s 900 °C RTA so that comparisons between the calculations and experiment will utilize the results of the *ab initio* calculations for the abrupt nitrided Si–SiO_2, and the interface with the smallest amount of sub-

oxide bonding. The interfaces characterized by optical SHG have been prepared on Si(111) surfaces, where as those discussed below for MOS devices were formed on Si(100) substrates. It was established earlier in Tables I and II that qualitative trends in interface potential steps as a function of interface bonding are the same for Si(111) and Si(100) interfaces. The differences are in the quantitative aspects of the calculations, namely, the magnitudes of the ΔVs, and differences in these values for different interface bonding on Si(111) and Si(100) substrates are small and within the experimental uncertainties of the experimental data.

It is tempting to compare the results of these calculations with barrier heights used to interpret tunneling studies on MOS devices with ultrathin gate dielectrics; however, the extraction of effective barrier heights from tunneling data is not unambiguous. For example, it depends on assumptions made about the magnitude of the effective mass of the tunneling electrons, and approximations made taking into account potential drops in the substrate and polycrystalline gate electrodes.[22] The results presented in Ref. 28 demonstrate that tunneling current densities are markedly different for nitrided and non-nitrided interfaces with nitrided interfaces showing reductions in tunneling current that are independent of the direction of the injection, substrate, or gate, and that are also quantitatively the same for tunneling in the Fowler–Nordheim (5 nm) and direct tunneling (2 and 3 nm) regimes. However, since the flatband voltages vary by no more than 0.05 eV, the reductions in tunneling current are not due to deferences in barrier heights at the substrate–Si/dielectric layer interface. Since the samples in Ref. 28 were subjected to postdeposition annealing in inert ambients at 900 °C, this leads us to conclude that the effective barrier heights at structurally and chemical relaxed, nitrided, and non-nitrided interfaces are essentially the same. This suggests that the effective barrier height at $Si–SiO_2$ interfaces with minimal suboxide bonding is essentially the same are for a fully nitrided interface. This result is consistent with the calculations presented in Tables I and II. The agreement between the predictions of the calculation and the tunneling experiments validates the average bonding model used for the interfaces with minimal suboxide bonding.

C. Stress relief at $Si–SiO_2$ interfaces

The results presented in this article, and in particular, the good agreement between the calculations and the tunneling experiments provide an important insight into the issue of the different roles of suboxide bonding and interface nitridation in stress relief at $Si–SiO_2$ interfaces. There is a large molar volume mismatch between Si and SiO_2 that means that the effective spacing of Si atoms in these two materials is markedly different. The Si–Si bond length in crystalline Si is 0.235 nm (2.35 A), whereas the second-nearest-neighbor Si–Si distance in SiO_2 is 0.305 ± 0.005 nm (3.05 ± 0.05 Å). The large mismatch in interatomic distances means that the $Si–SiO_2$ interface is in many respects an analog of lattice-mismatched semiconductor heterointerfaces. The formation of non-lattice-matched semiconductor heterointerfaces is accomplished in several different ways: (i) by pseudomorphic growth habits, where there are large strains in the epitaxial overgrowth layer and wherein the maximum thickness of the pseudomorphic overgrowth is inversely proportional to the strain; (ii) by the formation of interface defects in the form of dislocations; and (iii) by reduction of interface strain through the use of buffer layers with intermediate values of lattice constants.

The formation of non-nitrided $Si–SiO_2$ interfaces requires a transition region with suboxide bonding to reduce the very large strain mismatch between the Si and the SiO_2. From the XPS studies, it can be concluded that the minimum suboxide bonding region after postoxidation annealing at 900 °C contains about 1.5 ML of Si.[27] The incorporation of interfacial nitrogen at the monolayer level is an analog of the buffer layer approach, where a material with a different composition can accommodate a significant fraction of the strain mismatch and thereby reduce the strain and require reduced suboxide bonding.

Strain is developed differently at interfaces formed in different ways. For example, it has been shown that $Si–SiO_2$ interfaces with relatively thick oxide layers (>50 nm) formed by conventional thermal oxidation in dry O_2, display very nearly uniform levels of compressive stress in the oxide regions, thereby being analogous to pseudomorphic semiconductor–semiconductor heterostructures. Additional experiments have demonstrated strain relief in these oxides at annealing temperatures greater than or equal to about 1000 °C. After such anneals, there is a considerable strain gradient in the oxide layers extending to distances of approximately 20 nm. Finally, strain is pinned at the $Si–SiO_2$ interface at a level that corresponds to ~0.002–0.003.[29] Recent experiments on significantly thinner oxides, <50 nm, have identified transition regions and interfaces formed in this way, and additionally at interfaces formed by remote plasma-assisted oxidation.[26] This leads us to conclude that one class of optimally relaxed interfaces contain a relatively thin transition region between the Si substrate and SiO_2 layer in which the excess concentration of Si atoms with suboxide bonding is of the order of 1.5 ML. Oxides under about 100 nm do not appear to support significantly strain, and interface chemical and mechanical strain appear to be completely relaxed by the 900 °C RTAs discussed above. Finally, a monolayer of interfacial N atoms at $Si–SiO_2$ interfaces provides an alternative pathway for interface relaxation paralleling the introduction of graded strain relief layers in semiconductor–semiconductor heterostructures. As discussed in other publications, the preferential bonding of nitrogen atoms at $Si–SiO_2$ interfaces is determined by three factors.[30] The Si–N bond energy is estimated to be about 0.3 eV smaller than the Si–O bond energy. However, the mechanical and chemical strain produced by nitrogen atom bonding to Si atoms are both less than the corresponding Si–O strain energies. The preferential bonding of nitrogen means that the decrease in these strain energies more than compensates for the difference in bond energies.

V. SUMMARY

In summary, this article has demonstrated that interfacial transition regions at non-nitrided $Si-SiO_2$ interfaces reduce the partial charge of Si atoms on the *Si side* of that interface, thereby reducing the interfacial charge-transfer dipole below the value it would have for an abrupt $Si-SiO_2$ interface with no suboxide bonding beyond what is required to define the interface. Monolayer concentrations of nitrogen atoms incorporated at $Si-SiO_2$ interfaces have a similar effect in the context that they also reduce the partial charge on the Si atoms on the Si side of the interface below the value it would have had at an abrupt $Si-SiO_2$ interface. The model calculations based on empirical chemistry and *ab initio* approaches applied to the same interface cluster models, combined with the two-layer molecular approach indicate the conduction-band offset between Si and SiO_2 is only slightly increased by monolayer nitrogen-atom incorporation as compared to minimal suboxide bonding. For Si(111) interfaces, the calculated average difference in interfacial conduction-band offsets, $\delta\Delta V$, is estimated to be 0.05 ± 0.03 eV. Similar differences are obtained for Si(100) interfaces using the empirical cluster model. The results of recent experiments on tunneling currents in fully nitrided $Si(100)-SiO_2$ interfaces are in accord with the predictions of these calculations. The reduction in tunneling currents for nitrided interfaces reported in Ref. 29 must then be attributed to other factors including, for example, reductions in interface roughness that accompany interfacial nitridation.

ACKNOWLEDGMENTS

This research is supported by the Office of Naval Research (ONR), the National Science Foundation (NSF) Engineering Research Center (ERC) at North Carolina State University (NCSU), and the Semiconductor Research Corporation (SRC).

[1]D. Buchanan, J. Vac. Sci. Technol. B (to be published).
[2]D. R. Lee, G. Lucovsky, M. R. Denker, and C. Magee, J. Vac. Sci. Technol. A **13**, 1671 (1995); D. R. Lee, C. R. Parker, J. R. Hauser, and G. Lucovsky, J. Vac. Sci. Technol. B **13**, 1778 (1995).
[3]S. V. Hattangady *et al.*, Tech. Dig. Int. Electron Devices Meet., 495 (1996).
[4]Y. Wu, G. Lucovsky, and H. Z. Massoud, in Proc. IEEE Reliability Phys. Symp. (1998).
[5]P. Perfetti, C. Quaresima, C. Coluzza, C. Fortunato, and G. Margaritondo, Phys. Rev. Lett. **57**, 2065 (1986).
[6]H. Z. Massoud, Mater. Res. Soc. Symp. Proc. **105**, 265 (1988).
[7]H. Z. Massoud, J. Appl. Phys. **63**, 2000 (1988).
[8]R. T. Sanderson, *Chemical Periodicity* (Van Nostrand Reinhold, Princeton, NJ, 1960).
[9]R. T. Sanderson, *Inorganic Chemistry* (Reinhold, New York, 1967).
[10]R. T. Sanderson, *Chemical Bonds and Bond Energy*, 2nd ed. (Academic, New York, 1976).
[11]J. C. Carver, R. C. Gray, and D. M. Hercules, J. Am. Chem. Soc. **96**, 6851 (1984).
[12]L. Pauling, *The Nature of the Chemical Bond*, 3rd. ed. (Cornell University Press, Ithaca, NY, 1960).
[13]G. Lucovsky, Solid State Commun. **29**, 571 (1979).
[14]D. V. Tsu, B. N. Davidson, and G. Lucovsky, Phys. Rev. B **40**, 1795 (1989).
[15]G. N. Parsons and G. Lucovsky, Phys. Rev. B **41**, 1664 (1990).
[16]Z. Jing, G. Lucovsky, and J. L. Whitten, J. Vac. Sci. Technol. B **13**, 1613 (1995).
[17]P. Cremaschi and J. L. Whitten, Theor. Chim. Acta **72**, 485 (1987).
[18]J. L. Whitten and H. Yang, Int. J. Quantum Chem., Quantum Chem. Symp. **29**, 41 (1995).
[19]J. L. Whitten and H. Yang, Surf. Sci. Rep. **24**, 55 (1996).
[20]D. J. DiMaria, J. Appl. Phys. **50**, 5826 (1979).
[21]D. J. DiMaria (private communication).
[22]H. Y. Yang, H. Niimi, and G. Lucovsky, J. Appl. Phys. **83**, 2327 (1998).
[23]C. H. Bjorkman, T. Yasuda, C. E. Shearon, Jr., U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, J. Vac. Sci. Technol. B **11**, 1521 (1993).
[24]C. H. Bjorkman, C. E. Shearon, Jr., Y. Ma, T. Yasuda, G. Lucovsky, U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, J. Vac. Sci. Technol. A **11**, 964 (1993).
[25]U. Emmerichs, C. Meyer, H. J. Bakker, F. Wolter, H. Kurz, G. Lucovsky, C. Bjorkman, T. Yasuda, Yi Ma, Z. Jing, and J. L. Whitten, J. Vac. Sci. Technol. B **12**, 2484 (1994).
[26]G. Lucovsky, A. Banerjee, B. Hinds, B. Claflin, K. Koh, and H. Yang, J. Vac. Sci. Technol. B **15**, 1074 (1997).
[27]J. Rowe (unpublished).
[28]G. Lucovsky, H. Niimi, Y. Wu, C. G. Parker, and J. R. Hauser, J. Vac. Sci. Technol. A **16**, 172 (1998).
[29]C. H. Bjorkman, J. T. Fitch, and G. Lucovsky, Appl. Phys. Lett. **56**, 1983 (1990).
[30]K. Koh, H. Niimi, G. Lucovsky, and M. L. Green, Jpn. J. Appl. Phys. **83**, 2327 (1998).

Multilayer oxidation of AlAs by thermal and electron beam induced decomposition of H_2O in ultrahigh vacuum

S. I. Yi,[a),b),c)] W. J. Mitchell,[b)] C.-H. Chung,[c)] E. L. Hu,[b)] and W. H. Weinberg[b),c)]
University of California, Santa Barbara, Santa Barbara, California 93106

(Received 21 January 1998; accepted 21 May 1998)

The oxidation of AlAs by the decomposition of water has been investigated in ultrahigh vacuum using high-resolution electron energy loss spectroscopy, Auger electron spectroscopy, and temperature-programmed desorption. Significant oxidation of the AlAs near-surface region was observed after a single adsorption/anneal cycle in which multilayers (>10 monolayers) of water, adsorbed at 100 K, were annealed to above room temperature. Sputter profiling shows that repeated cycles of multilayer water exposure at 100 K followed by annealing results in a surface oxide that is at least 20–30 Å thick. The extent of surface oxidation, as measured by Auger electron spectroscopy, is not affected by annealing to 700 K. However, at temperatures in excess of 800 K, diffusion of subsurface AlAs to the oxide surface was observed. Moreover, it was found that irradiating a partially oxidized AlAs surface with a 3 keV electron beam for sufficiently long times would significantly increase the extent of surface oxidation. © *1998 American Vacuum Society.* [S0734-211X(98)07904-9]

I. INTRODUCTION

The excellent properties of AlAs oxide films formed by wet oxidation have recently attracted much attention.[1–4] At 700–800 K, exposure of gas-phase H_2O to AlAs films gives rise to robust, high quality oxide layers that have been incorporated into III-V metal–oxide–semiconductor (MOS) devices[5–7] and used as apertures in vertical cavity lasers, confining both optical mode and drive current.[8–11]

The majority of current research is directed at understanding the structural, compositional, and electrical properties of the fully oxidized films that result after atmospheric pressure oxidation in tube furnaces.[1–11] The evolution of the oxide at a fundamental level has been addressed only recently.[12] It was found that molecularly adsorbed H_2O on AlAs(100) at a surface temperature of 100 K both dissociates and desorbs upon annealing to 150–200 K, by virtue of overlapping kinetics. The formation of aluminum hydroxide, aluminum oxide, and arsenic hydride results from the dissociation of H_2O. The arsenic hydride ultimately desorbs over a wide temperature range as arsine, disrupting the near-surface AlAs matrix and exposing subsurface Al to further oxidation. The aluminum hydroxide either rehydrogenates and/or disproportionates between 400 and 650 K to produce desorbing water, or it dissociates between 500 and 750 K to produce hydrogen which desorbs. The latter two processes also result in aluminum oxide.

In this work, the interaction of large multilayers [>10 monolayers (ML)] of water with the AlAs(100) surface is investigated using Auger electron spectroscopy (AES) and high-resolution electron energy loss spectroscopy (HREELS).

[a)]Author to whom correspondence should be addressed; electronic mail: sanginyi@engineering.ucsb.edu
[b)]Center for Quantized Electronic Structures (QUEST).
[c)]Department of Chemical Engineering.

II. EXPERIMENTAL METHODS

The samples used in this study were 50 Å thick AlAs (undoped) films, grown on GaAs(001) substrates by molecular beam epitaxy, and which were As capped to prevent oxidation by air upon exposure to the atmosphere. The complete AlAs(100) surface preparation procedure has been outlined in detail elsewhere.[12] The sample temperature was monitored indirectly by a type-K thermocouple attached to the heater housing upon which the sample is mounted. An estimate within ±50 K of the true sample temperature from this thermocouple reading has been made [using the known GaAs(001) phase transition temperatures and low-energy electron diffraction (LEED)] and is used in all the spectra presented here. Sample heating to 1000 K is achieved by radiative heating from small W filaments located within a heater housing behind the sample. When using liquid nitrogen as the coolant in the cryostat reservoir, it is possible to cool the sample to 100 K.

AES was used primarily to monitor the extent of AlAs oxidation. Low-energy (10–100 eV), mid-energy (450–550 eV) and high-energy (1150–1500 eV) spectra were recorded in the differential mode [$dN(E)/dE$ vs E] using an electron beam energy of 3 keV. The energy of a given Auger transition was taken to be the minimum of the differential peak. The low-energy data include the reduced As(M_4VV) transitions at 31 and 38 eV, the reduced Al($L_1L_{23}V$) and Al($L_{23}VV$) transitions at 42 and 64 eV, and the oxidized Al(L_{23})O(L_{23})O(L_1) and Al(L_{23})O(L_{23})O(L_{23}) interatomic transitions at 39 and 54 eV. All of these transitions are surface specific since they exhibit electron escape depths of less than 8 Å.[13] To enhance signal resolution in the low-energy spectra, a 1 V peak-to-peak modulation voltage was used. The mid-energy spectrum monitors the oxygen coverage directly via the O(KL_1L_{23}) transition at 510 eV.

The HREEL spectra were all recorded in the specular direction. The resolution of the spectrometer was approxi-

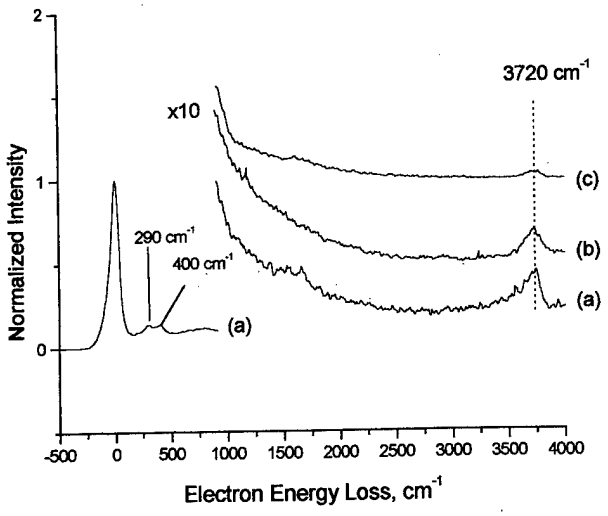


FIG. 1. HREEL spectra recorded after the indicated exposure of H_2O to AlAs(100) at 100 K followed by annealing to the indicated temperature. (a) 3 L exposure, 300 K anneal, (b) 37 L exposure, 300 K anneal, (c) 3 L exposure, 600 K anneal.

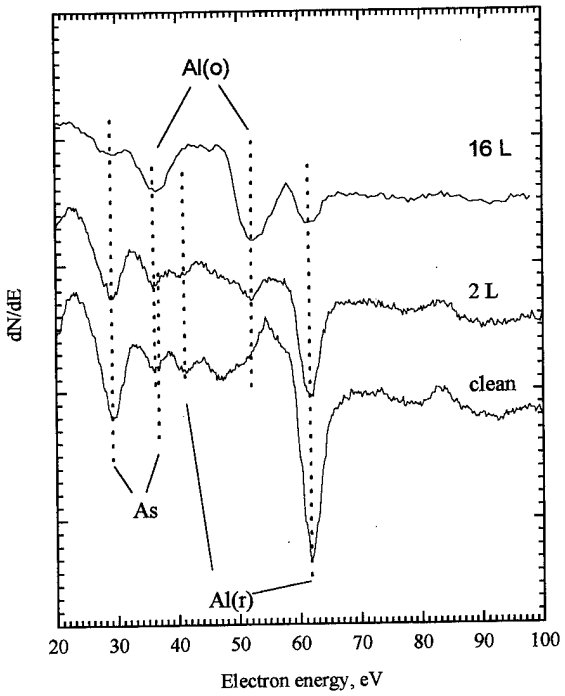


FIG. 2. Auger spectra recorded after exposing 2 and 16 L of H_2O to the AlAs(100) surface at 100 K followed by annealing to 300 K. Al(r) and Al(o) refers to the metallic Al and the oxidized Al^{3+} transitions, respectively.

mately 70 cm^{-1} (elastically scattered peak full width at half maximum) while maintaining a count rate of at least 10^5 Hz in the specular, elastically scattered beam from the clean AlAs(001) surface. Incident beam energies of 7.5 eV were used.

The H_2O was dosed through a directional 5 μm pin hole gas doser backed by a pressure of 3–10 Torr in the gas reservoir. To purify the H_2O, the reservoir which contains de-ionized H_2O was frozen, pumped, and thawed. The exposures were measured in Torrs, which is the product of the backing pressure in the reservoir and the exposure time. It was found in a previous calibration with acetylene that 80 Torrs of acetylene with this doser is equivalent to 1 L of background pressure dosing. Using the relationship that the flux is proportional to (mass)$^{-0.5}$, 96 Torrs of H_2O is equal to 1 L.

III. RESULTS

A. Oxide overlayer formed from a multilayer of H_2O on AlAs(100)

HREEL spectra recorded from overlayers formed by adsorption of H_2O at 100 K followed by annealing to the indicated temperatures are shown in Fig. 1. At surface temperatures of 200 and 300 K, the O–H stretching vibrational mode at 3720 cm^{-1} is observed without the H_2O scissoring mode at 1650 cm^{-1}, indicating that the surface is populated by hydroxides from the dissociation of H_2O.[12,14,15] The intensity of this O–H stretching vibrational mode is independent of H_2O exposures above ~3 L at 100 K, indicating that the coverage of surface hydroxide has saturated.[12] In comparison, the As hydrides are not well resolved; possible depletion of the As hydrides by desorption of AsH_3 below room temperature may be responsible.[12] Upon annealing to a temperature of 600 K, the O–H stretching vibrational mode intensity is nearly eliminated, as desorption of H_2O and H_2 occurs.[12]

The extent of oxidation, observed from Auger spectra, for AlAs(100) with 16 and with 2 L of adsorbed H_2O at 100 K followed by annealing to 300 K is shown in Fig. 2. With 16 L of adsorbed H_2O at 100 K, considerable oxidation of AlAs(100) occurs upon annealing, as judged from the emergence of the Al^{3+} Auger transitions (39 and 54 eV) and oxygen Auger transition (510 eV). The reduced Al state (64 eV), indicative of AlAs(100), is substantially decreased in intensity. Furthermore, nearly complete depletion of the As (32 and 38 eV) Auger transitions is observed, which suggests that several layers of AlAs, beyond the escape depth of the As Auger electron (~8 Å), is disrupted, i.e., the Al–As bonds are broken and an Al hydroxide/oxide has formed. In contrast, after a much smaller 2 L exposure of H_2O at a surface temperature of 100 K followed by annealing, the extent of oxidation is minimal. The growth of Al^{3+} (39 and 54 eV), oxygen (510 eV), and the reduction in As (38 and 32 eV) are much less, despite a near saturation of the surface hydroxide as deduced by HREELS (*vide supra*). The temperature-programmed desorption (TPD) intensities of H_2 and H_2O, produced by the rehydrogenation, disproportionation, and dissociation of surface hydroxide at 400–700 K, also saturate.[12] This shows that the growth of the Al^{3+} Auger transition (39 and 54 eV) does not necessarily suggest greater coverage of the Al hydroxide but rather the formation of Al oxides.

The oxide structures resulting from adsorption of H_2O at a surface temperature of 100 K, followed by annealing were

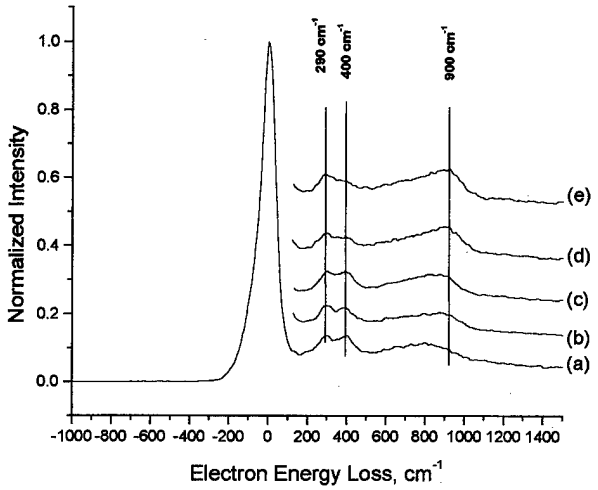


FIG. 3. HREEL spectra recorded after the indicated exposure at 100 K/anneal cycle. (a) 3 L exposure, 200 K anneal, (b) 3 L exposure, 600 K anneal, (c) 3 L exposure, 800 K anneal, (d) five cycles of H_2O exposure (3.5, 7, 14, 23, and 46 L), and 800 K anneal, (e) surface in (d) exposed to 46 L, 240 K anneal.

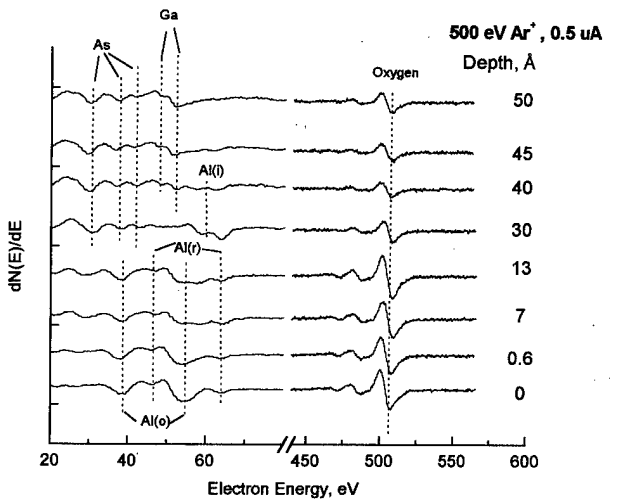


FIG. 4. Auger spectra recorded after sputtering (for the indicated depths) the oxidized AlAs film that results from eight cycles of 20 L of H_2O exposure at 100 K followed by annealing to 800 K. Al(r), Al(o), and Al(i) refers to the metallic Al, the oxidized Al^{3+}, and the intermediate oxide Al transitions, respectively.

also monitored with HREELS in order to compare the AlAs and GaAs phonon modes, and the Al oxide modes. Because of the strong dipolar characteristics of the phonon modes, the GaAs phonon mode at 290 cm^{-1} (from the buried GaAs layer) is observed adjacent to the AlAs phonon mode at 400 cm^{-1}.[16] In Fig. 3, the HREEL spectra recorded after various H_2O treatments are shown. The spectra recorded after exposing 3 L of H_2O at a surface temperature of 100 K followed by annealing to 200, 600, and 800 K are presented in Figs. 3(a), 3(b), and 3(c), respectively. The attenuation in the AlAs phonon mode is minimal, indicating that much of the AlAs lattice remains intact. In fact, upon annealing to a temperature of 800 K, the AlAs phonon mode has regained some of its lost intensity, presumably due to restructuring following desorption of the surface hydrides and the hydroxides.[12] Irreversible attenuation of the AlAs phonon mode is observed after many (>5) adsorption/anneal cycles in which the AlAs is exposed to a multilayer of water at 100 K and then annealed to 800 K, cf., Figs. 3(d) and 3(e). It is also clear after such a treatment that the Al oxide mode at 850–900 cm^{-1} has grown substantially.[17] The broad nature of the Al oxide mode suggests that suboxides may co-exist.

The extent of oxidation possible on this 50 Å thick AlAs overlayer on GaAs(100) was examined with an Auger sputtering profile, as shown in Fig. 4. The AlAs surface was oxidized by eight adsorption/anneal cycles in which the surface was exposed to 20 L of H_2O at 100 K followed by annealing to 800 K. It is clear that the oxygen (510 eV) signal does not significantly diminish in the near-surface region with sputtering, despite the emergence of an intermediate Al oxidic state at 58 eV (perhaps indicative of the range of oxidation states that could result from H_2O decomposition on AlAs) as well as the As signal at 31 eV. With further sputtering into the bulk, the Al^{3+} (39 and 54 eV) and oxygen (510 eV) peaks decrease substantially as the reduced Al(64 eV) peak emerges. However, some residual oxygen content within the sputtered matrix is maintained even as the subsurface GaAs layer is exposed. This shows that the AlAs lattice has been severely disrupted by oxidation and that an Al oxide has formed on the top 20–30 Å of the lattice.

B. Diffusion of the subsurface AlAs through the oxide overlayer

Shown in Fig. 5 are Auger spectra recorded after an AlAs surface that has been exposed to 15 L of H_2O at 100 K is annealed to 850 K and held at that temperature for the indicated times. The initial oxidation state of the surface is as expected from the analysis presented above. However, with increasing time at 850 K, we can see a clear increase in the metallic Al (64 eV) state with the annealing time. There is no

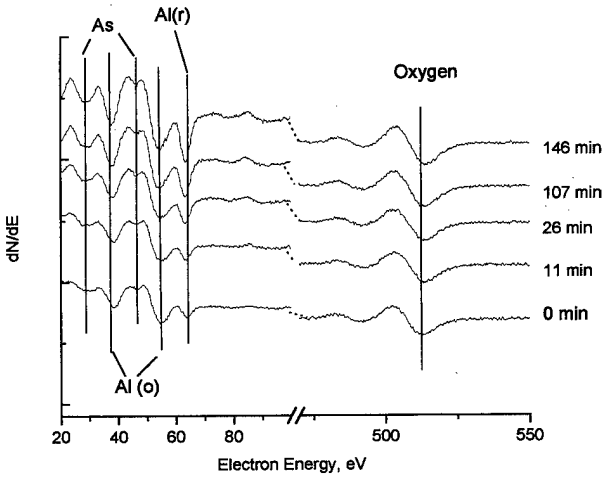


FIG. 5. Auger spectra recorded from the oxidized AlAs(100) surface that results from an exposure of 15 L of H_2O at a surface temperature of 100 K followed by annealing to 850 K for the indicated annealing times. Al(r) and Al(o) refers to the metallic Al and the oxidized Al^{3+} transitions, respectively.

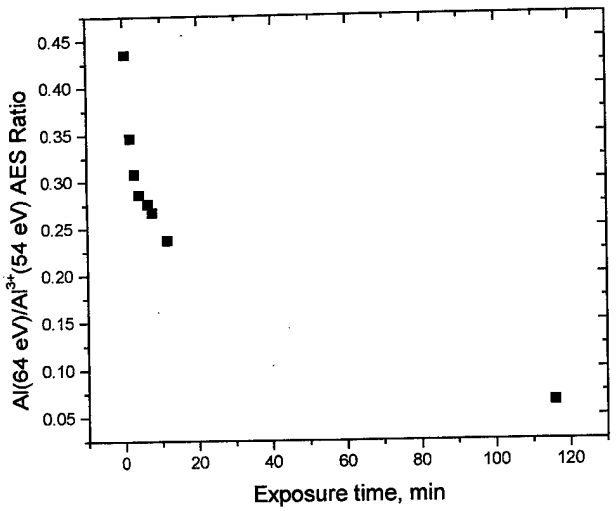


FIG. 6. Peak-to-peak Al(64 eV)/Al^{3+}(54 eV) Auger intensity ratio as a function of exposure time to a 3 keV electron beam at a surface temperature of 300 K. The AlAs(100) surface had been exposed to 16 L of H_2O at 100 K and annealed to 300 K.

desorption of Al oxides, such as Al_2O or Al_2O_3, detectable by mass spectrometry. Moreover, there is no decrease in the near-surface oxygen concentration (peak at 510 eV), as shown in Fig. 5 An As (31 eV) peak is also observed to emerge. Taken together, these results indicate that metallic Al and As can diffuse through the oxide layer to the surface (where they can react directly with H_2O) which explicitly reminds us that cationic diffusion also plays a part in the overall oxide growth mechanism.

C. Electron beam enhanced oxidation

We also considered the effects of electron beam exposure on the AlAs wet oxidation reaction. The AlAs surface was first exposed to 16 L of H_2O at 100 K and annealed to 300 K to produce a surface that is rich in surface hydroxides (*vide supra*). This partially oxidized surface was then exposed to a 3 keV electron beam (~3 μA current with a beam diameter on the order of 100 μm) for different times, whereupon the peak-to-peak intensity ratios of the reduced Al (64 eV) to the oxidized Al (54 eV) Auger transitions were recorded (see Fig. 6). The Auger spectra required ~70 s of electron beam exposure. It is clear from Fig. 6 that the surface Al undergoes further oxidation after a beam exposure time of ~2 h, as measured by the dramatically decreased peak-to-peak intensity ratio. The oxygen (510 eV) Auger transition intensity was found not to change with the electron beam exposure. Note that the surface regions which were not exposed to the electron beam maintained the initial oxidation state shown in Fig. 6. However, if the surface was annealed to temperatures at which the hydroxide phase is depleted (forming H_2 and H_2O which desorbs and Al oxide[12]), electron beam enhanced oxidation was not observed, indicating that the presence of surface hydroxides is critical to the effect.

IV. DISCUSSION

It has been observed that the near-surface region of AlAs is disrupted and an oxidic overlayer can form after a single multilayer water adsorption at 100 K/anneal to 300 K cycle. This is indicated by the substantial growth of the Al^{3+} state and depletion of the reduced Al and As transitions in Auger electron spectroscopy as well as the decrease in intensity of the AlAs phonon modes using HREELS.

These observations are the direct corollary of our earlier findings in which we observed overlapping dissociation and desorption kinetics of H_2O at a surface temperature of 150–250 K on AlAs(100).[12] Thus, from a multilayer of adsorbed H_2O, a greater coverage of H_2O is able to undergo the dissociation pathway even as the surface temperature is raised to 200 K. Furthermore, it is observed that oxidation of the subsurface region is possible, confirming the mechanism which disrupts the near-surface AlAs lattice. This is consistent with the observed desorption of AsH_3 at 150–250 K; AsH_3 is formed from the dissociation of H_2O, depleting As from the forming oxide matrix and exposing the subsurface to further oxidation.[12]

It was also found that subsurface Al and As can diffuse through the oxide overlayer at temperatures above 800 K. The high temperature required to observe this effect may be a result of a high activation barrier reaction that produces metallic Al and As at the bottom interface and/or diffusion of the metallic species through the oxide to the surface.[18] However, in an atmospheric steam furnace, propagation of the oxidation front occurs at temperatures as low as 700 K.[4] Thus, the present article cannot confirm cationic diffusion through the oxide as an integral part of the overall oxidation mechanism. Note, however, that more porous oxides are expected under atmospheric conditions due to the much higher flux of H_2O.

It has been observed that a partially oxidized AlAs surface populated with Al hydroxides can be further oxidized by prolonged exposure to a 3 keV electron beam. This suggests that the electron beam is able to induce dissociation of hydroxides and form Al oxides. This observation is analogous to that obtained from the isoelectronic system: electron beam enhanced sulfidization of GaAs from adsorption of H_2S.[19]

V. CONCLUSIONS

We have found that a large fraction of molecularly adsorbed H_2O on cold (100 K) AlAs surfaces undergoes dissociation upon annealing to form Al oxides and hydroxides. Thus, the extent of oxidation after a single adsorption/anneal cycle can be increased significantly by using water exposures that greatly exceed monolayer coverage. A substantial reduction in the AlAs phonon mode and growth of Al oxide modes at 850–900 cm^{-1} was observed (with HREELS) after repeated adsorption/anneal cycles. Sputter profiling shows that the Al^{3+} and O transitions (in AES) are present 20–30 Å below the surface. It was also found that electron beam irradiation of a partially oxidized surface could increase the surface oxidation, a result which may have implications for an *in situ* electron beam lithography application. Finally, after

annealing to temperatures in excess of 800 K, diffusion of the subsurface AlAs through the oxide overlayer was observed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of S. S. Shi in the MBE growth of the AlAs(100) samples used in this work. One of us (C.-H.C.) would like to thank the Korean Science Foundation for a successful post-Doctoral fellowship. Primary support of this research was provided by QUEST, a national Science Foundation and Technology Center (Grant No. DMR91-20007), and by the National Science Foundation (Grant No. DMR-9504400). Additional support was provided by the W. M. Keck Foundation.

[1]J. M. Dallesasse, N. Holonyak, A. R. Sugg, T. A. Richard, and N. Elzein, Appl. Phys. Lett. **57**, 2844 (1990).
[2]J. C. P. Chang, K. L. Kavanagh, F. Cardone, and D. K. Sadana, Appl. Phys. Lett. **60**, 1235 (1992).
[3]Y. S. Lee, Y. H. Lee, and J. H. Lee, Appl. Phys. Lett. **65**, 65 (1994).
[4]M. Ochiai, G. E. Giudice, H. Temkin, J. W. Scott, and T. M. Cockerill, Appl. Phys. Lett. **68**, 1898 (1996).
[5]W. T. Tsang, Appl. Phys. Lett. **33**, 426 (1978).
[6]E. I. Chen, N. Holonyak, and S. A. Maranowski, Appl. Phys. Lett. **66**, 2688 (1995).
[7]P. A. Parikh, S. S. Shi, J. Ibbetsen, E. L. Hu, and U. Mishra, Electron. Lett. **32**, 1724 (1996).
[8]A. R. Sugg, E. I. Chen, T. A. Richard, N. Holonyak, and K. C. Hsieh, J. Appl. Phys. **74**, 797 (1993).
[9]K. L. Lear, K. D. Choquette, R. P. Schneider, S. P. Kilcoyne, and K. M. Geib, Electron. Lett. **31**, 208 (1995).
[10]P. D. Floyd, B. J. Thibeault, E. R. Hegblom, J. Ko, L. A. Coldren, and J. L. Merz, IEEE Photonics Technol. Lett. **8**, 590 (1996).
[11]K. D. Choquette, R. P. Schneider, K. L. Lear, and K. M. Geib, Electron. Lett. **30**, 2043 (1994).
[12]W. J. Mitchell, C.-H. Chung, S. I. Yi, W. H. Weinberg, and E. L. Hu, Surf. Sci. **348**, 81 (1997).
[13]L. Lindau and W. E. Spicer, J. Electron Spectrosc. Relat. Phenom. **3**, 409 (1974).
[14]F. J. Szalkowski, J. Chem. Phys. **77**, 5224 (1982).
[15]P. A. Thiel, F. M. Hoffmann, and W. H. Weinberg, Phys. Rev. Lett. **49**, 501 (1982).
[16]J. L. Guyaux, A. Degiovanni, P. A. Thiry, R. Sporken, and R. Caudano, Appl. Surf. Sci. **56**, 697 (1991).
[17]J. G. Chen, J. E. Crowell, and J. T. Yates, Jr., Phys. Rev. B **33**, 1436 (1986).
[18]A. Mesarwi and A. Ignatiev, J. Appl. Phys. **71**, 1943 (1992).
[19]M. G. Noony, V. Liberman, and R. M. Martin, J. Vac. Sci. Technol. A **13**, 1837 (1995).

Exciton localization in InGaN quantum well devices

Shigefusa Chichibu[a)]
Materials Department, University of California, Santa Barbara, California 93106
and Faculty of Science and Technology, Science University of Tokyo, Noda, Chiba 278-8510, Japan

Takayuki Sota[b)]
Department of Electrical, Electronics, and Computer Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan

Kazumi Wada[c)]
Compound Semiconductor Materials Research, NTT System Electronics Laboratories, Atsugi, Kanagawa 243-0198, Japan

Shuji Nakamura[d)]
Department of Research and Development, Nichia Chemical Industries Limited, Kaminaka, Anan, Tokushima 774-8601, Japan

(Received 21 January 1998; accepted 11 May 1998)

Emission mechanisms of a device-quality quantum well (QW) structure and bulk three dimensional (3D) InGaN materials grown on sapphire substrates without any epitaxial lateral overgrown GaN base layers were investigated. The $In_xGa_{1-x}N$ layers showed various degrees of in-plane spatial potential (band gap) inhomogeneity, which is due to a compositional fluctuation or a few monolayers thickness fluctuation. The degree of fluctuation changed remarkably around a nominal InN molar fraction $x=0.2$, which changes to nearly 0.08–0.1 for the strained $In_xGa_{1-x}N$. This potential fluctuation induces localized energy states both in the QW and 3D InGaN, showing a large Stokes-like shift. The spontaneous emission from undoped InGaN single QW light-emitting diodes (LEDs), undoped 3D LEDs, and multiple QW (MQW) laser diode (LD) wafers was assigned as being due to the recombination of excitons localized at the potential minima, whose lateral size was determined by cathodoluminescence mapping to vary from less than 60 to 300 nm in QWs. Those structures are referred to as quantum disks (Q disks) or segmented QWs depending on the lateral size. Blueshift of the emission peak by an increase of the driving current was explained to be combined effects of band filling of the localized states by excitons and Coulomb screening of the quantum confined Stark effect induced by the piezoelectric field. The lasing mechanisms of the continuous wave $In_{0.15}Ga_{0.85}N$ MQW LDs having small potential fluctuations can be described by the well-known electron-hole-plasma (EHP) picture. However, the inhomogeneous MQW LDs are considered to lase by EHP in segmented QWs or Q disks. It is desirable to use entire QW planes with small potential inhomogeneity as gain media for higher performance LD operation. © *1998 American Vacuum Society.* [S0734-211X(98)05704-7]

I. INTRODUCTION

The InGaN alloys are attracting special interest because of their potential for the fabrication of light emitting devices operating in the red to ultraviolet (UV) energy region. Bright blue, bluish-green, and pure green light-emitting diodes (LEDs) have been put into practical use,[1] and the device lifetime of continuous wave (cw) operation of purplish-blue multiple quantum well (MQW) laser diodes (LDs) with modulation-doped strained-layer superlattice cladding layers grown on epitaxial lateral overgrown GaN (ELOG) substrates has been extended up to 10 000 h at room temperature (RT), as estimated from accelerated testing.[2]

All the good performance blue/green LEDs[1] and purplish-blue or UV LDs[1–9] reported to date have InGaN active layers. However, the material physics in InGaN is still unclear. One of the important issues in InGaN material is its phase-separating nature[10,11] due to a mismatch of thermodynamical and chemical stabilities between GaN and InN,[12–14] strain in InGaN caused by lattice mismatch between the InGaN and the GaN base layer,[15,16] or threading dislocation (TD) or defect-related driving forces.[15,17,18] Remarkable compositional separation and inhomogeneities are reported both in bulk three dimensional (3D) epilayers[10,11,14,16,17,19–22] and QWs.[17,18,23–28]

The practical LEDs and LDs grown on sapphire substrates have large TD densities up to 10^9 cm^{-2}.[1,23] However, they exhibit intense electroluminescence (EL) peaks or lasing operation. Thus it is interesting to clarify the emission mechanisms of InGaN-based QW and double-hetero (DH) structures for further improvement of device performance. A strong piezoelectric field in strained InGaN has been reported to significantly affect the optical properties of wurtzite nitride materials.[23,29,30] Moreover, since 3D wurzite GaN ex-

[a)]Corresponding author; electronic mail: chichibu@ee.noda.sut.ac.jp or chichibu@engineering.ucsb.edu
[b)]Electronic mail: tkyksota@mn.waseda.ac.jp
[c)]Present address: Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; electronic mail: k-wada@zx.netcom.com
[d)]Electronic mail: shuji@nichia.co.jp

hibits an excitonic photoluminescence (PL) peak even at RT,[31,32] it has been also important to investigate the contribution of excitons on the spontaneous emission mechanisms in InGaN QW structures and 3D epilayers.

This article describes the current understanding of emission mechanisms in InGaN materials and presents experimental results of optical and structural properties of device-quality InGaN single QW (SQW), MQW, and 3D layers grown on sapphire (0001) substrates with a low-temperature GaN buffer layer[1] only. These devices grown on ELOG substrates were not examined in this study. The spontaneous emission from them is assigned as being due to the recombination of excitons spatially localized at potential minima[23,25,28] whose lateral size varied from less than 60 to up to 300 nm (Ref. 27) due to the potential inhomogeneity.[18,19,21–23,25–28,33,34] Their sizes correspond to structures that are referred to as quantum disks (Q disks)[35] and segmented QWs depending on the lateral size. The lasing mechanisms of cw $In_{0.15}Ga_{0.85}N$ MQW LDs with small potential inhomogeneity are explained well by the electron-hole-plasma (EHP) model with strong Coulomb enhancement.[36] However, those LDs having inhomogeneous MQWs exhibited characteristic gain peaks within the Q disks or segmented QWs having lower resonance energies.

II. EXPERIMENT

The samples used in this study were grown on sapphire (0001) substrates by metalorganic vapor phase epitaxy (MOVPE) with low-temperature GaN buffer layers.[1] The AlGaN/InGaN 3D DH LED structures[1] were modified to have (i) a 50-nm-thick $In_{0.06}Ga_{0.94}N$ active layer codoped with Si and Zn, (ii) the same structure as (i) except for an undoped $In_{0.06}Ga_{0.94}N$ active layer, and (iii) the same structure as (i) except for a Si-doped (10^{19} cm^{-3}) $In_{0.06}Ga_{0.94}N$ active layer. For PL and PL excitation (PLE) measurements, a 50-nm-thick undoped 3D $In_{0.09}Ga_{0.91}N$ layer was grown on a 50-nm-thick $In_{0.01}Ga_{0.99}N$:Si/50-nm-thick $Al_{0.3}Ga_{0.7}N$:Si/2-μm-thick GaN:Si epilayer base.

The blue/green SQW LEDs have a 3-nm-thick undoped $In_xGa_{1-x}N$ QW ($x=0.3$ and 0.45), respectively.[1] For both spatially resolved and integrated cathodoluminescence (CL) measurements, 3-nm-thick undoped $In_xGa_{1-x}N$ ($x=0.05$, 0.2, and 0.5) SQWs were grown on a 3-μm-thick Si-doped (5×10^{18} cm^{-3}) GaN layer. The SQWs were subsequently capped by a 6-nm-thick undoped GaN layer, or left uncapped.

Several MQW LD wafers[1,2] were prepared. One of them has 10 periods of 2.5-nm-thick undoped $In_{0.2}Ga_{0.8}N$ wells and 7.5-nm-thick $In_{0.05}Ga_{0.95}N$ undoped barriers, whose pulsed lasing wavelength is 410 nm and whose spontaneous EL peak is 399 nm (3.11 eV) at RT.[1,23] Its threshold current density (J_{th}) was 11.4 kA/cm^2. The other cw LD wafer group had three 3.5-nm-thick Si-doped $In_{0.15}Ga_{0.85}N$ wells and Si-doped $In_{0.02}Ga_{0.98}N$ barriers. Two LD wafers, CW1 and CW2, were examined. A remarkable difference between the two is that the potential fluctuations in the QW plane in CW1 are much weaker than those in CW2. The J_{th} for RT cw operation was 7.4 and 9.4 kA/cm^2, and the wavelength was 398 and 409 nm for CW1 and CW2, respectively. The device lifetime was less than 1 h for both LDs. As will be shown later, CW1 has a small potential fluctuation, and showed a gain spectrum that can be fully explained by the EHP picture.

EL and photovoltaic (PV) spectra were measured using the LED or LD devices. PL, PLE, and modulated-electroabsorption (EA) spectra[19,23,27] were measured using the device wafers. For comparison, we also characterized single GaN epilayers by photoreflectance (PR) measurements.[37] Both the PR and EA spectra were analyzed using the Lorentzian line shape functional form. Static PL was mainly excited by the 325 nm line of a cw He–Cd laser. Time-resolved PL (TR-PL) measurements were also carried out using a N_2 pulsed laser or a frequency-doubled Ti:sapphire laser as an exciter. These measurements were carried out between 10 K and RT.

Spatially integrated or spatially resolved CL[27] was excited by a cw electron beam (e-beam) with or without e-beam scanning, and dispersed by a 25 cm focal-length grating monochromator coupled to a scanning electron microscope. The energy resolution of the CL spectrum was about 15 meV at 410 nm. The acceleration energy and current of the e-beam were typically 3 kV and 20 pA, respectively. Monochromatic spatially resolved CL images were taken under a fixed wavelength by e-beam scanning. The spatial resolution of the CL mapping is essentially limited by the diffusion length in the matrix, which corresponds approximately to less than 60 nm in this experiment, judging from the minimum dark area width. All CL measurements were carried out at 10 K.

To study gain spectra, spectrally resolved emission signals were measured using the variable excitation-stripe length (VEL) optical pumping method.[38] The excitation light was obtained from a frequency-tripled 10 Hz Q-switched Nd^+:YAG laser with a pulse duration of 10 ns. The excitation wavelength 355 nm nearly matches the exciton resonance energy of GaN at RT.[31,39] The net modal gain, $g(E)$, is obtained by measuring the stimulated emission intensity as a function of stripe length and pumping power intensity.[40]

III. RESULTS AND DISCUSSION

A. Excitonic structures in GaN

Excitonic features in optical spectra of GaN at RT are described as background to this work first. The optical absorption (OA), PR, and PL spectra of a GaN epilayer[31] measured at RT are summarized in Fig. 1. A clear excitonic resonance is found in both OA and PR spectra. Indeed, the resonance structures consist of two free excitonic (FE) resonances related to A and B transitions [FE(A) and FE(B), respectively]. The PL peak is a convolution of FE(A) and FE(B) emissions.[31] It is natural to observe FE emission from 3D GaN at RT, since this material satisfies several restrictions for excitons to survive at RT: (1) small numbers of active longitudinal optical (LO) phonons at RT, which is

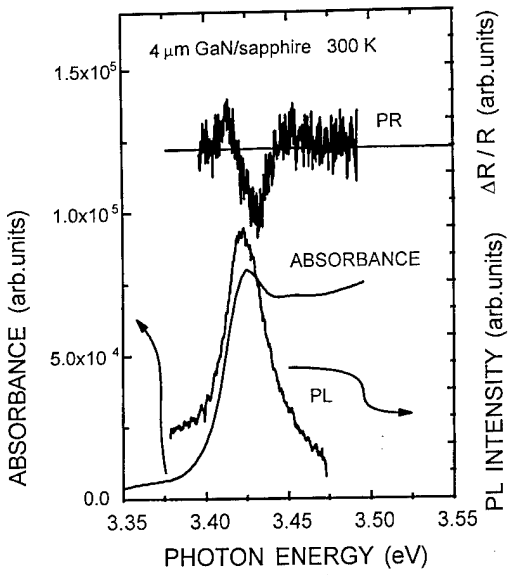


FIG. 1. Room-temperature PL, PR, and OA spectra of GaN/sapphire. The PL and PR spectra were obtained from 4-μm-thick GaN. The OA spectrum was obtained from a 1-μm-thick h-GaN/sapphire because the transmitted light intensity from the 4-μm-thick sample was so weak that we could not measure the spectrum. Since the degree of biaxial compressive strain differs for each, the spectral position (horizontal axis) of the OA spectrum is appropriately adjusted by referring to the FE(A) energy values of both samples, according to Ref. 39. FE emission is clearly observed from 3D GaN.

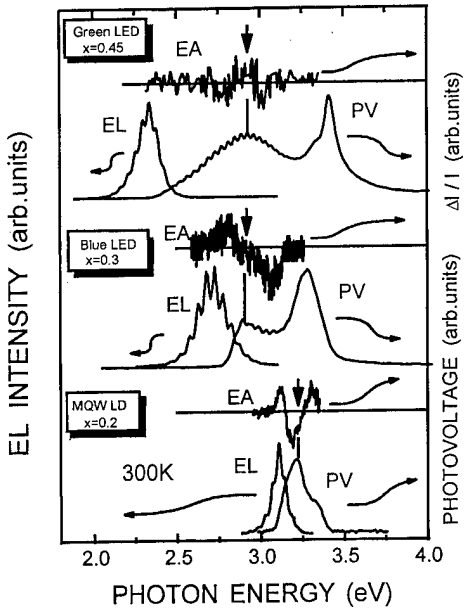


FIG. 2. EL, PV, and EA spectra for InGaN green and blue SQW LED structures and the MQW LD structure. The EL spectrum of the MQW LED was measured below J_{th}. The values of x in the $In_xGa_{1-x}N$ QW for green, blue, and MQW LEDs is 0.45, 0.3, and 0.2, respectively, assuming free-standing layers. The shoulderlike noticeable fringes in the EL and PV spectra are due to interference from multiple reflections. The structures in the EA spectra correspond to FE resonances.

satisfied if the LO phonon energy is far larger than the thermal energy k_BT of RT (26 meV), and (2) large binding energy of FEs (E_{ex}) greater than k_BT. Two other restrictions depend on the injected carrier density and the device structure; (3) exciton Bohr radius a_B smaller than the screening length, which is satisfied if the charge density is smaller than the critical charge density N_{crit} (Ref. 36) and (4) the electric field is weaker than the critical value (E_{ex}/a_B). The two former are easily satisfied, because the LO phonon energies in GaN are very large [91 meV for A_1 (LO) mode and 93 meV for E_1 (LO) mode][41] and few LO phonons are thermally activated even at RT.[42] Also, GaN has a large E_{ex} of 26 meV,[31,32,39] which is comparable to k_BT. The small dielectric constant ($\epsilon=8.2$) (Ref. 39) and a_B (3.4 nm),[39] and large effective masses[43] lead to N_{crit} as high as 1×10^{18} cm^{-3} according to Debye–Hückel screening.[36] Furthermore, the critical electric field strength is calculated to be 7.6×10^4 V/cm, which means that excitons in GaN are field-resistant particles. These excitonic material properties, which come from the strong Coulomb interaction in GaN, make it possible to observe FE emissions at RT.[31,32]

B. Optical properties of InGaN

Figure 2 summarizes the spontaneous EL, PV, and EA spectra of SQW ($x=0.3$, 0.45) and MQW ($x=0.2$) LEDs at RT. The EA spectra exhibit a derivativelike resonance structure. Because the broadening parameter Γ (Ref. 37) of the structures is very large, the structures due to the A and B transitions[31,39,44,45] are not resolved. The EA measurement monitors the FE resonance rather than the band-to-band transitions even at RT in wide gap semiconductors provided that a_B is small.[19,22,23,31,37,39,46] Therefore each PV peak and EA resonance corresponds to FE absorption in the QW. The result that the PV spectra exhibit a peaklike line shape also supports the FE resonance being observable at RT. The PV peak energy decreases from 3.21 to 2.91 eV with increasing x from 0.2 to 0.3. However, the peak energy is nearly unchanged for $x=0.3$ and $x=0.45$, and the full width at half maximum (FWHM) of all structures seems to increase with increasing x. These broadened EA and PV structures are discussed later. The spontaneous EL peak is located at the lower energy tail of the FE resonance, exhibiting the large Stokes-like shift in the QWs.

The observation of a Stokes-like shift in actual QWs, which are defined as ultrathin films whose thicknesses are smaller than the 3D exciton Bohr radius (in this case, thinner than 3 nm), directly proves the existence of potential fluctuation in the QW plane, since electron and hole quantized energy levels do not suffer from energy broadening due to the electric field normal to the QW plane.[47,48]

Supporting evidence for the potential fluctuation in InGaN materials was also obtained from the optical properties of 3D bulk InGaN layers with small x as follows. Figure 3 shows PL and PLE spectra at 10 K of a 50-nm-thick 3D $In_{0.09}Ga_{0.91}N$ epilayer grown on an $In_{0.01}Ga_{0.99}N/Al_{0.3}Ga_{0.7}N$ stacked buffer layer.[19,23] The PL peak energy is unchanged by changing the excitation intensity up to four orders of magnitude. The PL peak at 3.28 eV is effectively excited by the FE absorption in both the active and barrier layers, shown by the arrows followed by a band-to-band (B–B) con-

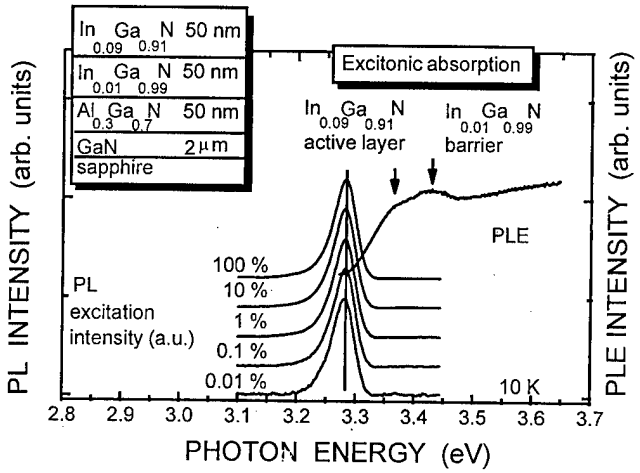


FIG. 3. PL and PLE spectra of a 50-nm-thick undoped 3D $In_{0.09}Ga_{0.91}N$ layer at 10 K. The PLE spectrum exhibits two excitonic absorption peaks originating from the $In_{0.09}Ga_{0.91}N$ top active layer and the $In_{0.01}Ga_{0.99}N$ barrier layer, as shown by the arrows. The Stokes-like shift is 85 meV.

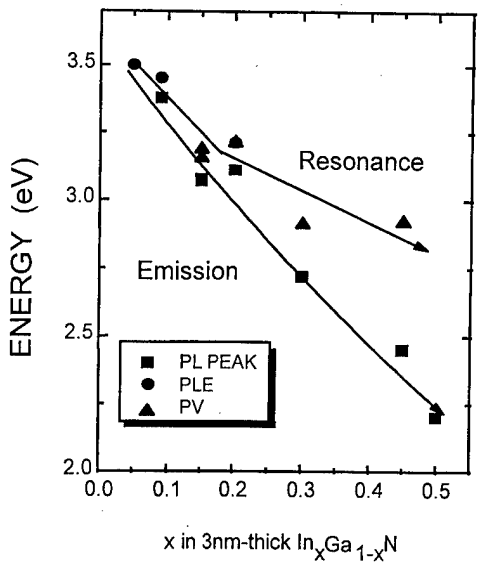


FIG. 5. Free excitonic resonance and emission peak energies in InGaN QW devices and wafers as a function of InN molar fraction x. The energy difference between the resonance and emission increases with increasing x.

tinuum. The PL peak is located at the lower energy tail of the FE absorption. The FWHM of the FE resonance in the $In_{0.09}Ga_{0.91}N$ active layer is estimated to be about 75 meV, showing a large potential fluctuation. The Stokes-like shift is 85 meV.[19,23]

Various optical spectra of the 50-nm-thick 3D $In_{0.06}Ga_{0.94}N$ layers at RT are shown in Fig. 4. The EA spectrum of the undoped layer exhibits a derivativelike, broadened resonance structure without any Franz–Keldysh[49] oscillations, as is the case with the PR spectra of 3D thick InGaN layers.[22,50,51] The broadening is attributed to the potential fluctuations in InGaN. This is proved later by the spatially resolved CL measurement. The FE energy in 3D $In_{0.06}Ga_{0.94}N$ at RT is 3.298 eV, which agrees well with the PLE peak energy in $In_{0.06}Ga_{0.94}N$ codoped with Si and Zn, as shown in the bottom trace. The decrease of the PLE signal in the higher-energy side is due to the absorption by the top p-GaN layer. Nevertheless, the PLE spectrum exhibits an excitonic absorption even at RT, as is the case at 10 K (Fig. 3). Again in the undoped 3D $In_{0.06}Ga_{0.94}N$ device structure, the spontaneous EL peak appeared in the lower-energy tail of the FE resonance even at RT, showing a Stokes-like shift of 40 meV.[19,23] We note here that radiative decay of localized excitons in 3D InGaN layers has been reported by TR-PL measurements,[52] temperature-dependent PL measurements,[53] and modulation spectroscopy measurements.[19,23]

The relation between the Stokes-like shifts and x is shown in Figs. 5 and 6. As shown in Fig. 5, the energy difference between the FE resonance and the emission increases with increasing x both in the QWs and in the 3D layers. The

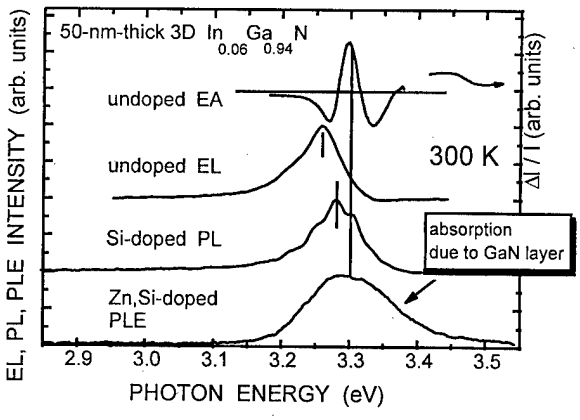


FIG. 4. Optical spectra of 50-nm-thick 3D $In_{0.06}Ga_{0.94}N$ device structures at RT. The PLE spectrum of the Si- and Zn-codoped InGaN monitored a blue emission band around 460 nm. The Stokes-like shift of the undoped $In_{0.06}Ga_{0.94}N$ active layer was nearly 40 meV at RT.

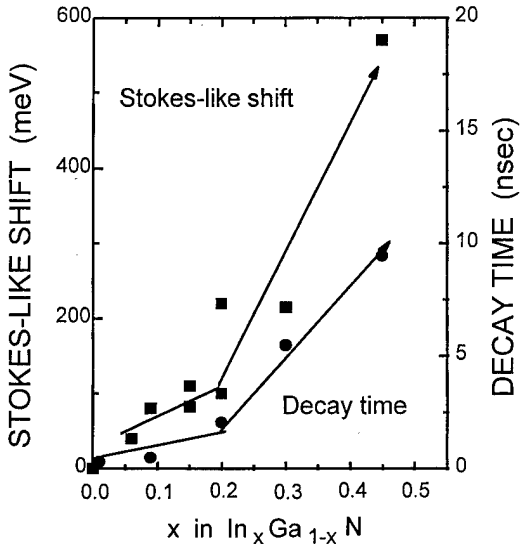


FIG. 6. Stokes-like shift and emission decay time of InGaN as a function of x. The emission decay time was measured directly by TR-PL measurements. There is a certain critical InN molar fraction, nominally nearly $x=0.2$, where the dependences of the Stokes-like shift and decay time on x change significantly.

longer component of the decay time obtained by the TR-PL measurements also increases with increasing x. Such increase of the decay time is one of the characteristics of localized electronic systems,[35] and detailed TR-PL measurements have been done by Narukawa *et al.*[25] and by Im *et al.*[54] However, further experiments are necessary to explain the long decay time; e.g., enhancement of the decay time by the piezoelectric field should be taken into account.

It should be mentioned here that there is a certain critical InN molar fraction, nominally nearly $x=0.2$, where the dependences of the FE resonance, Stokes-like shift, and decay time on x change remarkably. The Stokes-like shift, especially, increases rapidly for $x>0.2$. The 3-nm-thick well width MQW LD structure with $x=0.2$ exhibited a Stokes-like shift of 100 meV;[23] however, other 3-nm-thick QW samples with $x=0.2$ exhibited a larger Stokes-like shift of 250 meV.[25,27] This means that the degree of potential fluctuation changes around $x=0.2$, which can explain the inconsistency between the results of Narukawa *et al.*[25] and our results,[23,27] and which has been argued before. The $In_xGa_{1-x}N$ layers having $x>0.2$ are considered to have some extra mechanisms for producing the Stokes-like shifts.

Recently, Takeuchi *et al.*[29] reported that AlGaN and InGaN alloys grew coherently on thick GaN layers at least up to thicknesses of 650 and 40 nm, respectively. Therefore, x values in strained InGaN layers, especially QWs, must be carefully estimated. If this is also the case for our InGaN samples, the x values may be modified to about half of the original values determined by the x-ray diffraction (XRD) measurements, according to their reliable bowing parameter for strained InGaN (3.2 eV).[29]

C. Quantum confined Stark effects in InGaN QWs

Another characteristic observed for the emission properties of InGaN QW devices is a blueshift of the EL peak with increasing driving current or excitation intensity. Taguchi *et al.*[55] and the authors of Ref. 23 have found a blueshift of the EL peak in blue and green SQW LEDs[23,55] and MQW LDs.[23] EL spectra of the SQW LEDs are shown in Figs. 7(a) and 7(b) as a function of driving current. The emission intensity increases approximately linearly by increasing driving current. A similar relation between the intensity and the excitation level was observed for the PL and EL spectra of all the QW and 3D structures investigated here. The emission intensity is nearly constant from 10 K to RT in these SQW LEDs. Note that the distinct shoulderlike fringes in the EL spectra are due to internal multiple reflections at the surface and at the n-GaN/sapphire interface. The EL peaks shifted to higher energy by 60 and 110 meV, respectively, for the blue and green SQW LEDs.

Takeuchi and co-workers[29,56] have also found a blueshift of the PL peak in InGaN/GaN QWs by increasing the excitation level. They observed tremendous blueshifts up to 280 meV by exciting the sample above the threshold power to obtain stimulated emission, and concluded that the blueshift was due to Coulomb screening of the quantum confined Stark effect (QCSE)[47,48] induced by the piezoelectric field.[57]

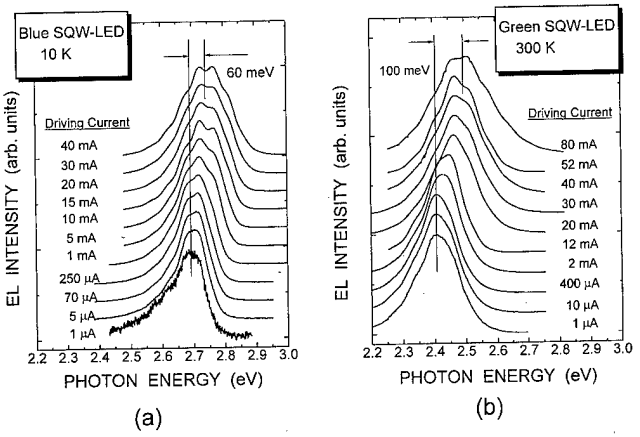


FIG. 7. EL spectra of (a) blue and (b) green InGaN SQW LEDs (product ID Nos. NSPB and NSPG) as a function of the driving current. The shoulder-like noticeable fringes in the spectra are due to the interference effect from multiple reflections. Spectral broadenings can be recognized as well as a blueshift of the emission peak with increasing driving current.

Actually, in hexagonal nitrides, the QCSE due to the piezoelectric field as well as that due to the built-in field in the (0001)-oriented strained QWs should be taken into account[23,29,30,56] to estimate the Stark shifts and to consider exciton wave functions in the QWs. It was reported that the directions of the piezoelectric and the built-in field are opposite.[29] Thus both effects are estimated as follows. Since the QWs are very thin (3 nm) and both n-GaN and p-$Al_{0.2}Ga_{0.8}N$ barriers are highly doped, a strong built-in electric field exists across the QW plane. The field strength is as high as 8.5×10^5 V/cm, which is estimated from the values of electron and hole concentrations in n- and p-type barriers (5×10^{18} and 1×10^{18} cm^{-3}, respectively). The piezoelectric field strength in a 1% strained GaN QW is estimated to be of the order of MV/cm.[23,29,56,58] Thus the Stark shift due to the QCSE in a 3-nm-thick undoped QW with an undoped barrier is estimated to be less than 100 meV for a field of 1 MV/cm and rapidly increases up to a few hundred meV for the fields more than a few MV/cm.[23,58] Under such a high electric field, the emission intensity should decrease and the emission lifetime should increase. However, we have observed very bright EL from InGaN QW devices, showing that the effect of the piezoelectric field is not so strong in actual devices. Both injected carriers and doped impurities in the barriers (sometimes also in the wells) are considered to result in effective Coulomb screening of those fields.[58]

Figure 8 shows PL spectra at RT of an $In_{0.45}Ga_{0.55}N$ green SQW LED structure as a function of external bias. The PL was excited by the 457.9 nm (2.71 eV) line of a cw Ar^+ laser (50 mW), which excites carriers only in the SQW. The peak intensity decreases with increasing reverse bias. The PL spectrum for $V=+1.991$ V corresponds to that taken under open-circuit conditions. By applying -2 V reverse bias, the PL intensity decreases to one-third that for the $+2$ V bias. The emission vanishes for a reverse bias of -10 V. How-

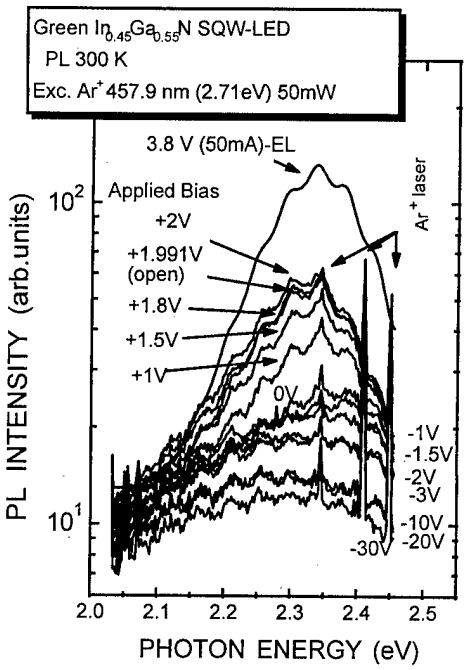


FIG. 8. PL spectra of a 3-nm-thick $In_{0.45}Ga_{0.55}N$ SQW structure as a function of external bias measured at RT. The PL was excited by the 457.9 nm (2.71 eV) line of a cw Ar^+ laser (50 mW), which excites carriers only in the SQW. From Fig. 2, it is clear that the excitation energy is in the energy range for free excitonic absorption in the QW. The EL spectrum is also shown for comparison. The open-circuit voltage with the weak laser irradiation is 1.991 V.

ever, it is difficult to judge the direction of the peak energy shift, since the QW thickness is only 3 nm.

As shown in Figs. 7(a) and 7(b), the blue and green EL peaks shifted to higher energy by 60 and 110 meV, respectively, in actual devices. These blueshifts could be explained by the QCSE assuming the appropriate degree of residual strain. However, both of the EL spectra shown in Fig. 7 exhibited spectral broadening with the blueshift (increasing the current). This phenomenon cannot be explained only by the QCSE in terms of Coulomb screening of the piezoelectric field by current injection. Therefore the blueshift in the EL peak energy may be due to combined effects of band filling of the local potential minima by excitons and Coulomb screening of the QCSE due to the piezoelectric and the built-in fields.

D. Potential inhomogeneity and exciton localization in InGaN

Usually, the energy difference between the absorption and emission, which is referred to as the Stokes shift, is caused by a potential inhomogeneity in the active layers. There are several mechanisms for carrier localization in QWs that exhibit the Stokes-like shift shown in Fig. 9: (i) monolayer thickness fluctuation, (ii) spatial compositional (or strain) undulation, and (iii) complete phase separation. These are referred to as cases A, B, and C, respectively, hereafter. Case A and relatively weak potential fluctuation (case B) usually produce a weak potential undulation in QWs. A strong com-

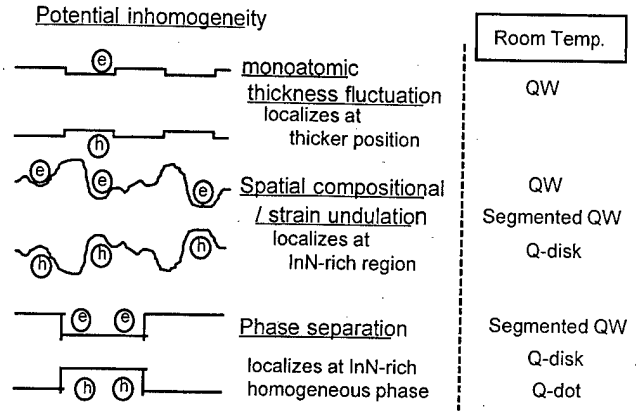


FIG. 9. Schematic illustration of the origin of the Stokes-like shift and carrier localization in InGaN. Each structure may act as a QW, segmented QW, Q disk, or Q dot depending on the lateral size and the degree of the potential inhomogeneity at RT.

plete phase separation (case C) produces well-defined quantum dots (Q dots), Q disks, or segmented QWs depending on the lateral size. The case in between B and C produces a strong potential undulation where a Q-disk potential or segmented QW condition is sustained.

In order to figure out the spatial distribution of the emission, three 3-nm-thick $In_xGa_{1-x}N$ SQWs ($x=0.05$, 0.2, and 0.5) capped by a 6-nm-thick GaN layer were prepared. The growth parameters were exactly the same as those for the blue and green SQW LEDs[1] except for the GaN capping.

First we observed cross-sectional transmission electron micrographs (TEM) and found very abrupt heterointerfaces between the SQW and GaN in all the samples.[27] The TD density was less than 10^9 cm^{-2}. Abrupt heterointerfaces were also observed in the $In_{0.2}Ga_{0.8}N/In_{0.05}Ga_{0.95}N$ MQW LD wafer.[1,23] The atomic force microscopy (AFM) images of GaN-capped and uncapped $In_{0.5}Ga_{0.5}N$ SQWs were also examined. The GaN-capped SQW exhibited a flat surface with nanoscale holes due to TDs. Also, a several monolayer step structure presumably due to spiral growth initiated by TDs[59] was found.[27] The spiral growth is considered to create several monolayers of thickness fluctuation in the QW plane. On the other hand, the uncapped bare SQW exhibited many grooves (10–50 nm in diameter, 1–2 nm in depth). These grooves are considered to be areas of reevaporated InN-rich materials in the SQW since the MOVPE recipe was terminated just after the SQW growth at a growth temperature where the vapor pressure of InN is higher than that of GaN.

A wide-area (500×500 μm^2) PL spectrum and its PLE spectrum of the GaN-capped $In_{0.2}Ga_{0.8}N$ SQW taken at 10 K are shown in Fig. 10. The PL peak energy is 3.017 eV, and the FWHM is nearly 130 meV. The averaged FE resonance energy in the SQW is 3.28 eV, showing a Stokes-like shift of 270 meV at 10 K. These values are larger than those reported before[25,27] and suggest the presence of large potential fluctuation, although we should consider the separation of the $e-h$ wave function in triangular potentials formed at opposite sides of the well.

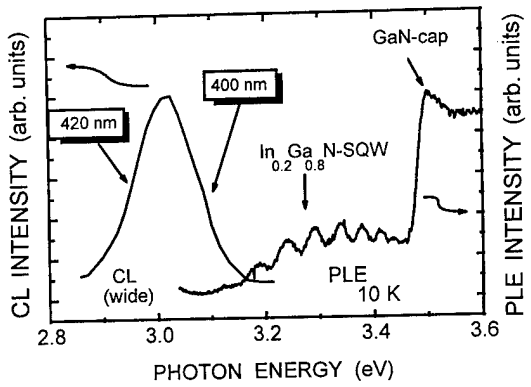


FIG. 10. Wide-area scanning CL spectrum and its PLE spectrum of a 3-nm-thick $In_{0.2}Ga_{0.8}N$ SQW capped by a 6-nm-thick GaN. Both spectra were measured at 10 K. The scanning area for the CL measurement was $10 \times 10\ \mu m^2$. The spectral modulation in the PLE spectrum is the interference effect. A large Stokes-like shift of 260 meV was observed.

A wide-area ($10 \times 10\ \mu m^2$) integrated scanning CL spectrum from two different positions on the same sample measured at 10 K is shown in Figs. 11(a) and 11(b). As the case with the result reported before,[27] the CL peak energies obtained from limited spot areas ($\sim 500 \times 60\ nm^2$) varied from position to position. The FWHM values of the spot CL peaks (~20 meV) are much smaller than that of the wide-area scanning CL peak [~90 (Ref. 27) or 130 meV (Ref. 23)]. Note that the spectral resolution is as large as 15 meV. These results clearly show that the wide-area broadened CL (and PL) peak consists of sharper emission peaks having various peak energies, which are separated by nearly 50 meV. The appearance of several CL peaks in the spectra (b)–(e) in Figs. 11(a) and 11(b) is not due to a multiple interference

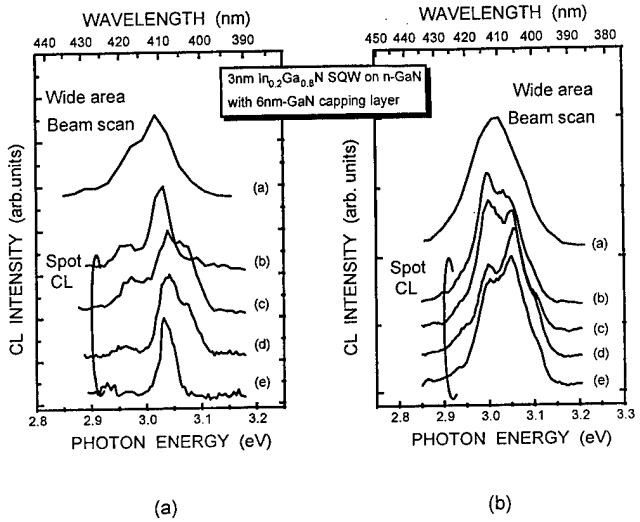


FIG. 11. Typical CL spectra of a GaN-capped $In_{0.2}Ga_{0.8}N$ SQW measured at 10 K. The wide area scanning CL spectra were obtained from an area $10 \times 10\ \mu m^2$. The fixed spot-area CL spectra were taken from an area of smaller than $\sim 500 \times 60\ nm^2$. The CL peak energies obtained from limited spot areas varied from position to position, as shown in traces (b)–(e) in both (a) and (b). The FWHM values of the spot CL peaks (less than 30 meV) are much smaller than those of the wide-area scanning CL peak (90–130 meV).

effect. Therefore, there exist several nanoscale structures having different net band gap energies. This structure can act as a segmented QW, a Q disk, or Q dots having compositional and/or size inhomogeneity depending on their lateral size.

Monochromatic scanning CL images of the GaN-capped $In_{0.2}Ga_{0.8}N$ SQW taken at wavelengths of 400 (3.100 eV) and 420 nm (2.952 eV), which correspond to the higher- and lower-energy sides of the wide-area integrated CL peak at 411 nm (3.017 eV), indicated the following results;[27] (i) each bright area emitting CL consists of emissions from regions of less than 60 up to 300 nm in lateral size; (ii) some dark areas in one wavelength correspond to bright ones in the other wavelength; (iii) some areas exhibit both 400 and 420 nm CL emissions; and (iv) approximately 40% of the entire areas are dark at both wavelengths. These results can be explained by the presence of potential inhomogeneity whose lateral size/interval is smaller than 60 nm. This value is the spatial resolution of the system, and is essentially limited by the carrier diffusion length.

The structures having lateral sizes smaller than the exciton resonance wavelength, which is the emission wavelength divided by the refractive index and corresponds approximately to 160 nm in GaN, are defined as Q disks.[35] In Q disks, the spontaneous emission lifetime of excitons is longer than that in two dimensional QWs.[35] Here we should mention that both Q disks and segmented QWs lase via EHP with the Coulomb enhancement,[36,60] since there is no or only very weak carrier/exciton confinement in the lateral directions. Therefore we call the structure Q disks and/or segmented QWs hereafter, because we cannot define their absolute sizes. Note that the electron and hole wave functions in Q disks or segmented QWs are quantized at least with respect to the z direction even under the electric field.[47,48] Result (iv) can also be interpreted as the presence of Q disks superimposed on the large scale inhomogeneity expanding up to 300 nm in lateral size.

At 420 nm, the bright area density is about $5 \times 10^8\ cm^{-2}$, which is close to the TD density. Also, the spacing between the large bright areas is also close to that of the TD spacing. Very recently, Sato *et al.*[17] showed featureless CL images for 110-nm-thick $In_{0.14}Ga_{0.86}N$ layers grown on homoepitaxially grown thick GaN layers having very low TD densities. Their result was quite different from the inhomogeneous CL images.[18,27] Thus the TDs are likely to act as some kind of driving force of the potential fluctuation in InGaN. The monochromatic CL image of the $In_{0.05}Ga_{0.95}N$ SQW also exhibited an inhomogeneous fine structure, and the Stokes-like shift was nearly 60 meV. The structure size was smaller than that of the SQW with $x = 0.2$.

There are two possible origins of the potential fluctuation: (1) compositional inhomogeneity and (2) monolayer thickness fluctuation. If we assume a coherent growth of InGaN,[29] the relative change in the x value is estimated to be less than a few percent to reproduce a band gap change of nearly 100 meV. It is rather difficult to distinguish such small compositional inhomogeneity. In addition, InN-rich regions should

suffer from a larger QCSE because the residual strain in InN-rich regions is larger than that in InN-poor regions, resulting in a stronger piezoelectric field in the InN-rich region. Thus the energy difference between two regions having different InN molar fractions is a sum of the change in the alloy energy gap and the change in the Stark shift due to the piezoelectric QCSE if we assume coherent growth. In such QWs, local strain distribution inhomogeneity may present. Ponce *et al.*[18] have found such local strain inhomogeneity in InGaN SQWs. Next, if we assume monolayer thickness fluctuation without any compositional inhomogeneity, the energy difference between two regions having different well thicknesses is a sum of the change in the quantized energy level and the change in the Stark shift due to the piezoelectric QCSE. Since several groups[27,59] have shown 4–6 monolayer thickness steps in AFM images of GaN-capped SQW, it might explain the appearance of several sharp CL peaks shown in Fig. 11. In both cases, the carrier localization due to the potential fluctuation is enhanced by the Stark effect. In this case, the electronic confinement system can be referred to as the Stark-enhanced Q-disk/segmented QW potential.

It has been argued that InGaN QWs contained InN-rich Q dots having a lateral size of a few nm.[1,18,21,24,26] Since the spatial resolution of our CL system is 60 nm, it is impossible to discuss that question. The large Stokes-like shift (260 meV) due to potential inhomogeneity in the SQW indicates that most parts of the SQW could have larger band gap energies. Those parts would be seen as dark areas at the wavelengths monitored here provided that the diffusion length is short enough. It follows from result (iv) that the real area emitting CL is much smaller than the bright area observed. Anyway, it has been proven that electrons injected in such wide gap areas move to narrow gap areas to emit CL. Such Q disks or segmented QWs can enhance the emission efficiency in the QW due to the limitation of exciton movement in small spaces, which can reduce nonradiative pathways. A schematic representation of the potential fluctuation, Stokes-like shift, and localized tail states in InGaN QWs is shown in Fig. 12.

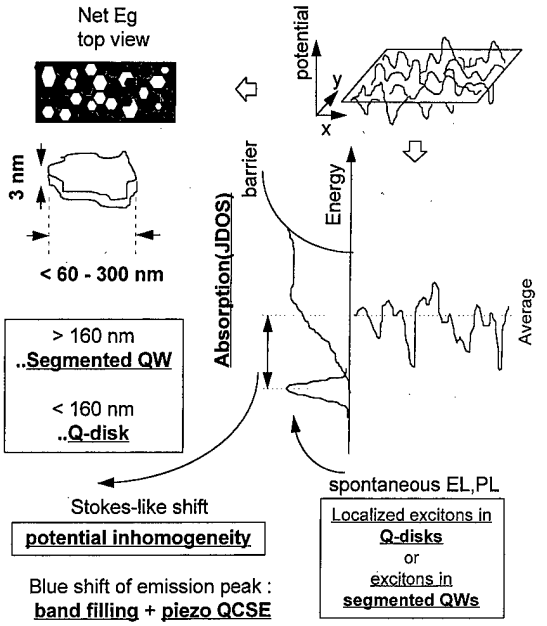


FIG. 12. Schematic illustration of the potential inhomogeneity, size, joint density-of-states, Stokes-like shift, and localized excitons in InGaN QWs.

E. Localized excitons in mesoscopic semiconductor quantum disks in inhomogeneous $In_xGa_{1-x}N$ QWs

The spontaneous emission peak from InGaN QWs and 3D layers is assigned as being due to the recombination of excitons localized at potential minima[1,19,23,25–27] as follows. We compared the environmental conditions for excitons assuming that LED devices are operated under the catalog dc values ($I=20$ mA) or less. (1) As is the case with GaN, few LO phonons are thermally activated in InGaN at RT due to their large energies [91.8 meV for GaN (Ref. 41) and 73.6 meV for InN (Ref. 61)]. (2) Because the InN molar fraction x of the undoped DH LED is as small as 0.06, which can be deduced to be nearly 3% provided that the 3D layer is strained, E_{ex} of a 3D layer is considered to be nearly the same as that of GaN (26 meV).[31,39,43–45]

For the undoped 3D DH LEDs, (3) the injected carrier density for dc EL operation is estimated to be less than 10^{17} cm^{-3} using the values of active layer thickness (50 nm), injected current density, and emission decay time ($\tau = 1$ ns). (4) The piezoelectric field is weaker than the critical value (7.6×10^4 V/cm), since there is less than 0.3% lattice mismatch. Under these conditions, 3D FEs can survive at RT in the undoped 3D LED. The localization energy for the particular sample was 40 meV at RT, which is larger than the thermal energy at RT. Note that emission is either localized or nearly free depending on the sample quality, since x is very small.

Two restrictions, (3) and (4), are critical for QWs, since there exist strong piezoelectric and built-in fields, and high carrier density is established in very thin 3-nm-thick SQWs. It is known that excitons can survive up to RT in QWs because of the increase of E_{ex} due to the confinement of the wave functions.[47,62–65] Thus we have calculated E_{ex} in 3-nm-thick GaN QWs by a variational approach taking the electron-hole pair motion in the QW into account.[23,66] The value obtained (47 meV) is about 1.8 times larger than that in the 3D case. Uenoyama[67] has also calculated E_{ex} of 4-nm-thick GaN QWs as 40 meV, which agrees with our calculation. Therefore the electric field needed to dissociate FEs in the GaN QWs is as high as 6.0×10^5 V/cm. Note that, even in GaAs QWs, excitons survive at RT under an electric field of 10^5 V/cm.[47,65] If we assume the entire Stokes-like shift is a superficial effect due to the piezoelectric QCSE, which will appear when the absorption and emission are measured under different field strengths, the piezoelectric field exceeds this value. However, we measured both the absorption and luminescence under low excitation conditions. Thus both the QWs and the 3D layers still have areas having a potential minimum even at RT, where carriers or excitons are spatially localized. The spectral broadening due to the increase of the

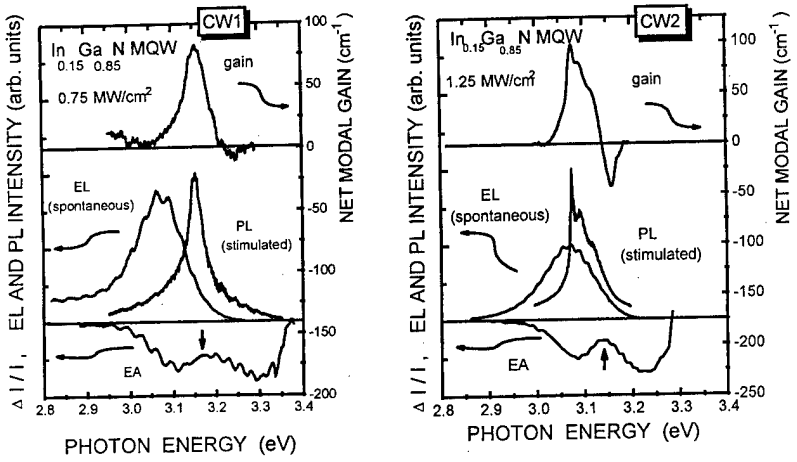


FIG. 13. Optical gain spectra of the cw $In_{0.15}Ga_{0.85}N/In_{0.02}Ga_{0.98}N$ MQW LD wafers with less inhomogeneous (a) and certainly inhomogeneous (b) QW potentials. The two samples are abbreviated as CW1 and CW2, respectively. Spontaneous EL, stimulated PL, and EA spectra are shown to compare the sample properties. The average InN molar fraction x in the well for CW2 is slightly higher than that of CW1. CW1 exhibits a gain spectrum which can be explained by EHP picture including the band gap renormalization. CW2 lases lower energy parts of the QW plane, which are segmented QWs and/or Q disks.

driving current shown in Fig. 7 is a sign of the band filling of such regions, where the electric field is weakened by the forward current injection. Chichibu *et al.*[19,23] observed clear FE absorption from both 3D and QW InGaN materials at RT, the data being consistent with the presence of excitonic absorption. Indeed, the piezoelectric field is already partly weakened by doping Si in the barrier layer.[1]

The carrier density in the SQW is estimated to be about 1.2×10^{18} cm^{-3} ($I=20$ mA, $\tau=3$ ns). This value is comparable to the critical charge density for FEs in 3D GaN. Thus observation of the excitonic emission seems to be possible in the potential minima in QWs, where the electron-hole distance is shorter than that in the 3D case.[62–64] The improvement of the emission intensity in QWs compared to that in the 3D case[23] may be attributable to the increased E_{ex} and oscillator strength of localized excitons in the Q disks or segmented QWs in the QW plane.

It should be noted here that electron and hole wave functions are slightly separated towards the p- and n-type barrier sides of the QW by the piezoelectric field,[58] but they maintain enough overlap to emit bright emissions.[1] The electron-hole pair present in the potential minima of QWs can be referred to as localized excitons, because the wave functions are confined in the QW and under such conditions the electron and hole produce a bound state, which are pictures of quantized excitons, even if the wave function overlapping is partly weakened.[62–64]

F. Optical gain in cw LDs with relatively homogeneous and inhomogeneous MQWs

The optical gain spectra of CW1 and CW2 are compared in Figs. 13(a) and 13(b), respectively. The arrows on the EA spectra indicate the average FE resonance energy in the QWs. The stimulated emission peak energy of CW1 is close but slightly lower than the FE resonance energy, and the gain spectrum can be well explained by the EHP model.[36,40,68,69] The FWHM of the gain peak for CW1 is smaller than that of CW2, which means that potential fluctuation for CW1 is smaller than that of CW2. It is obvious that CW2 contains more In in the well, since the EA resonance energy of CW2 is lower than that of CW1.

The appearance of sharp stimulated emission and gain peaks on the spontaneous EL peak for CW2, which appeared for high pumping energy density or long excitation-stripe lengths, indicates that there exist actual energy states below the averaged FE resonance energy. A remarkable induced-absorption peak near the FE resonance energy appeared[40] simultaneously with the appearance of the sharp gain peaks. This induced absorption might be due to the effective resonance between two degenerate levels having lower transition energy and one undegenerate level (three-level lasing model).[60,67,70,71] Mohs *et al.*[72] have also found induced absorption in the InGaN MQW LD structure. However, further experiments are necessary to disclose the physics underlying this gain/absorption mechanism. At any rate, the result shows that CW2 lases by EHP in segmented QWs or Q disks. The J_{th} value for CW1 is smaller than that of CW2, and therefore it is desirable to use entire QW planes as gain media for higher performance LD operation.

IV. CONCLUSION

Emission mechanisms of the device-quality InGaN SQW and MQW structures grown on sapphire substrates without any ELOG base layers were investigated. The $In_xGa_{1-x}N$ layers showed various degrees of spatial effective band gap inhomogeneity. The dependence of the degree of inhomogeneity on x had a critical InN molar fraction around, nominally, $x=0.2$ (nearly 0.08–0.10 for the strained $In_xGa_{1-x}N$). The potential undulation induces the localized energy states both in QW and 3D InGaN, and shows a large Stokes-like shift.

The spontaneous emission from undoped InGaN SQW LEDs, undoped 3D DH LEDs, and MQW LD wafers was assigned as being due to the recombination of excitons localized at the potential minima, whose area was determined by monochromatic CL mapping to vary from less than 60 up to 300 nm in lateral size.

The blueshift of the emission peak by increasing the driving current was explained as being due to combined effects of band filling of the localized states by excitons and Coulomb screening of the quantum confined Stark effect induced by the piezoelectric field.

The lasing mechanisms of the cw $In_{0.15}Ga_{0.85}N$ MQW LDs with small potential fluctuation can be described by the well-known EHP picture with strong Coulomb enhancement while LDs with a strongly inhomogeneous MQW exhibited a sharp gain peak in the lower energy parts of the QW planes. Such inhomogeneous systems are considered to lase by EHP in segmented QWs or Q disks (energy tail states). It is desirable to use entire QW planes as gain media for higher performance LD operation. Note that use of InGaN QWs with small potential fluctuation results in lasing operation while GaN QWs, which are supposed to have homogeneous potentials, have not shown electrically pumped lasing.

ACKNOWLEDGMENTS

The authors are grateful to Dr. T. Uenoyama, Dr. M. Sugawara, Dr. S. Keller, Dr. J. Speck, Dr. T. Azuhata, T. Deguchi, and Professor K. Suzuki for stimulating discussions. The authors would like to thank Dr. S. Shirakata for instructions about PR measurements. They are also grateful to Professor H. Nakanishi, Professor H. Ikoma, Professor U. K. Mishra, Professor L. A. Coldren, and Professor S. P. DenBaars for continuous encouragement.

[1]Important data and references are cited in S. Nakamura and G. Fasol, *The Blue Laser Diode* (Springer, Berlin, 1997); recent data are from S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, Y. Sugimoto, and H. Kiyoku, Jpn. J. Appl. Phys., Part 2 **36**, L1059 (1997).
[2]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, Y. Sugimoto, T. Kozaki, H. Umemoto, M. Sano, and K. Chocho, Jpn. J. Appl. Phys., Part 2 **36**, L1568 (1997); Appl. Phys. Lett. **72**, 211 (1998); S. Nakamura, The 2nd International Conference on Nitride Semiconductors, Tokushima, Japan, 27–31 October 1997; they adopted modulation-doped superlattice cladding layers and ELOG substrates and the acceleration testing was done at 50°.
[3]I. Akasaki, S. Sota, H. Sakai, T. Tanaka, M. Koike, and H. Amano, Electron. Lett. **32**, 1105 (1996).
[4]K. Itaya, M. Onomura, J. Nishio, L. Sugiura, S. Saito, M. Suzuki, J. Rennie, S. Nunoue, M. Yamamoto, H. Fujimoto, Y. Kokubun, Y. Ohba, G. Hatakoshi, and M. Ishikawa, Jpn. J. Appl. Phys., Part 2 **35**, L1315 (1996).
[5]G. E. Bulman, K. Doverspike, S. T. Sheppard, T. W. Weeks, H. S. Kong, H. M. Dieringer, J. A. Edmond, J. D. Brown, J. T. Swindell, and J. F. Schetzina, Electron. Lett. **33**, 1556 (1997).
[6]A. Kuramata, K. Domen, R. Soejima, K. Horino, S. Kubota, and T. Tanahashi, Jpn. J. Appl. Phys., Part 2 **36**, L1130 (1997).
[7]M. P. Mack, A. Abare, M. Aizcorbe, P. Kozodoy, S. Keller, U. K. Mishra, L. A. Coldren, and S. P. DenBaars, Mater. Res. Soc. Internet J. Nitride Semicond. Res. **2**, 41 (1997).
[8]F. Nakamura, T. Kobayashi, T. Asatsuma, K. Funato, K. Yanashima, S. Hashimoto, K. Naganuma, S. Tomioka, T. Miyajima, E. Morita, H. Kawai, and M. Ikeda, Proceedings of the 2nd International Conference on Nitride Semiconductors, Tokushima, Japan, 1997, p. 460.
[9]M Kreissl, D. P. Bour, N. Johnson, L. Romano, B. Krusor, R. Donaldson, J. Walker, and C. Dunnrowicz, Appl Phys. Lett. **72**, 1539 (1998).
[10]K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. **46**, 3432 (1975).
[11]R. Singh, D. Doppalapudi, T. D. Moustakas, and L. Romano, Appl. Phys. Lett. **70**, 1089 (1997).
[12]A. Koukitsu, N. Takahashi, T. Taki, and H. Seki, Jpn. J. Appl. Phys. Part 2 **35**, L673 (1996); A. Koukitsu and H. Seki, in Ref. 8, p. 18.
[13]I. Ho and G. B. Stringfellow, Appl. Phys. Lett. **69**, 2701 (1996).
[14]T. Matsuoka, Appl. Phys. Lett. **71**, 105 (1997); in Ref. 8, p. 20.
[15]J. S. Speck, X. H. Wu, P. Fini, E. J. Tarsa, B. Heying, S. Keller, U. K. Mishra, and S. P. DenBaars, in Ref. 8, p. 208.
[16]M. Shimizu, Y. Kawaguchi, K. Hiramatsu, and N. Sawaki, Solid-State Electron. **41**, 145 (1997); Jpn. J. Appl. Phys., Part 1 **36**, 3381 (1997).
[17]H. Sato, T. Sugahara, Y. Naoi, and S. Sakai, in Ref. 8, p. 28.
[18]F. Ponce, S. Galloway, W. Götz, and R. Kern, Mater. Res. Soc. Symp. Proc. **482**, 625 (1998); L. Romano, M. McCluskey, B. Krusor, D. Bour, and C. Chua, in Ref. 8, p. 32.
[19]S. Chichibu, T. Azuhata, T. Sota, and S. Nakamura, Appl. Phys. Lett. **70**, 2822 (1997); in Ref. 8, p. 224; S. Chichibu, T. Deguchi, T. Sota, K. Wada, and S. Nakamura, Mater. Res. Soc. Symp. Proc. **482**, 613 (1998).
[20]A. Wakahara, T. Tokuda, X. Dang, S. Noda, and A. Sasaki, Appl. Phys. Lett. **71**, 906 (1997).
[21]F. Scholz, V. Härle, F. Steuber, A. Sohmer, H. Bolay, V. Syganow, A. Doernen, J. Im, A. Hangleiter, J. Y. Duboz, P. Galtier, E. Rosencher, O. Ambacher, D. Brunner, and H. Lakner, Mater. Res. Soc. Symp. Proc. **449**, 3 (1997); F. Scholz, J. Off, A. Sohmer, V. Syganow, A. Dörnen, and O. Ambacher, in Ref. 8, p. 14.
[22]S. Chichibu, M. Arita, H. Nakanishi, J. Nishio, L. Sugiura, Y. Kokubun, and K. Itaya, J. Appl. Phys. **83**, (1998).
[23]S. Chichibu, T. Azuhata, T. Sota, and S. Nakamura, 38th Electronic Materials Conference, 26–28 June Santa Barbara, CA; Appl. Phys. Lett. **69**, 4188 (1996); Mater. Res. Soc. Symp. Proc. **449**, 653 (1997); in Ref. 8, p. 224; S. Chichibu, T. Deguchi, T. Sota, K. Wada, and S. Nakamura, Mater. Res. Soc. Symp. Proc. **482**, 613 (1998).
[24]C. Kisielowski and Z. Liliental-Weber (private communication); they have observed a dotlike nanoscale compositional disorder in the SQW LED wafers using the electron scattering potential mapping method.
[25]Y. Narukawa, Y. Kawakami, Sz. Fujita, Sg. Fujita, and S. Nakamura, Phys. Rev. B **55**, R1938 (1997).
[26]Y. Narukawa, Y. Kawakami, M. Funato, Sz. Fujita, Sg. Fujita, and S. Nakamura, Appl. Phys. Lett. **70**, 981 (1997).
[27]S. Chichibu, K. Wada, and S. Nakamura, Appl. Phys. Lett. **71**, 2346 (1997); S. Chichibu, K. Wada, M. Arita, T. Sota, and S. Nakamura, in Ref. 8, p. 186.
[28]J. Im, S. Heppel, H. Kollmer, A. Sohmer, J. Off, F. Scholz, and A. Hangleiter, in Ref. 8, p. 228.
[29]T. Takeuchi, H. Takeuchi, S. Sota, H. Sakai, H. Amano, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **36**, L177 (1997); T. Takeuchi, S. Sota, H. Sakai, H. Amano, and I. Akasaki, in Ref. 8, p. 418.
[30]T. Honda, T. Miyamoto, T. Sakaguchi, H. Kawanishi, F. Koyama, and K. Iga, in Ref. 8, p. 74.
[31]S. Chichibu, T. Azuhata, T. Sota and S. Nakamura, J. Appl. Phys. **79**, 2784 (1996); Proceedings of the International Symposium on Blue Laser and Light Emitting Diodes, Ohmsha, Tokyo, 1996, p. 202; S. Chichibu, H. Okumura, S. Nakamura, G. Feuillet, T. Azuhata, T. Sota, and S. Yoshida, Jpn. J. Appl. Phys., Part 1 **36**, 1976 (1997).
[32]B. Monemar, J. P. Bergman, H. Amano, I. Akasaki, T. Detchprohm, K. Hiramatsu, and N. Sawaki, in Proceedings in Ref. 31, p. 135.
[33]M. Smith, G. Chen, J. Y. Lin, H. Jiang, M. A. Khan, and Q. Chen, Appl. Phys. Lett. **69**, 2837 (1996).
[34]K. Domen, A. Kuramata, and T. Tanahashi, Appl. Phys. Lett. **72**, 1359 (1998).
[35]M. Sugawara, Phys. Rev. B **51**, 10 743 (1995).
[36]For a review see, for example, H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (World Scientific, Singapore, 1990) and W. W. Chow, S. W. Koch, and M. Sargent III, *Semiconductor-Laser Physics* (Springer, Berlin, 1994); W. W. Chow, A. F. Wright, and J. S. Nelson, Appl. Phys. Lett. **68**, 296 (1996).
[37]M. Cardona, in *Modulation Spectroscopy*, Solid State Physics Suppl. 11, edited by S. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1969); D. E. Aspnes, *Handbook on Semiconductors*, edited by T. S. Moss (North-Holland, Amsterdam, 1980), Vol. 2, Chap. 4A, p. 109; D. E. Aspnes, Surf. Sci. **37**, 418 (1973).
[38]K. L. Shaklee and R. F. Leheny, Appl. Phys. Lett. **18**, 475 (1971); K. L. Shaklee, R. F. Leheny, and R. E. Nahory, *ibid.* **19**, 302 (1971).
[39]S. Chichibu, A. Shikanai, T. Azuhata, T. Sota, A. Kuramata, K. Horino, and S. Nakamura, Appl. Phys. Lett. **68**, 3766 (1996); A. Shikanai, T. Azuhata, T. Sota, S. Chichibu, A. Kuramata, K. Horino, and S. Nakamura, J. Appl. Phys. **81**, 417 (1997); S. Chichibu, T. Azuhata, T. Sota, H. Amano, and I. Akasaki, Appl. Phys. Lett. **70**, 2085 (1997).
[40]T. Deguchi, T. Azuhata, T. Sota, S. Chichibu, and S. Nakamura, Mater. Sci. Eng. B **50**, 251 (1997); T. Deguchi, T. Azuhata, T. Sota, S. Chichibu, M. Arita, H. Nakanishi, and S. Nakamura, Semicond. Sci. Technol. **13**, 97 (1998).
[41]T. Azuhata, T. Sota, K. Suzuki, and S. Nakamura, J. Phys.: Condens. Matter **7**, L129 (1995); T. Azuhata, T. Matsunaga, K. Shimada, K.

Yoshida, T. Sota, K. Suzuki, and S. Nakamura, Physica B **219/220**, 493 (1996).
[42]S. Chichibu, T. Mizutani, T. Shioda, H. Nakanishi, T. Deguchi, T. Azuhata, T. Sota, and S. Nakamura, Appl. Phys. Lett. **70**, 3440 (1997); M. Arita, S. Chichibu, H. Nakanishi, T. Deguchi, T. Azuhata, T. Sota, and S. Nakamura, in Ref. 8, p. 426.
[43]M. Suzuki, T. Uenoyama, and A. Yanase, Phys. Rev. B **52**, 8132 (1995).
[44]R. Dingle, D. D. Sell, S. E. Stokowski, and M. Ilegems, Phys. Rev. B **4**, 1211 (1971).
[45]B. Monemar, Phys. Rev. B **10**, 676 (1974).
[46]S. Shirakata and S. Chichibu, J. Appl. Phys. **80**, 2043 (1996).
[47]D. A. Miller, D. S. Chemla, T. C. Damen, A. C. Gross, W. Wiegmann, T. H. Wood, and C. A. Burrus, Phys. Rev. Lett. **53**, 2173 (1981); Phys. Rev. B **32**, 1043 (1985).
[48]D. A. B. Miller, D. S. Chemla, and S. Schmitt-Rink, Phys. Rev. B **33**, 6976 (1986).
[49]L. V. Keldysh, Sov. Phys. JETP **7**, 788 (1958); W. Franz, Z. Naturforsch. A **13A**, 484 (1958).
[50]W. Shan, B. Little, J. Song, Z. Feng, M. Schuman, and R. Stall, Appl. Phys. Lett. **69**, 3315 (1996).
[51]C. Wetzel, T. Takeuchi, H. Amano, and I. Akasaki, in Ref. 8, p. 100.
[52]C. I. Harris, B. Monemar, H. Amano, and I. Akasaki, Appl. Phys. Lett. **67**, 840 (1995).
[53]T. Taguchi, T. Maeda, Y. Yamada, S. Nakamura, and G. Shinomiya, in Proceedings in Ref. 31, p. 372.
[54]J. Im, V. Härle, F. Scholz, and A. Hangleiter, Mater. Res. Soc. Internet J. Nitride Semicond. Res. **1**, 37 (1996).
[55]T. Taguchi, 43rd Spring Meeting of the Japan Society of Applied Physics and Related Societies, Asaka, Japan, 29 March 1996 (unpublished).
[56]H. Sakai, T. Takeuchi, S. Sota, M. Katsuragawa, M. Komori, H. Amano, and I. Akasaki, in Ref. 8, p. 456.
[57]D. L. Smith and C. Mailhiot, Phys. Rev. Lett. **58**, 1264 (1987).
[58]K. Torii, S. Chichibu, T. Sota, and S. Nakamura (unpublished); the piezoelectric field was calculated with the values of piezoelectric constants of GaN according to M. P. Halsall, J. E. Nicholls, J. J. Davies, B. Cockayne, and P. J. Wright [J. Appl. Phys. **71**, 907 (1992)], and the Stark shift induced by the electric field was calculated by variational method neglecting exciton binding energy. The calculations are based on Ref. 47. Finally, the degree of Coulomb screening of the piezoelectric field was estimated by solving the Schrödinger and Poisson equations simultaneously self-consistently.
[59]S. Keller, B. P. Keller, M. S. Minsky, J. E. Bowers, U. K. Mishra, S. P. DenBaars, and W. Seifert, in Ref. 8; (private communication).
[60]T. Uenoyama, Phys. Rev. B **51**, 10 228 (1995); (private communication).
[61]H. Kwon, Y. Lee, O. Miki, H. Yamano, and A. Yoshida, Appl. Phys. Lett. **69**, 937 (1996).
[62]E. E. Mendez, G. Bastard, L. L. Chang, L. Esaki, H. Morkoç, and R. Fisher, Phys. Rev. B **26**, 710 (1982).
[63]M. Yamanishi and I. Suemune, Jpn. J. Appl. Phys., Part 2 **22**, L22 (1983).
[64]G. Bastard, E. E. Mendez, L. L. Chang, and L. Esaki, Phys. Rev. B **26**, 1974 (1982).
[65]J. S. Weiner, D. A. B. Miller, D. S. Chelma, T. C. Damen, C. A. Burrus, and T. H. Wood, Appl. Phys. Lett. **47**, 1148 (1985).
[66]T. Sota, S. Chichibu, and S. Nakamura (private communication). The E_{ex} value in the $GaN/Al_{0.1}Ga_{0.9}N$ QW was calculated by the variational method according to Ref. 64. In the calculation, we started from the Hamiltonian suitable for QWs including mass and dielectric constant anisotropies, assuming an infinite barrier height to simplify the calculation. After variable transformations for Z coordinates (perpendicular to the QW plane) which formally remove the mass anisotropy in the Hamiltonian, E_{ex} was calculated using the trial function for the excitonic relative motion given by $\exp\{-[\rho^2+(Z_e-Z_h)^2]^{1/2}/\lambda\}$, taking into consideration the fact that the 3D exciton Bohr radius a_B is as small as 3.4 nm. Here λ is the variational parameter, ρ is the absolute value of the relative position of the electron and the hole in the QW plane, and $Z_e(Z_h)$ is the transformed Z coordinate of the electron (hole).
[67]T. Uenoyama, Mater. Res. Soc. Symp. Proc. **482**, 805 (1998).
[68]G. Frankowsky, F. Steuber, V. Härle, F. Scholz, and A. Hangleiter, Appl. Phys. Lett. **68**, 3746 (1996).
[69]M. Kuball, E. Jeon, Y. Song, A. Nurmikko, P. Kozodoy, A. Abare, S. Keller, L. A. Coldren, U. K. Mishra, S. P. DenBaars, and D. A. Steigerwald, Appl. Phys. Lett. **70**, 2580 (1997).
[70]W. Chen, M. Fritze, A. V. Nurmikko, D. Ackley, C. Covard, and H. Lee, Phys. Rev. Lett. **64**, 2434 (1990).
[71]J. F. Mueller, Phys. Rev. B **42**, 11 189 (1990).
[72]G. Mohs, T. Aoki, M. Nagai, R. Shimano, M. Kuwata-Gonokami, and S. Nakamura, Solid State Commun. **104**, 643 (1997).

Dynamics of localized excitons in InGaN/GaN quantum wells

Hongbin Yu, H. Htoon, Alex deLozanne, and C. K. Shih
Department of Physics, The University of Texas at Austin, Austin, Texas 78712

P. A. Grudowski and R. D. Dupuis
Microelectronics Research Center, The University of Texas at Austin, Austin, Texas 78712

K. Zeng, R. Mair, J. Y. Lin, and H. X. Jiang
Department of Physics, Kansas State University, Manhattan, Kansas

(Received 21 January 1998; accepted 21 May 1998)

By using a "nanoscale aperture" method, we studied the dynamics of localized excitons in active InGaN layers in an InGaN/GaN multiple quantum well (MQW). Local photoluminescence (PL) from local excitation shows very different characteristics from far-field luminescence. PL spectra from these nanoapertures are, in general, blueshifted relative to the spectra acquired in the unmasked regions. The spectra vary from aperture to aperture, reflecting the spatial fluctuation of the In composition in the MQW. In addition, one typically observes spectra with clearly resolved multiple peaks from the small apertures, which may be attributed to compositional and size fluctuations over a length scale longer than the exciton diameter but smaller than the aperture size. The excitation intensity dependence of PL indicates a band-filling effect in the multiple-peak structure. Finally, time-resolved studies reveal a very interesting spectral weight shifting between the higher-energy peaks and the lower-energy peaks. © *1998 American Vacuum Society.* [S0734-211X(98)08104-9]

The recent breakthrough in the development of nitride-based blue and green light-emitting diodes (LEDs) and laser diodes (LDs) opens a new era in modern optoelectronic technology.[1–3] However, many important materials issues remain to be investigated. The lack of detailed understanding of the optoelectronic processes in these GaN-based devices presents scientific challenges and research opportunities. An important issue is related to the InGaN active layers in InGaN/GaN quantum wells.[1] Due to the large mismatch of the lattice constant between GaN and InN, it is very difficult to grow homogeneous InGaN layers. Phase separation into regions of different In composition is the rule rather than exception.[4,5] In such highly inhomogeneous active layers, optoelectronic processes such as carrier transport and exciton dynamics are highly complex. It has been suggested that the localized In-rich regions act as quantum dots (QDs), which are responsible for the lasing action in GaN-based LDs.[6–8]

The samples used in this study are InGaN multiple quantum wells (MQWs) grown on a 1.5-μm-thick GaN:Si layer deposited on a sapphire (0001) substrate by low-pressure metalorganic chemical vapor deposition. The MQW structure consists of five periods of 3.5-nm-thick $In_{0.13}Ga_{0.87}N$ Si-doped quantum wells and 7-nm-thick $In_{0.03}Ga_{0.97}N$ Si-doped barrier layers.[9,10] A 16-nm GaN:ud (undoped GaN) cap is grown on top of the MQW structures. The dynamic behavior of localized excitons in the InGaN MQW is studied using a "nanoaperture" method. Such a method has been used previously by others to study optical properties of single GaAs QDs.[11,12] However, instead of using electron-beam lithography, here we use a much simpler way to create nanoapertures by utilizing polystyrene (PS) beads.

The fabrication of the nanoscale apertures is shown schematically in Fig. 1. A suspension 0.3- or 0.6-μm-diam polystyrene beads diluted with alcohol was dropped onto the sample. After the alcohol evaporated, sparsely distributed PS beads were then left on the sample surface. After the deposition of 100 nm Al thin films followed by the removal of the PS beads using dichloromethane solution and ultrasonic agitation, an Al metal film with small apertures is formed. Most of the apertures have the same size as the diameter of the PS bead as revealed by the atomic force microscope (AFM) image (Fig. 2). However, some apertures are larger with an apparent shape due to the aggregate of a few PS beads. Using this method, we create apertures with size ranging from 0.3 to 3 μm. The holes were spaced sufficiently far apart to allow optical probing of a single hole. The photoluminescence (PL) was excited by temperatures from 10 to 300 K through a given hole with a picosecond laser system, and the PL was detected from the same hole with 1.33-m-long spectrometer and microchannel-plate photomultiplier tube together with a single-photon counting system.[5] Hence, the excitation and PL detection was performed via the same aperture.

The cw PL spectra acquired at 10 K from the submicron apertures and the spectrum from the unmasked region are shown in Fig. 3. It can be clearly seen that the PL peak positions of the masked region shift to higher energy as compared to the unmasked region. This blueshift trend can be explained in the context of exciton diffusion out of the aperture region, from higher-energy sites to lower-energy sites. Excitons localized in random wells induced by composition fluctuation will transfer from higher-energy site to lower-energy site by thermal relaxation. Those excitons which diffuse out of the aperture then will have lower energies and will not be detected due to the opaque Al cover. While for the unmasked sample, PL is dominant from the lowest-energy region due to the localization effect. Thus, PL emitted from an aperture will have a relatively higher-energy portion,

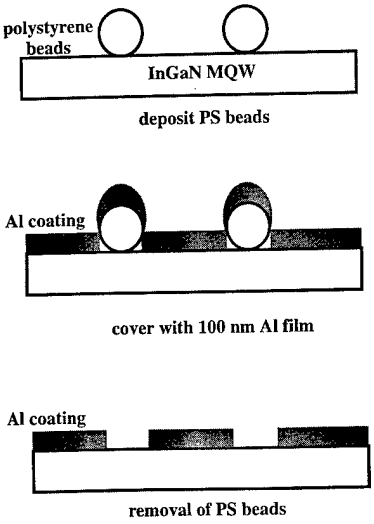


FIG. 1. Schematic diagrams of the preparation of the nanoscale apertures.

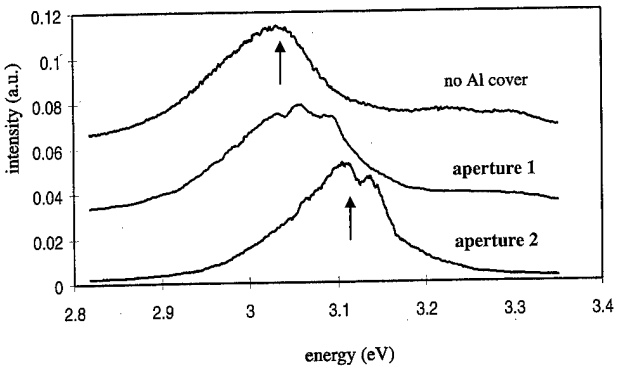


FIG. 3. Photoluminescence spectra taken from unmasked and masked regions.

i.e., exhibit a blueshift. In the masked sample, if there is no such potential difference, or no composition fluctuation when excitons diffuse out of the aperture, there will be no peak position shift of the spectra. Therefore, blueshift of the spectra indicates there is a composition fluctuation in the MQW and excitons are localized to the lower-band-gap regions of MQW corresponding to higher In composition. It also indicated that the exciton diffusion and localization process covers at least 1 μm.

One also sees in Fig. 3 multiple peaks on the spectra from the masked area. The appearance of several PL peaks in the spectra is not due to an interference effect, because there is no spacing in the structure that will give rise to the peak energy separation in the wavelength region we are studying. These peaks are also not phonon replicas, as has been observed in other studies on InGaN/GaN and GaN/AlGaN MQWs. The longitudinal optical phonon energies of InN and GaN are 86 and 91 meV, respectively,[13] while the energy separations between neighboring peaks in our spectra are, typically, about 30 meV, though the separations vary from one to another. The multiple-peak structure may be attributed to the compositional and size fluctuation over a length scale longer than the exciton diameter but smaller than the aperture size. It is interesting to note that in contrast to the Gammon and co-workers work, where sharp peaks from single GaAs quantum dots with a linewidth as small as 80 μeV can be resolved,[11,12] here the smallest linewidth is about 20 meV. The reasons for this may be due to the following. First, the size of the aperture here is not as small as in the Gammon and co-workers work. Second, the Bohr radius of the excitons in GaN is about 3.4 nm,[4] much smaller than that in GaAs, i.e., 20 nm.[11] Third, in their work, the excitons are formed due to the well width fluctuations of the quantum wells in GaAs, which are of a lateral size larger than 20 nm.[11] In InGaN MQW the composition fluctuations are expected to be of the size of several nanometers, consequently, the exciton localization effect will be much stronger than that in GaAs quantum dots, and give rise to broader peaks due in part to localization and in part due to phonon coupling to higher-energy phonons in InGaN.

Figure 4 shows PL spectra from the same aperture at 10 K for two different excitation intensities. The excitation intensity I_{exc} is proportional to 10^{-D}. The arrows indicate the multiple peak positions of the PL spectra. It is clear that the multiple emission peak positions remain at the same energies when changing the intensity, while the spectral weight shifts toward higher energies with increasing intensity. This excitation intensity dependence indicates a band-edge-filling effect of the multiple-peak structure, where the band edge is due to the local fluctuation of In composition.

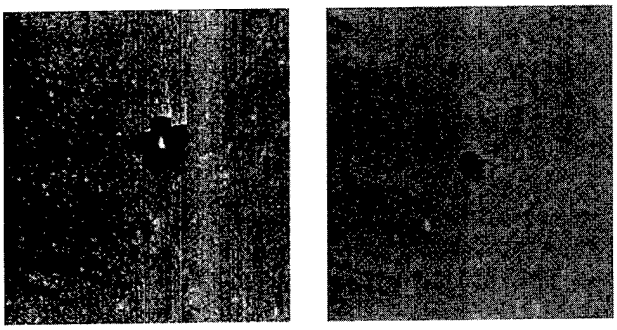


FIG. 2. 2.5 μm×5 μm AFM images showing apertures fabricated using PS beads of 0.3 μm size. Most of the apertures made are 0.3 μm, in some cases, apertures are of 1 μm size due to clustering of the PS beads.

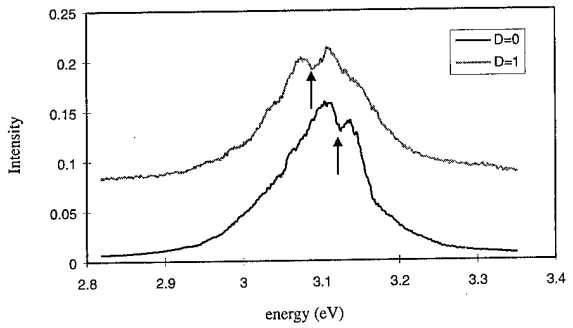


FIG. 4. Photoluminescence spectra from the same aperture with different excitation intensities.

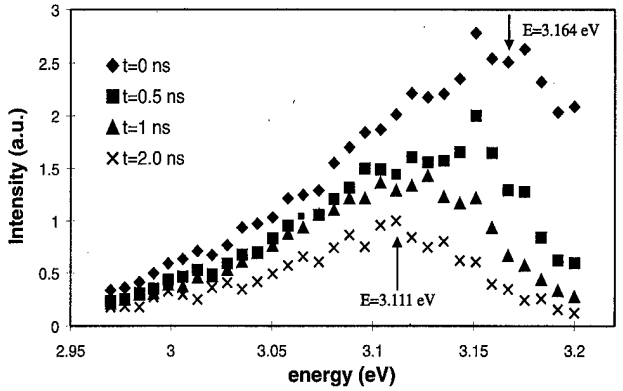


FIG. 5. Time-resolved photoluminescence spectra from the same aperture as in Fig. 4 showing the spectral shifting from high energy to lower energy with increasing delay time.

Time-resolved PL spectra have also been measured on both masked and unmasked regions to study the dynamics of the localized excitons. Figure 5 shows time-resolved emission spectra of the 3.164 eV emission line measured on the same aperture as in Fig. 4 at 10 K at several representative delay times. The arrows in Fig. 5 indicate the spectral peak positions at different delay times. It can be seen that the spectral peak position shifts toward lower energies as the delay time increases and the linewidth of the emission line also increases with delay times. This luminescence spectral shift with delay time is a direct consequence of a strong emission energy dependence of the recombination lifetime. If the optical transition involves only a single exciton energy level, the spectral shift with delay time and the variation of the recombination lifetime with emission energy are not expected. Thus, these results indicate that localized exciton energy levels in the aperture under study have a distribution. With the increasing delay time, the localized excitons are transferred from high-energy levels to low-lying levels within the aperture region.

Two important effects caused by this alloy fluctuation and local excitation/detection of microapertures have to be considered. First, in energy space, excitons relax towards the lower-energy regions due to the exciton localization effect. Second, in real space, excitons simultaneously diffuse away from the aperture due to the presence of the microapertures. Thus, only those excitons which radiatively recombine within the apertures will be detected in this special local excitation/detection configuration. This is the key difference between our results obtained here and previous results of unmasked samples. For samples without apertures, the cw PL emission spectra are always dominated from those excitons localized in the lower-energy regions, reflecting the fact that exciton localization and diffusion are very fast processes.

In conclusion, a "microaperture" method has been employed to study the dynamics of the localized excitons in InGaN/GaN MQWs. The cw spectra from masked and unmasked regions revealed that excitons are localized into spatially different regions due to the phase separation into different In composition regions in InGaN and the resulting composition fluctuation in quantum wells. The blueshift trend and the spectral shift with delay time have demonstrated that the localized excitons are transferred spatially from high potential energy sites to low potential sites and temporally from high-energy levels to low-energy levels.

The authors thank Professor Josef Kas for his help in making these nanoapertures. The work at The University of Texas is supported by NSF-DMR-9402938, NSF-Science and Technology Center, Grant No. CHE 89-20120, the State of Texas Advanced Technology Program, the Army Research Office under DAAH04-93-0317, and DARPA under Contract No. MDA972-95-3-0008. The work at Kansas State University is supported by ARO, BMDO/ONR, DOE No. (93ER45604/A000), and NSF(DMR) 9528226.

[1]S. Nakamura and G. Fasol, *The Blue Laser Diode* (Springer, Berlin, 1997), and references therein.
[2]S. N. Mohammad and H. Morkoc, Prog. Quantum Electron. **20**, 361 (1996).
[3]*Gallium Nitride and Related Materials*, Materials Research Society Symposium Proceedings, edited by F. A. Ponce, R. D. Dupuis, S. Nakamura, and J. A. Edmund (Materials Research Society, Pittsburgh, PA, 1996), Vol. 395.
[4]S. Chichibu, T. Azuhata, T. Sota, and S. Nakamura, Appl. Phys. Lett. **69**, 4188 (1996).
[5]M. Smith, G. D. Chen, J. Y. Lin, H. X. Jiang, M. Asif Khan, and Q. Chen, Appl. Phys. Lett. **69**, 2837 (1996).
[6]S. Chichibu, K. Wada, and S. Nakamura, Appl. Phys. Lett. **71**, 2346 (1997).
[7]Y. Narukawa, Y. Kawakami, S. Fujita, S. Fujita, and S. Nakamura, Phys. Rev. B **55**, R1938 (1997).
[8]Y. Narukawa, Y. Kawakami, M. Funato, S. Fujita, S. Fujita, and S. Nakamura, Appl. Phys. Lett. **70**, 981 (1997).
[9]P. A. Grudowski, C. J. Eiting, J. Park, B. S. Shelton, D. J. H. Lambert, and R. D. Dupuis, Appl. Phys. Lett. **71**, 1537 (1997).
[10]P. A. Grudowski, C. J. Eiting, and R. D. Dupuis, J. Cryst. Growth (to be published).
[11]D. Gammon, E. S. Snow, and D. S. Katzer, Appl. Phys. Lett. **67**, 2391 (1995).
[12]D. Gammon, E. S. Snow, B. V. Shanabrook, D. S. Katzer, and D. Park, Science **273**, 87 (1996).
[13]M. Smith, J. Y. Lin, H. X. Jiang, A. Khan, Q. Chen, A. Salvador, A. Botchkarev, W. Kim, and H. Morkoc, Appl. Phys. Lett. **70**, 2882 (1997).

Investigation of the chemistry and electronic properties of metal/gallium nitride interfaces

C. I. Wu and A. Kahn[a)]
Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544

(Received 21 January 1998; accepted 26 May 1998)

We present a systematic investigation of the formation of Schottky barriers between *n*- and *p*-GaN(0001)−(1×1) grown by metalorganic chemical vapor deposition and a series of high and low work function metals (Mg, Al, Ti, Au, and Pt). Al, Ti, and Mg react at room temperature with nitrogen, whereas Au and Pt form abrupt, unreacted interfaces. We find that the Fermi level movement on both *n*- and *p*-GaN is consistent with variations in metal work functions, but limited by surface or interface states. Upon annealing, the incorporation of Mg increases the density of acceptors as seen on both *n*- and *p*-GaN. In spite of similar work functions and chemical reaction with nitrogen, Ti and Al show drastic differences in Schottky barrier formation due to differences in the nature of the products of reaction. AlN is a wide band gap semiconductor whereas TiN is a metallic compound. © *1998 American Vacuum Society.* [S0734-211X(98)10504-8]

I. INTRODUCTION

The structural and electronic quality of GaN and related compounds has been substantially improved during the past few years, leading to major advances in blue–green light emitting devices. GaN-based light emitting diodes have been commercialized and blue laser diodes, with lifetimes of 10 000 h under continuous wave (cw) mode operation at 20 °C, have been projected.[1] However, the performance of GaN-based laser diodes is still limited by several materials problems related to substrates and to the difficulty in making low resistance ohmic contacts, especially to *p*-GaN. The large ionization energy of this semiconductor causes difficulty in making low resistance contacts to *p*-GaN by just using high work function metals.[2,3] The highest *p*-doping levels achieved to date are too low for making contact via tunneling through the depletion region. For *n*-GaN, although several reacted metal–nitride interfaces have been found to produce reasonably good contacts,[4–6] the relationship between injection and interfaces electronic structure is still not clear. Furthermore, the interface chemistry, which has direct bearing on the stability of these devices, has not been explored in detail. Only a few studies on the chemistry and electronic properties at metal–GaN interfaces have been reported.[7–9]

We present a systematic investigation of the formation of Schottky barriers between GaN and a series of metals (Mg, Al, Ti, Au, and Pt) with a range of low to high work functions. Clean and ordered (1×1) surfaces are obtained by nitrogen ion sputtering and annealing (ISA). The chemistry and electronic properties of the surfaces and interfaces are studied by low energy electron diffraction (LEED), Auger electron spectroscopy (AES), x-ray photoemission spectroscopy (XPS), and ultraviolet photoemission spectroscopy (UPS). We find that Al, Ti, and Mg react at room temperature with GaN, whereas Au and Pt form abrupt, unreacted interfaces. The Fermi level movements on *n*- and *p*-type GaN are found to be consistent with the changes in the metal work function, but are limited by the interface states. Chemical reactions also play a dominant role in the Schottky barrier behavior.

II. EXPERIMENT

N- and *p*-type wurzite GaN layers were grown by metalorganic chemical vapor deposition (MOCVD) on SiC substrates at Philips Laboratories, Briarcliff Manor, NY. The orientation of the layers is labeled (0001), even though the polarity of these MOCVD-grown nitride layers remains uncertain. The *n*-type layers were 1 μm thick and doped with 1×10^{19} cm^{-3} Si, whereas the *p*-type layers were 0.1 μm thick and doped with 5×10^{17} cm^{-3} Mg. After removal from the growth reactor, the samples were stored in air for periods of days, then loaded in an ultrahigh vacuum (UHV) system (base pressure of 10^{-10} Torr) equipped with LEED, AES, XPS, and UPS. The samples were cradled in tantalum foils which could be heated resistively up to 1000 °C. The substrate temperature in UHV was measured by a thermocouple and an infrared pyrometer. The GaN(0001)-(1×1) surfaces were prepared by multiple cycles of nitrogen ion sputtering (0.5 keV) and annealing (900 °C, 10 min).[10] All the metals were deposited by thermal evaporation.

The AES, UPS, and XPS spectra were recorded with a double-pass cylindrical mirror analyzer. The GaN surface composition was estimated from the ratio of the 1070 eV Ga *LMM* and 379 eV N *KLL* AES peak-to-peak amplitudes corrected for elemental sensitivities.[10] The valence band spectra were recorded using the He II photon line (40.8 eV) from a gas discharge lamp. The He pressure in the analysis chamber during UPS measurement was kept in the low 10^{-9} Torr range. The onset of photoemission, which corresponds to the vacuum level (E_{vac}) of the sample, was measured with a −3 V bias on the sample to clear the detector work function. The overall resolution of the UPS measurement was 0.15 eV. The valence band maximum (VBM) was measured by linear extrapolation of the high kinetic energy

a)Corresponding author; electronic mail: kahn@ee.princeton.edu

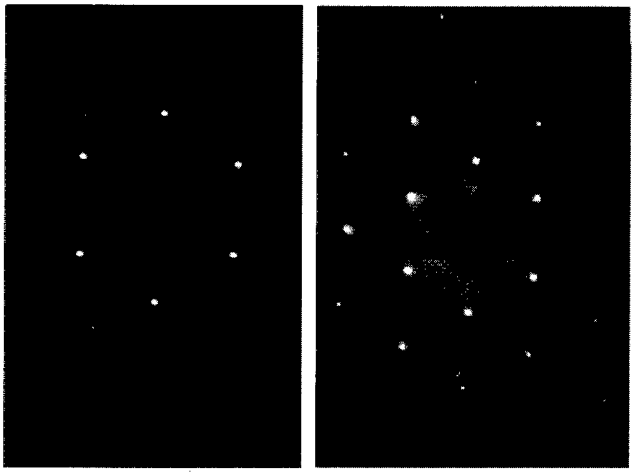


FIG. 1. LEED pattern from the clean GaN(0001)-(1×1) surface for primary beam energy equal to (a) 96 and (b) 149 eV.

side of the valence band on spectra corresponding to clean and ordered surfaces. The band bending caused by metal deposition was evaluated from the displacement of the Ga $3d$ core level measured using the $M\zeta$ photon line (151.4 eV) from the Zr anode of the x-ray source. The absolute resolution of the XPS measurement was 0.8 eV, but the accuracy of the core level displacement was 0.1 eV. The position of the Fermi level (E_F) was measured by UPS on clean gold films. All the measurements were done at room temperature.

III. RESULTS AND DISCUSSION

A. GaN(0001)-(1×1) surface

The as-loaded GaN surface is disordered and heavily contaminated by carbon and oxygen. Following several ISA cycles, the surface exhibits a sharp 1×1 LEED pattern with very low background (Fig. 1). The Ga-to-N AES peak ratio levels off to a value which we have previously associated with a nearly stoichiometric surface.[10] This preparation produces surfaces which appear to be comparable to those prepared under Ga flux.[11] The (1×1) structure is unexpected in view of the fact that GaN(0001) is a polar surface and, according to electron counting arguments,[12] should reconstruct to lower its energy and remain semiconducting. Whether the observed structure is a real or a slightly disordered (1×1) structure, and whether the surface is truly semiconducting or impurity stabilized is unknown at this point. However, a recent scanning tunneling microscopy study has confirmed the existence of an ordered (1×1) phase and has suggested that it corresponds to a Ga layer on top of a N-terminated surface.[13,14]

Valence band spectra from n-GaN following only two cycles of ISA and from the clean and ordered (1×1) surface are shown in Fig. 2. The peak at 12.8 eV below E_F corresponds to the Ga $3d$ core level excited by the He III (48.4 eV) photon line. The spectra in Figs. 2(a) and 2(b) are nearly identical, except for a residual shoulder at the top of the valence band indicated by the arrow in Fig. 2(a). Given that

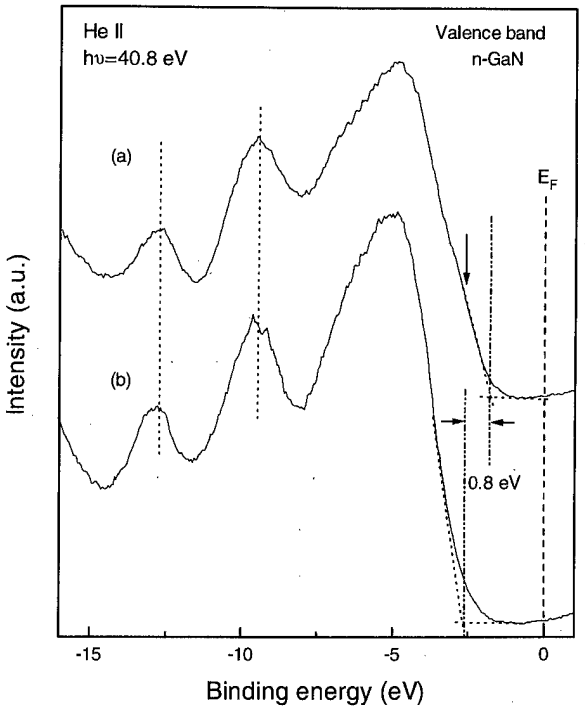


FIG. 2. UPS valence band spectra for a (a) contaminated and nonstoichiometric and (b) clean and ordered (1×1) n-type GaN surface. The binding energy is referenced to the Fermi level. The vertical arrow indicates a surface state-related feature.

the positions of the main valence features and of the Ga $3d$ core level are identical on these two surfaces, we attribute the shoulder to surface states. AES shows that the surface [Fig. 2(a)] still contains carbon and oxygen and is N deficient (Ga-to-N ratio of about 1.2). The observed surface states at or near the VBM are therefore contamination related and, if not eliminated, lead to errors in the determination of the VBM. The VBM determined by linear extrapolation of the high kinetic energy edge of the valence band on spectrum (b) of Fig. 2 is at 2.6 eV below E_F, whereas it is nearly 1 eV higher on spectrum (a) of Fig. 2. Given a 3.4 eV band gap and the fact that E_F is approximately 50 meV below the conduction band minimum in the bulk of this n-type material, spectrum (b) leads to a value of 0.75±0.1 eV for the upward band bending at the clean and ordered n-type surface. This value is in reasonable agreement with previous results.[11] The ionization energy (IE=$E_{vac}-E_v$) is 6.9 ±0.1 eV and the electron affinity (EA) is 3.5±0.1 eV, also in good agreement with previous results.[11] Finally, the binding energy of the Ga $3d$ core level with respect to the VBM is 17.8 eV.

The p-type surface was studied according to the same procedure. Spectrum (a) in Fig. 3 shows that surface states are again detectable near the VBM on the contaminated and nonstoichiometric surface. The VBM of the clean surface [spectrum (b) of Fig. 3] is found at 1.0 eV below E_F. With E_F about 250 meV above the VBM in the bulk of the sample, that is, the Mg acceptor level in GaN,[15] the downward band bending is estimated to be 0.75±0.1 eV. The

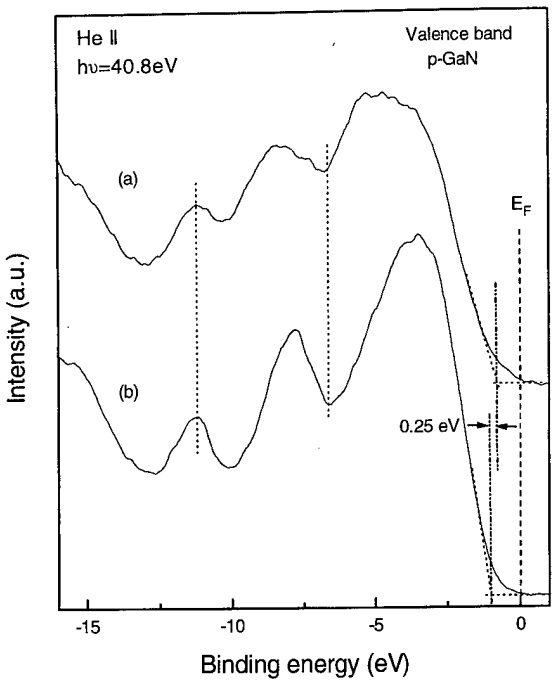
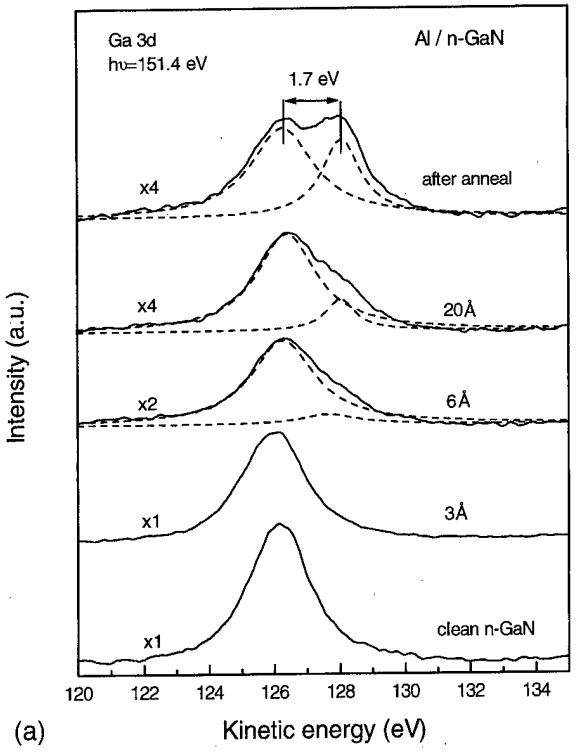


FIG. 3. Same data as in Fig. 2 for a p-type GaN surface.

measured ionization energy and Ga $3d$ binding energy are in complete agreement with the values found for n-GaN.

B. Al/GaN and Ti/GaN

The Al/GaN interface undergoes a Ga–Al exchange reaction driven by the large heat of formation of AlN as compared to that of GaN (see Table I). In the process, Ga is released from the substrate. A free Ga component, shifted by ~1.7 eV toward lower binding energy with respect to the bulk component, is clearly seen for Al coverages larger than 6 Å (Fig. 4). At low Al coverage, the Al $2p$ core level appears in a high binding energy position corresponding to reacted Al (presumably with N), indicating a strong interaction with the substrate and a two-dimensional initial growth. A weak metallic Al $2p$ component shifted by 2.0 eV with respect to the reacted component appears at ~10 Å, pointing out the beginning of three-dimensional growth and clustering on top of the reacted layer. A Fermi edge is observed in the UPS spectra for coverages above 10 Å. The metallic component dominates the Al $2p$ spectrum at high coverage. Annealing the interface at high temperature results in an enhancement of the chemical reaction and the nearly complete elimination of the metallic Al component. The free Ga component is increased at annealing temperatures below 900 °C, but is eliminated due to re-evaporation beyond 900 °C. The Ti/GaN interface leads to an almost identical result from the point of view of chemical reaction.



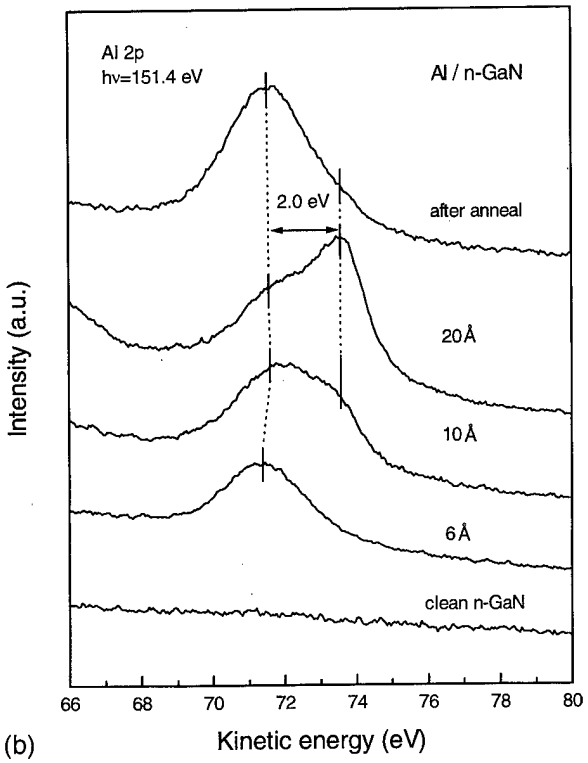


FIG. 4. (a) Ga $3d$ core level and (b) Al $2p$ core level as a function of Al deposition on n-GaN. The Ga-in-GaN and free Ga, and the metallic and reacted Al components are indicated. The annealing was done at 900 °C for 1 min.

TABLE I. Heat of formation of various nitrides (after Ref. 16).

	GaN	AlN	TiN	Mg_3N_2
ΔH (kJ/mol)	−110.9	−318.1	−265.5	−461.1

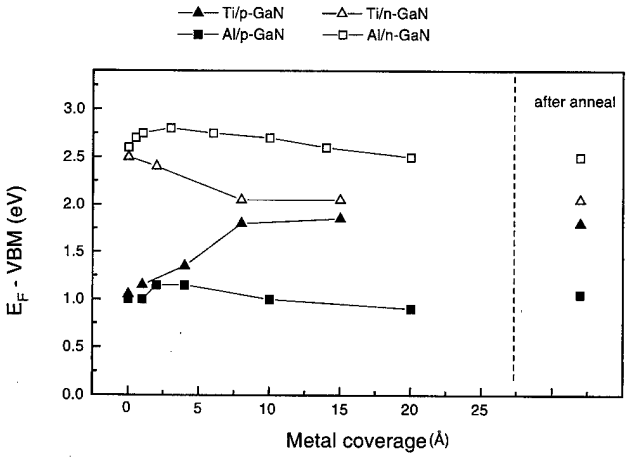


FIG. 5. Movement of the surface Fermi level on n- and p-GaN(0001) as a function of Al and Ti coverage; the right-hand side panel shows positions following annealing at 900 °C for 1 min.

The movement of E_F at the n- and p-GaN surfaces as a function of Al and Ti coverage is determined by XPS and is shown in Fig. 5. For both n- and p-GaN, E_F moves by less than 0.15 eV throughout the deposition of the first 20 Å of Al. On the other hand, Ti causes an additional band bending of 0.5 eV for n-GaN and 0.8 eV for p-GaN in spite of having a similar metal work function (4.28 eV for Al and 4.33 eV for Ti) and undergoing similar chemical reactions with GaN. The difference can be explained by comparing the electrical properties of the products of the reaction, i.e., AlN and TiN. As soon as Al is evaporated onto GaN, it forms AlN, an insulator with a band gap of 6.2 eV. As the Al coverage increases, the interface consists of a three-component metal–insulator–semiconductor structure, Al/AlN/GaN, and the Al metal-induced states are decoupled from the semiconductor. In the case of Ti/GaN, the reaction product is TiN, a metallic compound with conductivity 1.3 times that of Ti. The interface is a simple metal–semiconductor system and the Fermi levels of n- and p-GaN converge at high Ti coverage to nearly the same energy position with respect to valence band maximum. The relation between Fermi level movement and interfaces states will be discussed in Sec. III E. Annealing after metal deposition only enhances the chemical reaction but does not affect the relative band positions of either Al or Ti.

C. Mg/GaN

The Mg/GaN interface shows both similarities and differences with the two preceding cases. Like Al and Ti, due to large differences in heats of formation between Mg_3N_2 and GaN (see Table I), Mg replaces Ga at the GaN surface, as is evidenced by the appearance of a free Ga component in the Ga 3d core level spectrum [Fig. 6(a)]. Note that the intensity of the Ga 3d core level is considerably attenuated by the Mg layer on the 15 and 25 Å spectra. Because of its low work function (3.7 eV), Mg pulls E_F upward at the interface by about 0.95 eV on p-GaN, but does not substantially affect the band position on n-GaN. The initial 3 Å Mg peak in Fig.

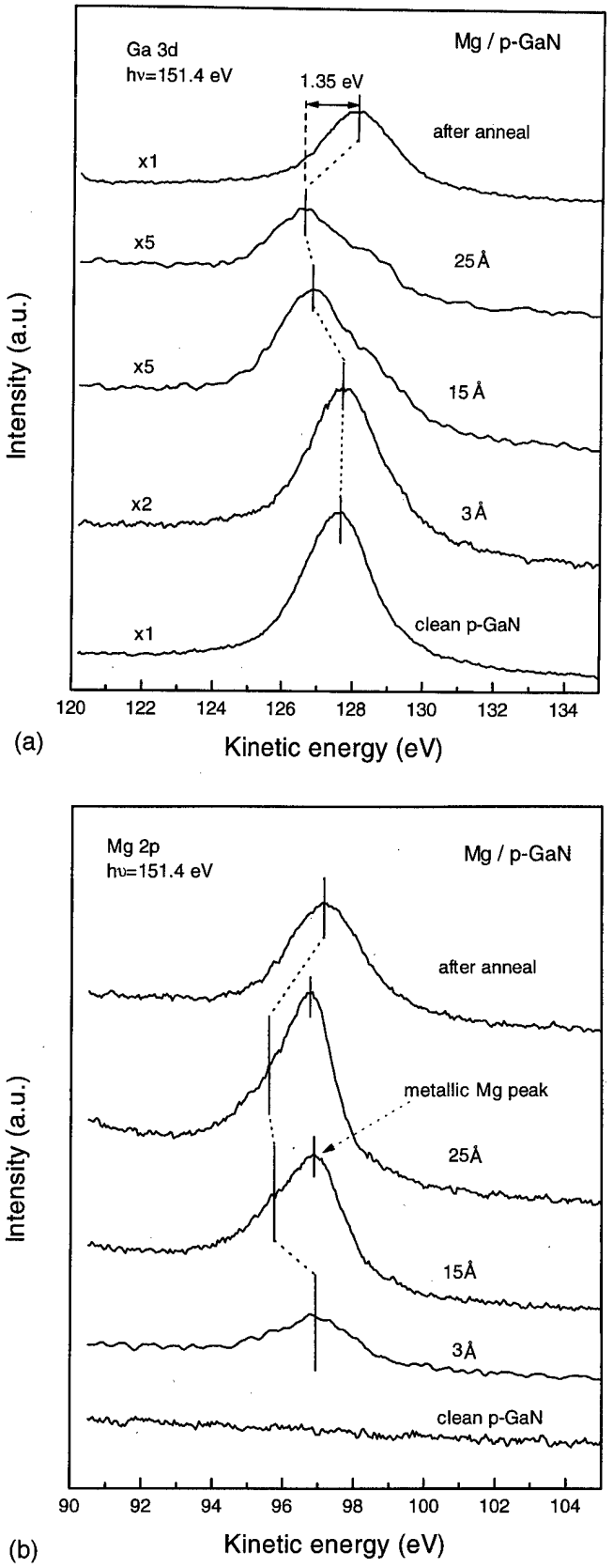


FIG. 6. (a) Ga 3d core level and (b) Mg 2p core level as a function of Mg deposition on p-GaN. The metallic and reacted Mg components are indicated. The annealing was done at 900 °C for 1 min.

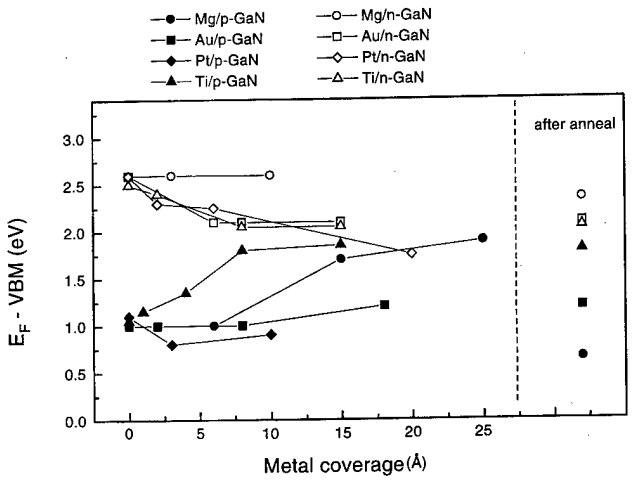


FIG. 7. Movement of the surface Fermi level on n- and p-GaN as a function of Mg, Ti, Au, and Pt coverage. The final stabilized position gives the Schottky barrier height. The right-hand side panel shows positions following annealing at 900 °C for 1 min.

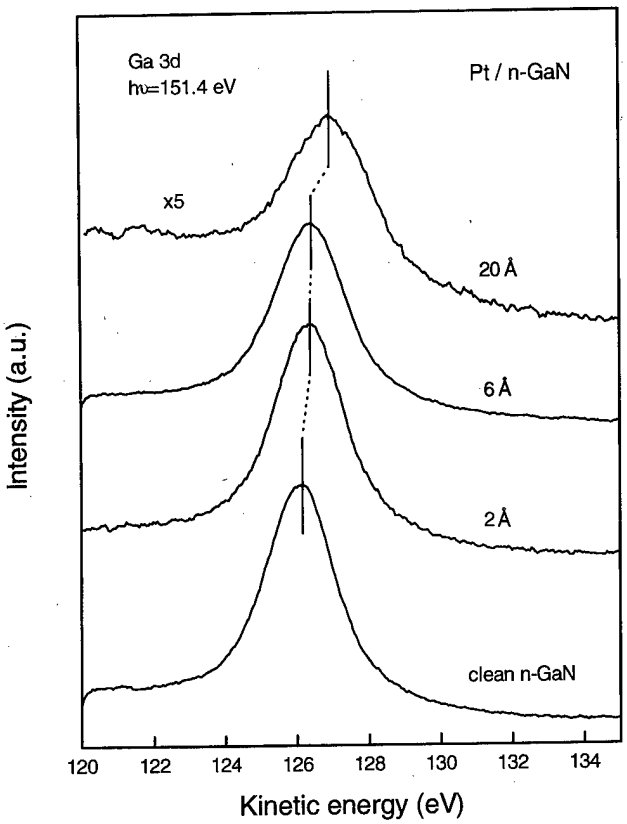


FIG. 8. Ga $3d$ core level as a function of Pt deposition on n-GaN.

6(b) corresponds to Mg reacted with N. The 15 Å spectrum is dominated by the metallic Mg $2p$ core level and a Fermi edge is clearly seen in the UPS spectrum. The high binding energy reacted component is shifted by 0.95 eV with respect to the 3 Å spectrum. This shift represents the band bending induced by metallic Mg, as already seen on the Ga $3d$ core level [Fig. 6(a)]. Unlike with Al and Ti, however, annealing the sample after Mg deposition not only enhances the Mg–GaN reaction but also drastically changes the E_F position. The Ga $3d$ and reacted Mg $2p$ core levels shift toward higher kinetic energy by about 1.35 eV, as the sample is annealed at 900 °C [Figs. 6(a) and 6(b)]. Concomitantly, E_F moves down in the gap by the same amount (Fig. 7). As evidenced by the Mg signal, there is still a substantial amount of Mg near the GaN surface following annealing, even though the annealing temperature is considerably higher than the sublimation temperature of Mg (~250 °C).[17] This indicates that part of the Mg diffuses into the GaN bulk and is not re-evaporated. Furthermore, E_F at the annealed Mg/p-GaN interface is 0.35 eV lower than at the clean p-GaN surface, indicating that the incorporation of Mg increases the density of acceptors in GaN. For the same reason, E_F also shifts downward by 0.25 eV at the Mg/n-GaN interface after annealing (Fig. 7).

D. Pt/GaN and Au/GaN

Pt and Au show considerably less chemical reactivity with GaN than the three metal described above. Figure 8 presents the evolution of the Ga $3d$ core level as a function of Pt deposition. No free Ga component is detectable at this interface. In addition, the Pt $4f$ core level appears from the first deposition at the energy of the metallic Pt and exhibits the expected spin–orbit splitting (not shown here). The interface is therefore unreacted and abrupt. Pt and Au are high work function metals and cause 0.9 and 0.5 eV downward movements of E_F on n-GaN, respectively. The E_F movements are relatively small for both metals on p-GaN. The Schottky barrier heights evaluated at high metal coverage on p-GaN are 0.9 and 1.2 eV for Pt and Au, respectively, consistent with the known difficulty in making low resistance ohmic contacts to p-GaN, even using high work function metals like Pt and Au.

E. Fermi level position and interface states

Figures 5 and 7 show the E_F position at n- and p-GaN(0001) interfaces as a function of metal coverage. Several remarks can be made. First, the sputtered and annealed surfaces, although atomically well ordered, include defects which introduce both donor and acceptor states and cause a ~0.8 eV band bending on both n- and p-GaN(0001). Second, the surface E_F position (and thus the Schottky barrier) depends on the metal work function. There is a ~1.0 eV difference in barrier heights between Mg and Pt on n- and p-GaN. The barrier dependence, however, appears to be limited by states which are induced by the metal deposition and which cause the convergence of the E_F at about 2.0 eV above the VBM for a number of interfaces. Third, the difference between the Al and Ti results is indicative of the nature of these states. As discussed above, Al and Ti undergo very similar reactions with the substrate and have an essentially identical work function. However, the former gives AlN, an insulator, and the latter TiN, a metal. With Ti, E_F converges to 2.0 eV above the VBM, whereas a 1.7 eV split

remains between the E_F positions on n- and p-GaN with Al. Such behaviors suggests that metal-induced gap states play an important role in the definition of the barriers, and tend to stabilize E_F near 2.0 eV above the VBM. With the density of metal-induced gap states being relatively small in wide-gap semiconductors,[18] the low and high work function metals, e.g., Mg, Au, and Pt, are able to pull E_F above or below the 2.0 eV mark, respectively. The final E_F position is therefore likely to be the result of several competitive processes involving states due to defects, to products of reactions, and to the metal–semiconductor boundary conditions.

IV. SUMMARY

We used photoemission spectroscopy to investigate the chemistry and electronic properties of the GaN(0001)-(1 $\times$1) surface and interfaces with various metals. Mg, Al, and Ti were found to react with GaN at room temperature, whereas Au and Pt were found to produce abrupt and unreacted interfaces. A substantial dependence of the Fermi level position on the metal work function was found, in agreement with previous studies on wide-gap semiconductors. The Ti/GaN and Al/GaN interfaces differ significantly in spite of similarities in the metal work function and chemical reaction with nitrogen. The difference is accounted for by the nature of the reaction products (AlN and TiN) and by the presumed role of metal-induced gap states in the definition of Schottky barriers. Upon annealing, the incorporation of Mg increases the acceptor density as seen in both n- and p-GaN. Finally, we verified that high work function metals, i.e., Au and Pt, stabilize the Fermi levels around 1 eV above the valence band maximum of p-GaN, giving large barriers for holes and exemplifying the difficulty of making ohmic contacts to this material.

ACKNOWLEDGMENTS

This work was supported by a grant of the National Science Foundation (DMR97-18771). The authors gratefully acknowledge Dr. Nikkil Taskar in Philips Laboratory for providing high quality GaN samples.

[1]S. Nakamura, 2nd International Conference on Nitride Semiconductors, Tokushima, Japan, 1997.
[2]Y. Koide, H. Ishikawa, S. Kobayashi, S. Yamasaki, S. Nagai, J. Umezaki, M. Koike, and M. Murakami, Appl. Surf. Sci. **117/118**, 373 (1997).
[3]H. Ishikawa, S. Kobayashi, Y. Koide, S. Yamasaki, S. Nagai, J. Umezaki, M. Koike, and M. Murakami, J. Appl. Phys. **81**, 1315 (1997).
[4]Z. Fan, S. N. Mohammad, W. Kim, O. Aktas, A. E. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **68**, 1672 (1996).
[5]J. Burm, K. Chu, W. A. Davis, W. J. Schaff, and T. J. Eustis, Appl. Phys. Lett. **70**, 464 (1997).
[6]Y. F. Wu, W. N. Jiang, B. P. Keller, S. Keller, D. Kapolnek, S. P. DenBaars, U. K. Mishra, and B. Wilson, Solid-State Electron. **41**, 165 (1997).
[7]B. P. Luther, S. E. Mohney, T. N. Jackson, M. A. Khan, Q. Chen, and J. W. Yang, Appl. Phys. Lett. **70**, 57 (1997).
[8]V. M. Bermudez, T. M. Jung, K. Doverspike, and A. E. Wickenden, J. Appl. Phys. **79**, 110 (1996).
[9]T. U. Kampen and W. Mönch, Appl. Surf. Sci. **117/118**, 388 (1997).
[10]C. I. Wu and A. Kahn, J. Appl. Phys. **83**, 4249 (1998).
[11]V. M. Bermudez, J. Appl. Phys. **82**, 1190 (1996).
[12]M. D. Pashley, Phys. Rev. B **40**, 10 481 (1989).
[13]A. R. Smith, V. Ramachandran, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, J. Vac. Sci. Technol. A **16**, 1641 (1998).
[14]A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, Appl. Phys. Lett. **72**, 2114 (1998).
[15]S. Strite and H. Morkoç, J. Vac. Sci. Technol. B **10**, 1237 (1992).
[16]O. Kubaschewski, C. B. Alock, and P. J. Spencer, *Materials Thermochemistry* (Pergamon, Oxford, 1993).
[17]R. E. Honig and P. A. Kramer, RCA Rev. **30**, 285 (1969).
[18]S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976).

Negative electron affinity of cesiated *p*-GaN(0001) surfaces

M. Eyckeler and W. Mönch
Laboratorium für Festkörperphysik, Gerhard-Mercator-Universität Duisburg, D-47048 Duisburg, Germany

T. U. Kampen[a)]
Professur für Halbleiterphysik, TU Chemnitz, D-09107 Chemnitz, Germany

R. Dimitrov, O. Ambacher, and M. Stutzmann
Walter Schottky Institut, Technische Universität München, D-85748 Garching, Germany

(Received 20 January 1998; accepted 20 May 1998)

The adsorption of cesium on clean *n*- and *p*-GaN(0001)-1×1 surfaces at 150 K was investigated using x-ray photoemission spectroscopy, photoemission spectroscopy with monochromatized He I radiation ultraviolet photoelectron spectroscopy (UPS) and a Kelvin probe (contact potential difference, CPD). The CPD measurements gave work functions of 3.88±0.15 and 3.6±0.15 eV for clean *n*- and *p*-GaN(0001) surfaces, respectively. The widths of UPS energy distribution curves yield an ionization energy of 6.8±0.15 eV. Thus, depletion and inversion layers exist at clean surfaces of *n*- and *p*-GaN(0001) surfaces, respectively. As a function of Cs coverage, the work function displays the well-known behavior in that it first decreases, passes through a minimum, and eventually reaches a value of 2.1 eV, the work function of metallic cesium. In the submonolayer coverage regime, the ionization energy decreases by 2.3±0.15 eV. At clean *p*-GaN(0001) surfaces the vacuum level lies by only 0.3 eV above the conduction-band minimum in the bulk. Already minute amounts of Cs suffice to produce negative electron affinity. The Schottky barrier height of metallic Cs films on *n*-GaN(0001) is determined as 0.2±0.15 eV. This value is in good agreement with what is predicted by the MIGS and electronegativity model. © *1998 American Vacuum Society.* [S0734-211X(98)09004-0]

I. INTRODUCTION

A most challenging proposal is the development of superflat field-emission displays based on group-III nitrides.[1] Such devices are based on the concept of negative electron affinity (NEA).[2] Quite generally, the electron affinity of a clean or an adsorbate-covered semiconductor surface is defined as the energy distance between the vacuum level and the bottom of the conduction band right at the surface. Provided the surface band bending is such that the vacuum level is below the conduction-band minimum in the bulk then the surface is said to have a NEA. Such surfaces are efficient electron emitters.

The electronic surface properties of semiconductors are generally modified by adsorbate-induced surface dipoles and surface states.[3] Adatoms form covalent bonds with surface atoms of the substrate. Such bonds are partly ionic and the bond charge is displaced towards the more electronegative atom. The respective adatom-induced surface dipoles will lower or increase the electron affinity of the semiconductor depending on whether the surface bond charge is shifted towards the surface atom or towards the adatom. Adatom-induced surface states will have either acceptor or donor character. Their charging state varies from neutral to negative and from positive to neutral, respectively, when their energy levels move from above to below the Fermi level. A space charge below the surface compensates the net charge in surface states. Space charges of positive and negative sign are correlated with an upward and a downward band bending, respectively.

Adatom-induced surface dipoles and band bending may move the vacuum level at the surface to below the conduction-band minimum in the bulk although right at the surface the vacuum level is still above the conduction-band bottom. This situation is commonly referred to as NEA. Electrons excited over the band gap in the space charge region and in the bulk will be freely emitted into vacuum provided they are not scattered to energies below the vacuum level. NEA is commonly achieved on samples doped *p* type.

Hydrogen has been reported to create NEA on (111)[4–6] and (001)[7,8] surfaces of *p*-diamond. Based on *ab initio* modeling, the change in the position of the conduction-band edge with respect to the vacuum level was believed to be the result of surface dipoles induced by the electropositive H adatoms.[7] Küttel *et al.*[9] evaporated cesium on (100) surfaces of *p*-diamond and investigated the resulting NEA with ultraviolet photoemission spectroscopy and electron field emission measurements. Moreover, nitrides such as AlN (Ref. 10) and $Al_xGa_{1-x}N$ (Ref. 11) for $x \geq 0.75$ were reported to exhibit NEA.

Previously, we studied the change of the electronic properties of *n*-GaN(0001) surface as the function of Cs coverage.[12] As to be expected, the Cs-induced surface dipoles reduced the work function. Our present investigations are focused on the electronic properties of Cs-covered *p*-GaN(0001) surfaces. Our experimental results clearly reveal cesiated *p*-GaN(0001) surfaces to exhibit NEA. Furthermore, we determined the surface band bending at

[a)]Author to whom correspondence should be addressed; electronic mail: kampen@physik.tu-chemnitz.de

Cs/*p*-GaN(0001) interfaces. The respective Schottky barrier height confirms a prediction resulting from the theory of interface-induced gap states.

II. EXPERIMENT

The *p*-type GaN(0001) layers were grown on *c*-plane sapphire by plasma-induced molecular beam epitaxy (PIMBE). Ga and Mg effusion cells and a rf plasma source (Oxford Applied Research CARS) providing nitrogen radicals were used in the deposition process. The layers were grown at 810 °C with a growth rate of ~0.5 μm/h and without a buffer layer. The layer thickness of 500 nm was determined by scanning electron microscopy with an accuracy of ±10 nm.[13] The hole concentration of 1.2×10^{17} cm^{-3} was determined from Hall measurements. As *n*-type substrates we used GaN(0001) epilayers on sapphire with a thickness of ~1.5 μm obtained from Cree Research Inc. The doping level was specified as below 5×10^{16} cm^{-3}.

To obtain clean surfaces the following procedures were applied. The samples were dipped for 1 min into hydrofluoric acid (HF), rinsed in de-ionized water, blown dry with N_2 gas, and immediately transferred into a ultrahigh vacuum (UHV) system via a fast load-lock. There they were heated to a temperature of about 800 °C with a rate lower than 1.5 °C/s. At 800 °C they were exposed to a flux of 1×10^{16} Ga atoms per cm^2/s for 10 min and subsequently annealed at the same temperature for another 30 min. The samples were then carefully cooled down to room temperature. This procedure gave clean and well-ordered surfaces as was monitored by using x-ray photoelectron and Auger electron spectroscopy (XPS, AES) and low-energy electron diffraction (LEED), respectively.

For the deposition of Cs and the subsequent measurements, the samples were cooled to 150 K. Cesium was evaporated from dispensers (SAES Getters) which were thoroughly degassed. They were heated with a constant current of 6 A while the durations of Cs exposures were controlled by a shutter. The Cs uptake of the samples was monitored by recording x-ray photoemission spectra.

The Kelvin method was used to measure the contact potential between the GaN surfaces under study and a vibrating probe. The work function of the Kelvin probe itself was calibrated as 4.28±0.1 eV utilizing cleaved *p*-GaAs(110) surfaces as reference. Their work function is known to be insensitive to cleavage-induced defects. The Kelvin probe was operated in the dark to avoid any surface photovoltage.

In the submonolayer coverage range, the ionization energy was determined from the width of the energy distribution curves (EDCs) of photoemitted electrons which were excited by using monochromatized He I radiation, i.e., 21.2 eV photons.

III. RESULTS

The adsorption of Cs on GaN(0001) at 150 K was followed by XPS. Figure 1 shows the ratio of the emission intensities from the Cs($3d_{3/2}$) and the Ga($3p$) core levels as a function of Cs exposure. The data points may be fitted by two straight lines. Such sequence of straight-line segments indicates layer-by-layer or Frank–van der Merwe growth. The break at approximately 400 s is then assigned to the completion of the first continuous adsorbate layer. Layer-by-layer growth of cesium at low temperatures is well known from other semiconductors.[14,15] The evaporation time necessary for deposition of the first complete cesium layer varied with the specific Cs dispenser used. During the formation of the second layer the Cs film becomes metallic. A Fermi edge was observed in ultraviolet photoemission spectra (UPS) excited with monochromatized He I radiation.

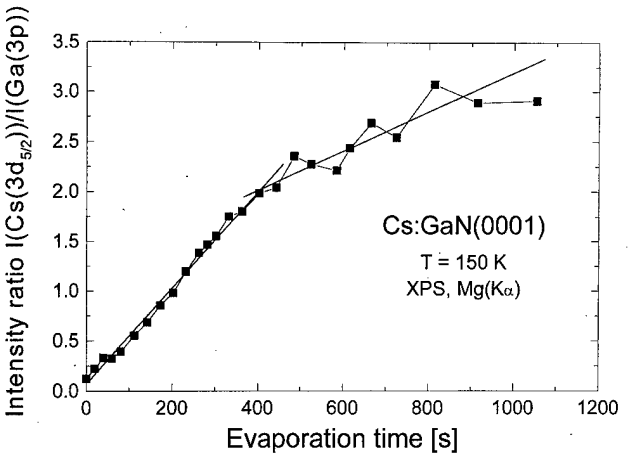


FIG. 1. Ratio of emission intensities of Cs($3d_{3/2}$) and Ga($3p$) core levels recorded with a *n*-GaN(0001) surface kept at 150 K as a function of the Cs evaporation time.

Figure 2 displays the kinetic energies of the N($1s$), Ga($3p_{1/2}$), Ga($3p_{3/2}$), and Cs($3d_{3/2}$) core-level signals as a function of Cs exposure. Initially, the core signals of the substrate atoms move to higher kinetic energies. For exposures larger than 250 s the kinetic energies off all core-level signals decrease by up to 0.5 eV and eventually become constant for exposures larger than 500 s, i.e., after the completion of the first continuous Cs layer. Core-level shifts may be caused by surface band bending or chemical bonding. Since the N($1s$), Ga($3p_{1/2}$), Ga($3p_{3/2}$), and Cs($3d_{3/2}$) core-level signals behave the same their shifts are attributed to Cs-induced changes in surface band bending.

Using a Kelvin probe the work functions of clean and subsequently Cs-covered *n*- and *p*-GaN(0001) surfaces were measured. The results are plotted in Fig. 3. Clean *n*- and *p*-GaN(0001) surfaces have a work function of 3.88±0.15 and 3.6±0.15 eV, respectively. As a function of Cs coverage, the work functions first decrease, then pass through a minimum, and eventually reach a constant value of 2.1 ±0.1 eV irrespective of the type of doping. This final value agrees with the work function, 1.81 eV (Ref. 16) and 2.1 eV (Ref. 17) reported for metallic cesium.

Figure 4 displays some of the EDCs recorded with clean and subsequently Cs-covered *n*- and *p*-GaN(0001) surfaces. The ionization energy I of a surface under study may be determined from photoemission spectra. The width W of the EDCs of the photoemitted electrons equals the difference of

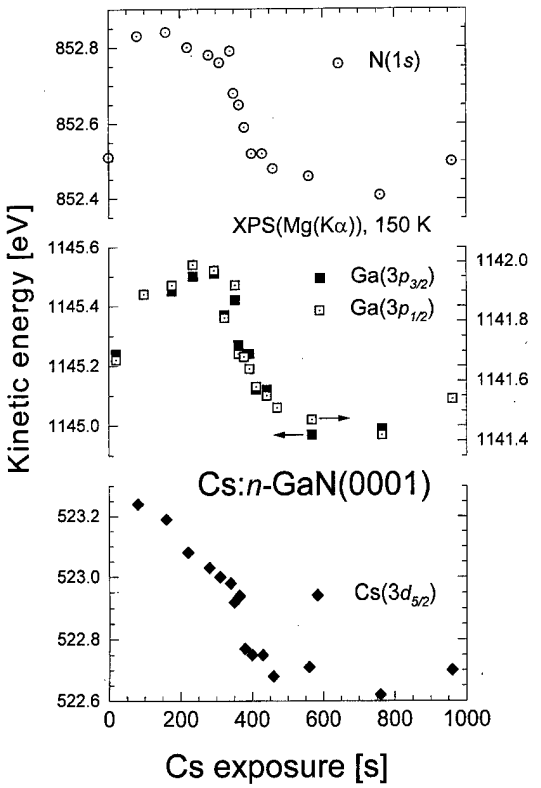


FIG. 2. Energy position of the N($1s$), Ga($3p_{1/2}$), Ga($3p_{3/2}$), and Cs($3d_{3/2}$) core-level signals as a function of Cs exposure.

the photon energy $\hbar\omega$ of the exciting light and the ionization energy, i.e.,

$$W=\hbar\omega-I. \tag{1}$$

Using Eq. (1) and the spectra displayed in Fig. 4 one obtains the ionization energy of clean GaN(0001) surfaces as 6.8 $\pm$0.1 eV, irrespective of the type of bulk doping. The ionization energy and the electron affinity of a semiconductor differ by the band gap of the semiconductor W_g which is 3.45 eV for GaN at 150 K.[18] Then the electron affinity of clean GaN(0001) surfaces results as 3.35$\pm$0.1 eV. Our finding is in good agreement with the values of 3.2 and 3.3 eV reported by Bermudez[19] and Benjamin,[10] respectively. With

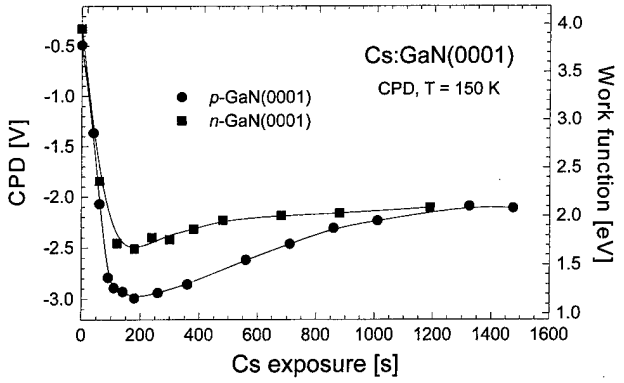


FIG. 3. Work function of n- and p-GaN(0001) surfaces as a function of the Cs evaporation time at 130 K.

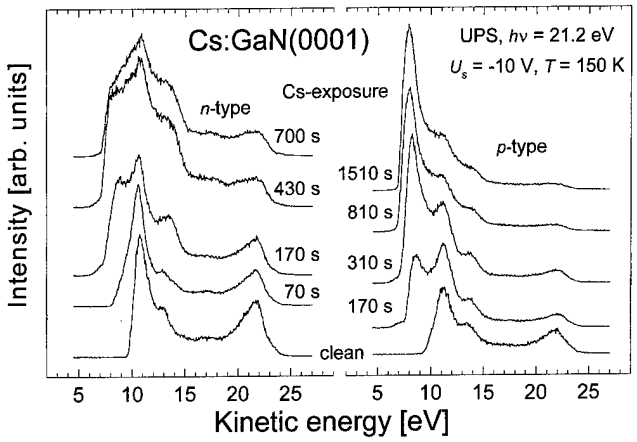


FIG. 4. EDCs of photoemitted electrons excited by monochromated He I radiation from cleaned n- and p-GaN(0001) surfaces and after successive exposures to increasing amounts of Cs at 150 K.

increasing Cs coverage the ionization energy decreases. After the completion of the first continuous Cs layer it reaches a value of 4.5$\pm$0.15 eV, again irrespective of the type of bulk doping. Consequently, the electron affinity now amounts to 1.05$\pm$0.15 eV.

Three features of the EDCs displayed in Fig. 4 will be considered in more detail. First, the intensity of the intense structure slightly below the valence-band maximum (VBM) decreases as a function of Cs deposition. This behavior is typical for the emission of electrons from surface states that are removed by the adsorption of adatoms. Thus we attribute this structure to occupied surface states just below the valence-band top. Second, the feature at 11 eV below the VBM of the clean surfaces occurs at the same kinetic energy in photoemission spectra excited with He II radiation. Therefore, we attribute this structure to emission of secondary electrons which accumulate in a region of high density of states at an energy of $\hbar\omega-W_g-11$ eV=6.8 eV above the conduction-band minimum. Third, the peak of the true secondaries observed with the clean surfaces apparently remains present after the subsequent Cs exposures. Therefore, we conclude from this observation that there is another region of high density of states in the range of 4.9 eV above the conduction-band minimum.

IV. DISCUSSION

A. Position of the Fermi level on clean and Cs-covered GaN(0001) surfaces

The energy distribution of surface states determines the surface band bending or, in other words, the energy position of the Fermi level with respect to the valence-band maximum. The work function Φ and the ionization energy I are defined as the energy distance from the Fermi level W_F and the valence-band maximum W_{vs}, right at the surface, respectively, to the vacuum level. Therefore, the position of the Fermi level with respect to the valence-band maximum at the surface is given by

$$W_F-W_{vs}=I-\Phi. \tag{2}$$

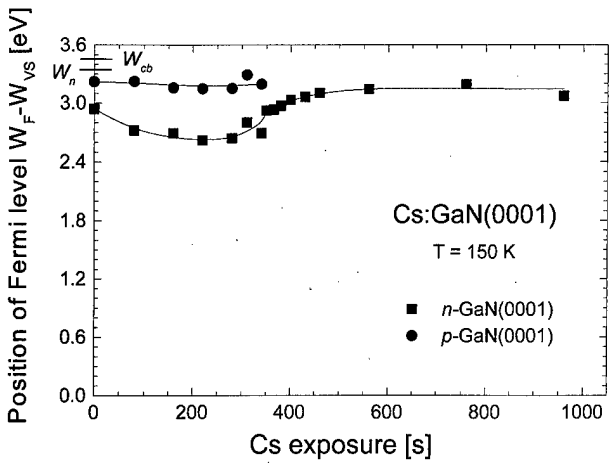


FIG. 5. Energy difference between Fermi level and valence-band maximum at clean and cesiated *n*- and *p*-GaN(0001) surfaces at 150 K as a function of Cs exposure.

On clean *n*- and *p*-GaN(0001) surfaces the Fermi level is positioned 2.94 and 3.22 eV above the valence-band maximum, respectively. Thus, depletion layers exist on *n*- but strong inversion layers on *p*-GaN(0001) surfaces.

The core-level shifts displayed in Fig. 2 are attributed to Cs-induced variations in surface band bending on *n*-GaN(0001) surfaces. They thus allow for a determination of the Fermi-level position relative to the valence-band maximum on Cs-covered *n*-GaN(0001) surfaces. The respective data are displayed in Fig. 5. In the submonolayer coverage range, the energy position of the Fermi level on Cs-covered *p*-GaN(0001) surfaces, on the other hand, may be determined using the work function data displayed in Fig. 3. On clean as well as on Cs-covered surfaces, the ionization energies are the same on *n*- and *p*-type samples. Therefore, the differences of the work functions of *n*- and *p*-type samples measured at identical Cs coverages equal the differences between the Fermi-level positions with regard to the valence-band maximum on *n*- and *p*-GaN(0001) surfaces. As displayed in Fig. 5 the band bending on *p*-type GaN is constant while it is increased for *n*-type GaN. For metallic Cs films the position of the Fermi level is the same for both types of doping and constant. The barrier heights for Cs:*n*-GaN(0001) and Cs:*p*-GaN(0001) contacts are 0.2 $\pm$0.15 and 3.2$\pm$0.15 eV, respectively.

B. Cs-induced surface dipoles

Adatom-induced surface dipoles will vary the ionization energy of semiconductor surfaces. Cesium is much more electropositive than both gallium and nitrogen. Cesium-induced surface dipoles on GaN are thus oriented such as to lower the surface barrier for electrons in agreement with what is found experimentally. The adatom-induced reduction of the ionization energy as a function of coverage Θ is given by[20]

$$\delta I = \frac{e_0}{\epsilon_0} \frac{\mu_0 \Theta}{1+9\alpha_{ad}\Theta^{3/2}}, \tag{3}$$

where e_0 and ϵ_0 are the electronic charge and the permittivity of vacuum, μ_0 is the dipole moment of the adatom-induced dipoles, and α_{ad} is the polarizibility of the adatom. The dipole moment of the adatom-induced dipoles may be written as

$$\mu_0 = \Delta q e_0 (r_s + r_{ad}), \tag{4}$$

where r_s and r_{ad} are the covalent radii of the substrate atoms and the adatoms, respectively. Pauling[20] described the partial ionic character Δq_1 of covalent single bonds by the electronegativity difference of the respective atoms. A revised version[21] of Pauling's original correlation gives

$$\Delta q_1 = 0.16|X_s - X_{ad}| + 0.035(X_s - X_{ad})^2, \tag{5}$$

where X_s and X_{ad} are the electronegativities of the substrate atom and the adatom, respectively.

Experimental and theoretical studies concluded that GaN(0001) surfaces should be Ga-terminated.[22] Therefore, we assume that covalent Cs–Ga bonds are present after the deposition of Cs on Ga(0001) surfaces. The maximum reduction of the ionization energy is obtained when the first continuous layer of covalently bonded Cs atoms is completed. Using the polarizability $\alpha_{Cs} = 1.63 \times 10^{-23}$ cm^3 of adsorbed Cs,[23] the covalent radii $r_{Ga} = 0.126$ nm and $r_{Cs} = 0.235$ nm and the electronegativity difference $X_{Ga} - X_{Cs} = 1.02$ one obtains $\delta I_{max} = 2.5$ eV. This value is very close to the experimental result $\delta I_{max}^{exp} = 2.3 \pm 0.2$ eV.

C. Negative electron affinity on Cs-covered GaN

The position of the conduction-band minimum W_{cb} in the bulk with respect to the vacuum level is given by

$$(W_{vac} - W_{cb})_n = \Phi_n - W_n \tag{6a}$$

on *n*-type substrates and

$$(W_{vac} - W_{cb})_p = \Phi_p + W_p - W_g \tag{6b}$$

on *p*-type substrates, where $W_n = W_{cb} - W_F = 0.03$ eV and $W_p = W_F - W_{vb} = 0.05$ eV are the energy separations between the respective edges of the majority-carrier bands and the Fermi level. The data plotted in Fig. 3 always give positive $(W_{vac} - W_{cb})_n$ values. On *p*-GaN(0001) surfaces, on the other hand, very small cesium coverages already reduce the clean-surface value $(W_{vac} - W_{cb})_p \approx 0.3$ eV to negative values. Thus, minute Cs coverages suffice to achieve negative electron affinity on *p*-GaN(0001) surfaces. After completion of the first continuous Cs layer, the vacuum level is by 2.1 eV below the conduction-band minimum in the bulk.

D. Schottky barrier height of Cs/GaN(0001) contacts

In the coverage regime above the first continuous Cs layer on GaN(0001) surfaces, the UPS spectra reveal a pronounced Fermi edge, i.e., the Cs overlayer becomes metallic. This finding is further supported by the observations that in this coverage range the work function assumes the value of compact cesium metal and the band bending reaches a constant value. Metal–semiconductor contacts are characterized by their Schottky barrier height $\phi_{Bn} = W_{ci} - W_F$ and $\phi_{Bp} = W_F$

$-W_{vi}$ which are the energy differences between the Fermi level and the edge of the respective majority-carrier band right at the interface. The data plotted in Fig. 5 give $\phi_{Bn} = 0.2 \pm 0.15$ eV for Cs/*n*-GaN(0001) contacts at 150 K.

The continuum of metal-induced gap states determines the electronic properties of ideal metal–semiconductor interfaces.[3] The charge transfer between the semiconductor and the metal may be modeled by applying Pauling's concept mentioned which was originally designed for small molecules. The variation of the Schottky barrier height as a function of the difference $X_m - X_s$ of the metal and the semiconductor electronegativities is then given by[24]

$$\phi_{Bn} = \phi_{bp} + S_x(X_m - X_s), \tag{7}$$

where $\phi_{bp} = W_c - W_{bp}$ is the zero-charge-transfer barrier height.

The energy positions of the branch point $W_{bp} - W_v$ in the continuum of interface-induced gap states were calculated by using an empirical tight-binding method.[25] The branch point energy of GaN resulted as $W_{bp} - W_v = 2.35$ eV, i.e., the zero-charge-transfer barrier height ϕ_{bp} of GaN measures 1.1 eV since the GaN band gap measures 3.45 eV. The slope parameters $S_x = \partial\phi_B / \partial X_m$ are related to the optical dielectric constants ϵ_∞ of the respective semiconductors by[26]

$$A_x / S_x - 1 = 0.1(\epsilon_\infty - 1)^2, \tag{8}$$

with $A_x = 0.86$ eV/Miedema unit. With $\epsilon_\infty = 5.45$ (Ref. 27) for GaN the slope parameter results as $S_x = 0.29$ eV/Miedema unit. With the Miedema electronegativities of 5.3 and 1.95 for GaN and Cs, respectively, the MIGS-and-electronegativity model predicts a barrier height of 0.13 eV for Cs:*n*-GaN(000) contacts. This value is in excellent agreement with the experimentally determined Schottky barrier height of 0.2 ± 0.15 eV. This finding corroborates our earlier conclusion[28] that the MIGS-and-electronegativity model provides an excellent description of the chemical trend of barrier heights of ideal GaN Schottky contacts.

V. CONCLUSION

We investigated the electronic properties of clean and Cs-covered *n*- and *p*-type GaN(0001) surfaces at low temperatures. Clean *n*- and *p*-type GaN(0001) surfaces exhibit depletion and inversion layers, respectively, and the Fermi level is pinned by surface states close to the conduction-band minimum. Cs-induced surface dipoles reduce the ionization energy from its clean-surface value of 6.8–4.5 eV after the completion of the first continuous Cs layer. At clean *p*-GaN(0001) surfaces the vacuum level lies only 0.3 eV above the conduction-band bottom in the bulk. Already minute amounts of Cs suffice to produce negative electron affinity. After the completion of a continuous Cs layer the vacuum level is by 2.1 eV below the condition-band minimum in the bulk. The Schottky barrier height of Cs/*n*-GaN(0001) contacts is determined as 0.2 eV. This value is in excellent agreement with the predictions of the MIGS-and-electronegativity model.

ACKNOWLEDGMENTS

This study was partly supported by Grants Nos. Mo 318/28-1 and Stu 139/3-1 of the Deutsche Forschungsgemeinschaft and of the Bayrische Forschungsgemeinschaft (FOROPTO II).

[1]C. R. Bolognesi, Compd. Semicond. **March/April**, 10 (1997).
[2]R. L. Bell, *Negative Electron Affinity Devices* (Clarendon, Oxford, 1973).
[3]W. Mönch, *Semiconductor Surfaces and Interfaces*, 2nd ed. (Springer, Berlin, 1995).
[4]F. J. Himpsel, J. A. Knapp, J. van Vechten, and D. E. Eastman, Phys. Rev. B **20**, 624 (1979).
[5]B. B. Pate, M. H. Hecht, C. Binns, I. Lindau, and W. E. Spicer, J. Vac. Sci. Technol. **64**, 268 (1982).
[6]J. van der Weide and R. J. Nemanich, Appl. Phys. Lett. **62**, 1878 (1993).
[7]D. P. Malta, J. B. Posthill, T. P. Humphreys, R. E. Thomas, G. G. Fountain, R. A. Rudder, G. C. Hudson, M. J. Mantini, and R. J. Markunas, Appl. Phys. Lett. **64**, 1929 (1994).
[8]J. van der Weide, Z. Zhang, P. K. Baumanns, M. G. Wensell, J. Bernholc, and R. J. Nemanich, Phys. Rev. B **50**, 5803 (1994).
[9]O. M. Küttel, O. Gröning, E. Schaller, L. Diederich, P. Gröning, and L. Schlapach, Diamond Relat. Mater. **5**, 807 (1996).
[10]M. C. Benjamin, C. Wang, R. F. Davis, and R. J. Nemanich, Appl. Phys. Lett. **64**, 3288 (1994).
[11]M. C. Benjamin, M. D. Bremser, J. T. W. Weeks, S. W. King, R. F. Davis, and R. J. Nemanich, Appl. Surf. Sci. **104/105**, 455 (1996).
[12]M. Eyckeler, T. U. Kampen, and W. Mönch, Appl. Surf. Sci. **123/124**, 28 (1998).
[13]H. Angerer, O. Ambacher, R. Dimitrov, T. Metzger, W. Rieger, and M. Stutzmann, MRS Internet J. Nitride Semicond. Res. **1**, 15 (1996).
[14]R. Linz, H. J. Clemens, and W. Mönch, J. Vac. Sci. Technol. B **11**, 1591 (1993).
[15]V. van Elsbergen, T. U. Kampen, and W. Mönch, J. Appl. Phys. **79**, 316 (1996).
[16]G. Ertl and J. Küppers, *Low Energy Electrons and Surface Chemistry* (VCH, Weinheim, Germany, 1985).
[17]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, CA, 1980).
[18]Y. Li, Y. Lu, H. Shen, M. Wraback, M. G. Brown, M. Schurman, L. Koszi, and R. A. Stall, Appl. Phys. Lett. **70**, 2458 (1997).
[19]V. M. Bermudez, J. Appl. Phys. **80**, 1190 (1996).
[20]J. Topping, Proc. R. Soc. London **68**, 171 (1946).
[21]N. B. Hanney and C. P. Smith, J. Am. Chem. Soc. **68**, 171 (1946).
[22]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
[23]W. Mönch, Phys. Status Solidi **40**, 257 (1970).
[24]W. Mönch, Phys. Rev. Lett. **58**, 1260 (1987).
[25]W. Mönch, J. Vac. Sci. Technol. B **14**, 2985 (1996).
[26]W. Mönch, Appl. Surf. Sci. **92**, 367 (1996).
[27]Landolt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung (Springer, Berlin, 1982), Vol. 17a.
[28]T. U. Kampen and W. Mönch, Appl. Surf. Sci. **117/118**, 388 (1997).

Nucleation and growth of GaN layers on GaAs, Si, and SiC substrates

Klaus H. Ploog,[a] Oliver Brandt, Hui Yang, Bin Yang, and Achim Trampert
Paul Drude Institute for Solid State Electronics, D-10117 Berlin, Germany

(Received 21 January 1998; accepted 11 May 1998)

We summarize our results on plasma-assisted molecular beam epitaxy of cubic GaN on GaAs(001) and on Si(001) and of hexagonal GaN on 6H-SiC(0001) with emphasis on the nucleation process. A two-step growth sequence must be used to optimize and control the nucleation and the subsequent growth independently. While a perfect epitaxial orientation exists for GaN-on-GaAs due to the coincidence lattice relationship of the two constituents. The same effect is impeded for GaN-on-Si by the growth of Si_xN_y inclusions at the interface which act as nucleation cores for the formation of the hexagonal GaN phase. A suitable template, such as a thin GaAs or SiC insertion layer, avoids formation of the Si_xN_y inclusions. Finally, growth of hexagonal GaN-on-6H-SiC without any buffer layer requires very careful adjustment of the N-to-Ga flux ratio and the substrate temperature, independently for the nucleation stage and for the subsequent layer-by-layer growth. The structural perfection and the optical properties of the resulting 1 μm thick GaN films then reach state-of-the-art quality even without a buffer template. © *1998 American Vacuum Society.*
[S0734-211X(98)04904-X]

I. INTRODUCTION

The lack of GaN substrates of suitable size and structural perfection has prompted numerous investigations of heteroepitaxial growth of GaN under conditions of a large mismatch of lattice constants, thermal properties, etc. The existence of a metastable cubic GaN phase in addition to the stable hexagonal phase imposes added problems to the fabrication of GaN films suitable for device application. In this article we will summarize our results on plasma-assisted molecular beam epitaxy (MBE) growth of cubic GaN on GaAs(001) (Refs. 1–4) and on Si(001) (Refs. 5 and 6) and of hexagonal GaN on 6H-SiC(0001),[7,8] with special emphasis on the nucleation process. While an epitaxial "cube-on-cube" orientation exists for GaN-on-GaAs due to the coincidence lattice of the two components, the same effect is hampered for GaN-on-Si by the formation of amorphous Si_xN_y inclusions at the interface which act as nucleation centers for the formation of the hexagonal GaN phase. A suitable template, such as a thin GaAs or SiC insertion layer, prevents formation of the Si_xN_y inclusions. Finally, we discuss the conditions for the direct nucleation of hexagonal GaN on 6H-SiC without the frequently employed buffer layer. This is an important step towards the realization of vertically conductive hexagonal group III nitride devices. As the nucleation process has a direct impact not only on the structural but also on the electronic properties of the resulting GaN films, we also highlight some key optical features of the GaN-on-SiC heterostructures.

II. EXPERIMENT

The GaN films were grown in a homemade MBE system equipped with a dc glow discharge nitrogen plasma source (NPS) (Riber NPS 200) operated at a power of 30 W. The substrate preparation, nucleation process, and film growth were monitored *in situ* by reflection high-energy electron diffraction (RHEED), using an incident angle of 1°–2° at an acceleration voltage of 15 keV. The RHEED patterns were recorded by a charge-coupled device (CCD) camera connected to an image processing system. In addition to *in situ* RHEED, the morphology and the structure were investigated *ex situ* by atomic force microscopy (AFM), high-resolution x-ray diffraction (XRD), and high-resolution transmission electron microscopy (HRTEM). A two-step growth sequence was employed to optimize and control the nucleation stage and the subsequent layer growth independently. For the latter, the growth rate was typically 0.1 monolayer (ML)/s.

III. RESULTS AND DISCUSSION

A. Cubic GaN-on-GaAs(001)

Cubic GaN(001) exhibits two Ga-terminated surface reconstructions with the RHEED patterns shown in Fig. 1, one with (2×2) symmetry and the other with $c(2\times2)$, while the N-terminated (1×1) surface remains unreconstructed.[1–3] The (2×2) reconstructed surface is thermally stable in ultrahigh vacuum (UHV) up to 700 °C. It transforms to the $c(2\times2)$ reconstruction after Ga deposition at $T_s \leqslant 600$ °C. As the transitions between the different reconstructions can be monitored *in situ* by RHEED, the RHEED transients have been used for a detailed study of the GaN(001) surface phase diagram and hence the GaN growth kinetics.[1,2] Moreover, these results allowed us to determine the effective N-to-Ga flux ratio in real time by means of the RHEED transients.[9,10]

The observed surface phase transitions as a function of the impinging Ga or N flux led to the conclusion that the (2×2) and the $c(2\times2)$ reconstructions correspond to surfaces covered with 0.5 and 1.0 ML Ga, respectively.[1,2] Simple surface models[1] assumed that the (2×2) structure is formed by Ga_2 dimer rows along [110] which are separated by one missing dimer row. Filling the missing dimer rows at

[a]Electronic mail: ploog@pdi-berlin.de

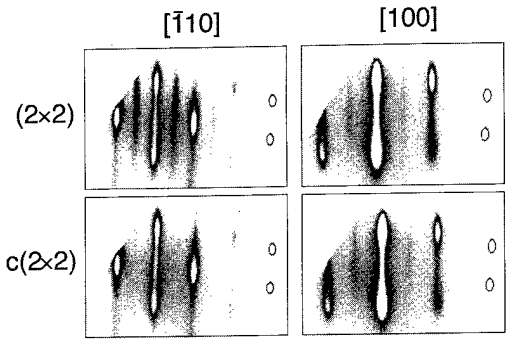


FIG. 1. RHEED patterns of cubic GaN(001) taken along [$\bar{1}$10] and [100] azimuths at two different N-to-Ga flux ratios. The patterns are grayscale encoded for better visibility. The two spots at the right of each pattern stem from light emitting diodes indicating the "open" state for both the Ga and the N shutters.

the center positions of four adjacent Ga_2 dimers at full monolayer coverage should then correspond to the $c(2\times2)$ structure. While our more recent scanning tunneling microscopy (STM) studies[11] confirm the surface structure model of the (2×2) reconstructions with the Ga_2 dimers along [110] at half-monolayer coverage, the $c(2\times2)$ surface previously noticed exhibits the more complicated $(\sqrt{10}\times\sqrt{10})$ $R18.4°$ reconstruction with a Ga coverage of 0.8 instead of 1.0 ML. Using a ball-and-stick model the $(\sqrt{10}\times\sqrt{10})$ $R18.4°$ reconstruction can be constructed as a vacancy structure of the $c(2\times2)$ reconstruction by removing one of the five Ga_2 dimers. Neither reconstruction can be distinguished by RHEED.

Prior to the deposition of GaN, a GaAs buffer layer is grown under conditions suitable to obtain an atomically flat surface.[1,12] The nucleation of GaN on this β-(2×4) reconstructed GaAs(001) surface is then initiated by closing the As_4 shutter and immediately opening the N shutter. The growth conditions of GaN must be adjusted such that a complete coverage of the GaAs surface by GaN is most rapid. This demand is met by minimizing the size and maximizing the density of the nuclei. We have found[10,12] that at substrate temperatures of 580–620 °C, a growth rate of 0.03 ML/s and an effective N-to-Ga flux ratio of 3 yield the best nucleation results. After deposition of 5–7 ML GaN a continuous film has been formed, and the growth rate and the growth temperature can be increased (0.1 ML/s at 680 °C). To prevent the formation of hexagonal domains in the cubic GaN matrix, growth now has to be carried out with a N-to-Ga flux ratio close to unity, yielding the transition from the (2×2) to the $c(2\times2)$ reconstruction. Monitoring these distinct GaN(001) surface structures in real time by RHEED provides a unique means for the realtime control of the surface stoichiometry, and hence enables us to synthesize single-phase cubic GaN films. The phase purity degrades with deviations from the optimum flux ratio, and the phase mixture is an important factor that governs the crystal quality of cubic GaN films.

To overcome the large lattice mismatch of 20% between cubic GaN and GaAs and to guarantee phase purity, accurate

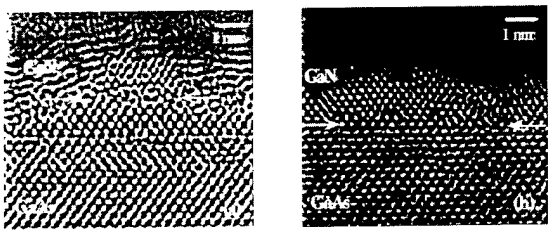


FIG. 2. Nucleation state of cubic GaN-on-GaAs(001) observed by cross-sectional HRTEM. GaN nuclei after deposition of nominally (a) 0.5 and (b) 5 ML thicknesses (the images are Fourier filtered).

control of the nucleation state (i.e., the first 5 ML) with the small cubic GaN nuclei exhibiting the desired epitaxial relation is most crucial. In the following we show that the resulting microstructure of the cubic GaN film is a direct consequence of the atomic configuration of the interface,[13] which can be understood in terms of a coincidence lattice model. The cross-sectional HRTEM image of a nominally 0.5 ML GaN sample [Fig. 2(a)] demonstrates that the nucleation occurs by the formation of nanoscale three-dimensional (3D) islands, which exhibit a well-defined epitaxial relationship. Even the smallest islands observed, which have a volume of less than 10 nm^3, are relaxed by misfit dislocations. The island shown in Fig. 2(a), for example, contains an edge-type dislocation in the center and a second one which is about to form at the edge of the island. The atomic structure close to the interface within the GaN island is highly distorted because the extension of the strong lattice plane bending near the dislocation core is of the same order of magnitude as the radius of the observed island. Additionally, elastic relaxation processes at the island surface may also influence the lattice image of the interface. The cross-sectional HRTEM image of the 5 ML GaN sample in Fig. 2(b) evidences that this sample exhibits a connected filmlike morphology with a thickness comparable to that of the 0.5 ML sample. The initial nuclei thus seem to grow almost exclusively laterally, with little growth at this stage along the [001] direction. Moreover, misfit dislocation cores are detected with a distance of five {111} GaN lattice planes. The density of these edge-type dislocations is just sufficient to relieve the entire misfit strain of 20%.

These results directly visualize that pure edge-type misfit dislocations are formed instantaneously during the initial nucleation of islands via the incorporation of extra {110} lattice planes into the edge of the growing island without any climb or glide mechanism.[13] On the other hand, it is clear that the epitaxial cubic interface must be energetically more favorable than an interface involving the thermodynamically stable hexagonal lattice. Assuming that the strain energy represents the most important part of the total interfacial energy (the chemical portion should be comparable between cubic and hexagonal phases), the results can be explained by a coincidence lattice or a "magic mismatch" occurring between cubic GaN and GaAs. A true coincidence between the epilayer (a_{ept}) and substrate lattice (a_{sub}) occurs when $a_{epi}/a_{sub}=m/n$, where m and n are integers. If $m=n\pm1$,

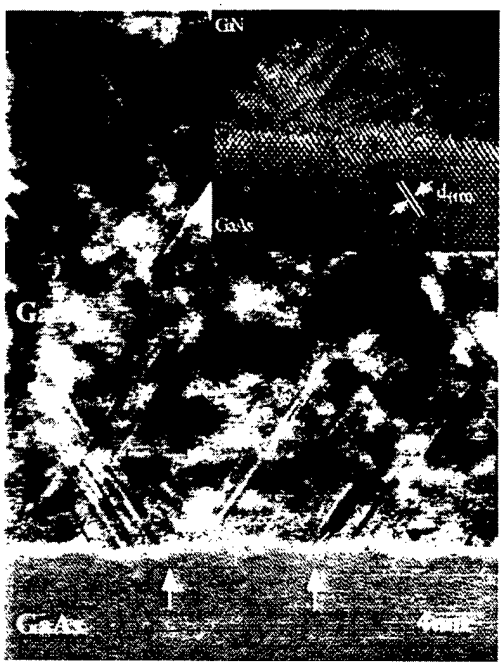


FIG. 3. Cross-section HRTEM image of the GaN–GaAs(001) heterostructure taken along the ⟨110⟩ direction. The arrows depict an atomically flat part of the interface containing regularly arranged misfit dislocations (magnified in the inset).

there is one extra lattice plane, i.e., a simple edge dislocation, in each unit cell of the coincidence lattice. However, the epitaxial system is in general not at true coincidence, and the coincidence–lattice mismatch f expresses this deviation from true coincidence as $f=(ma_{\text{sub}}-na_{\text{epi}})/ma_{\text{sub}}$. This deviation introduces elastic strain at the interface in addition to the strain accommodated by the misfit dislocations. Therefore, the energy of heteroboundaries is small only if f does not deviate substantially from true coincidence. For the present heterosystem $f=-0.0002\pm0.0020$ by taking $m/n=4/5$ and the most accurate values for the lattice constants available at growth temperature, i.e., $a_{\text{GaN}}=0.455\pm0.01$ nm and $a_{\text{GaAs}}=0.568886$ nm. This system is thus close to true coincidence (magic mismatch), and an array of edge dislocations with a period of five GaN lattice planes indeed accounts for the entire misfit. The occurence of a magic mismatch between cubic GaN and GaAs thus provides an explanation of the phenomenon of epitaxial growth for a strain at which epitaxy of covalently bonded materials is usually no longer achieved.

The subsequent growth and the evolution of the defect microstructure detected in the TEM micrographs (Fig. 3) are determined by the initial nucleation of the islands. While no planar defects are observed in the isolated nuclei [cf. Fig. 2(a)], stacking faults and microtwins are generated during the coalescence stage of the nuclei as observed in the HRTEM images of the 5 ML sample [Fig. 2(b)]. In fact, since the spacing of the individual nuclei is not necessarily in phase with respect to their dislocation array, the periodicity of the dislocation array is in general broken upon their coalescence. These locations are centers of high local strain (e.g., a 4/3 ratio between GaN and GaAs lattice planes corresponds to a residual misfit of $f=6.7\%$). These local strain concentrations are probably responsible for the generation of secondary defects in the layers, i.e., stacking faults and microtwins, which are able to fit the coincidence lattice in the region of the coalesced islands. The interface is sharp on an atomic scale with an overall vertical roughness of only a few monolayers. A more careful inspection of the interface structure (see the inset in Fig. 3) reveals that the boundary contains atomically straight parts. Within these regions no planar defects originate from the interface and, simultaneously, regularly arranged pure edge-type dislocations appear at the interface. The cores of these dislocations are observed at every fifth GaN {111} lattice plane as expected from the nucleation state described above (Fig. 2).

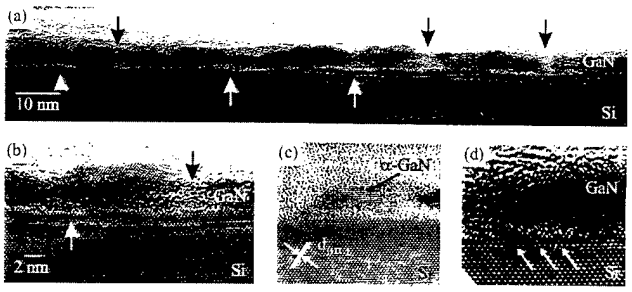


FIG. 4. ⟨110⟩ cross-sectional HRTEM micrographs of a 10-ML-thick GaN-on-Si(001) nucleation layer. (a) Overview, (b) higher magnification, showing the presence of amorphous material at the interface, (c) higher magnification, showing a hexagonal GaN grain nucleating on top of the amorphous material, and (d) higher magnification, showing an area where a well-defined epitaxial relation between cubic GaN and Si is achieved. In (a) and (b), the arrows point to the locations of the amorphous material; in (d) the arrows indicate the edge-type dislocations formed every 5 Si–Si interatomic distances.

Cubic GaN-on-GaAs(001) films grown under optimized nucleation conditions exhibit a bright luminescence at room temperature[4] which makes them promising for application in optoelectronic devices.

B. Cubic GaN-on-Si(001)

The GaN films are grown on Si(001) also in two steps, i.e., a lower temperature nucleation at 520–550 °C is followed by a higher temperature growth at 650–680 °C for the rest of the film.[5,6] However, achieving homogeneous wetting of cubic GaN on Si(001) is much more difficult than on GaAs(001). Therefore, only when nucleation is performed on a single-domain (1×2)-Si(001) surface (which requires substrates oriented 4° off toward ⟨110⟩), the RHEED patterns indicate a well-defined cubic symmetry after growth of 10 ML GaN. Atomic force micrographs reveal a smooth and connected surface morphology of these films. However, a more detailed analysis of the layers grown even under optimized conditions by HRTEM shows that epitaxy is only partially achieved, as evidenced by the HRTEM micrographs of one of these layers in Figs. 4 and 5. Up- and downward arrows in Fig. 4(a) indicate regions directly at the interface and within the layer, respectively, which exhibit contrast different from that of cubic GaN. Some portions of this micro-

graph shown at higher magnification in Figs. 4(b)–4(d) reveal the following interesting phenomena. The nucleation layer is actually not a homogeneous and continuous GaN layer, but is interrupted by randomly distributed amorphous domains [as indicated by the downward arrow in Fig. 4(b)] in between the cubic GaN grains. Our detailed studies of the interaction between the active N plasma and the Si(001) surface demonstrate that this amorphous material at the GaN–Si interface is Si_xN_y. At some GaN–Si interfacial regions [indicated by the upward arrow in Fig. 4(b)], the cubic GaN grains are heavily misoriented and contain a very high density of stacking faults. Hexagonal GaN grains whose basal plane is parallel to the Si(001) surface are observed to grow on top of the amorphous patches at the interface [Fig. 4(c)]. For those GaN–Si interfacial regions where a good epitaxial relation between GaN and Si is established [Fig. 4(d)], the large lattice mismatch (≈17%) between cubic GaN and Si is accommodated by misfit dislocations, which indeed form a 5:6 coincidence lattice as expected from the ratio of the lattice constants.

Continuous growth on top of the nucleation layer is found to lead to a transition from the cubic towards the hexagonal phase, as evidenced by RHEED. This transition is independent of both the thickness of the GaN nucleation layer and the conditions used for subsequent growth. The phase transformation is confirmed by x-ray diffractometry of a 700-nm-thick GaN film, where the dominant peak stems from the GaN(0002) reflection. Azimuthal scans across asymmetric reflections reveals the in-plane orientation of the characteristic columnar structure of hexagonal GaN. The columns are separated by straight boundaries and have well-defined surface facets. Within these hexagonal GaN grains, a high density of planar defects exists, which are stacking faults lying in the close-packed basal plane. These stacking faults are running parallel to the GaN–Si interface, showing that the out-of-plane orientation spread of the columns is small, resulting in a highly textured, fiberlike microstructure. The change of contrast close to the GaN–Si interface originates from grains of cubic GaN, as demonstrated by means of HRTEM (see the inset of Fig. 5) and selected-area electron diffraction (SAD) patterns. They are separated from each other by the hexagonal GaN columns which extend from the amorphous Si_xN_y patches at the interface.

The nucleation and growth of GaN on Si(001) hence take place via two different parallel paths. Cubic GaN with an epitaxial relationship to Si(001) is formed on the bare Si surface, but complete coverage is inhibited by the simultaneous formation of amorphous Si_xN_y. GaN condensing on top of this Si_xN_y-covered Si does not experience an epitaxial constraint and grows thus in its hexagonal modification with random in-plane orientation. Further growth leads to largely columnar growth of hexagonal GaN which rapidly overgrows the initial cubic GaN grains. This latter fact implies a higher growth rate of the basal plane and/or a lower surface energy compared to the (001) plane of cubic GaN. In any case, the phase transformation seems inevitable unless the Si surface is protected against impinging N.

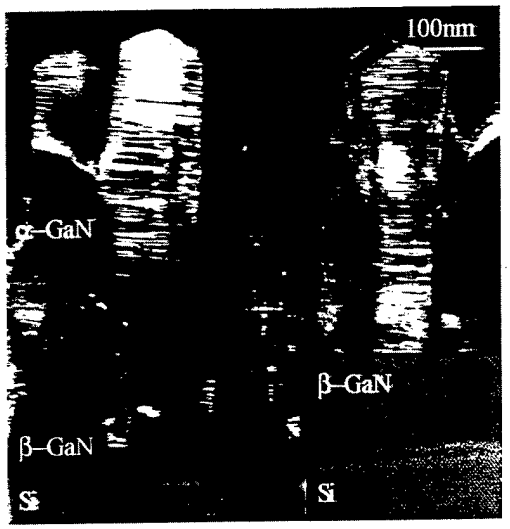


FIG. 5. ⟨110⟩ cross-section TEM dark-field image of a 0.7-μm-thick GaN film grown directly on Si(001). The inset displays a high-resolution image of a cubic GaN grain at the interface.

C. Cubic GaN on GaAs-buffered Si(001)

In order to inhibit the formation of Si_xN_y at the nucleation stage, a GaAs buffer is employed in our experiments.[5] Optimized growth conditions are used to grow both the GaAs buffer layer and the cubic GaN epilayers. The x-ray diffraction profile of a 0.7-μm-thick GaN film grown on GaAs-buffered Si(001) is dominated by the (002) reflection of cubic GaN. The phase content and crystallinity of the GaN film are also characterized by taking SAD patterns from a ⟨110⟩ cross-sectional specimen of the GaN–GaAs heterostructure. The superposition of the diffraction patterns of the GaAs buffer layer and the GaN film along their ⟨110⟩ zone axes visualizes the epitaxial orientation relationship. The lattice mismatch of 20% measured from the distance of the corresponding GaN and GaAs diffraction spots is in agreement with the result obtained by x-ray diffraction. The streaks observed along the ⟨111⟩ directions reveal that the major structural defects in this film are stacking faults.

The cross-sectional HRTEM micrographs taken from both the GaAs–Si and the GaN–GaAs interfaces [Figs. 6(a) and 6(b), respectively] demonstrate the well-defined epitaxial relationships. The large lattice mismatches at both interfaces are primarily accommodated by misfit dislocations. Further-

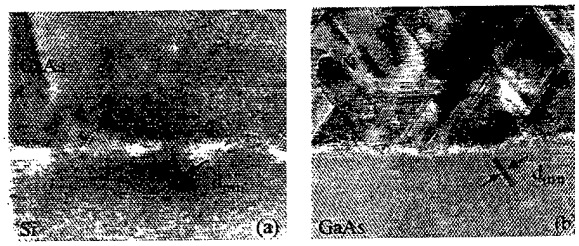


FIG. 6. ⟨110⟩ cross-sectional HRTEM images of 0.7-μm-thick GaN film grown on GaAs–Si(001) showing (a) the GaAs–Si interface and (b) the GaN–GaAs interface.

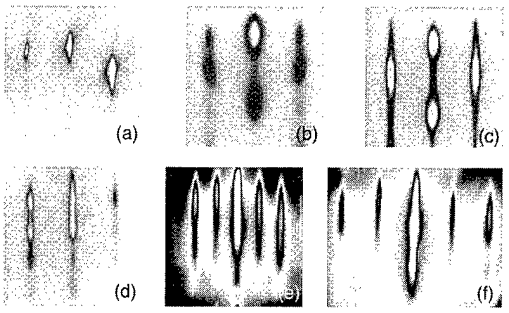


FIG. 7. *In situ* RHEED patterns taken from the (a) [11$\bar{2}$0] azimuth, 6*H*-SiC substrate; (b) [11$\bar{2}$0] azimuth, 3 ML GaN nucleated under nonoptimized conditions; (c) [11$\bar{2}$0] azimuth, subsequent growth of GaN layer under nonoptimized conditions; (d) [11$\bar{2}$0] azimuth, 3 ML quasi-2D nucleation layer; (e) [11$\bar{2}$0] azimuth, subsequent 2D growth, 2×; and (f) [1$\bar{1}$00] azimuth, subsequent 2D growth, 2×.

more, stacking faults and microtwins penetrate into the epilayers. These and the above results are essentially identical to those obtained for cubic GaN layers grown directly on GaAs(001). However, TEM micrographs covering a larger interfacial length (not shown here) reveal the additional presence of bundles of threading dislocations originating at the GaN–GaAs interface which are not observed for cubic GaN layers on on-axis GaAs(001) substrates. We speculate that the origin of these threading dislocations is related to the vertical mismatch between GaN and GaAs at the steps on the off-axis GaAs surface. The presence of these additional defects might also explain that, while band-edge photoluminescence (PL) of cubic GaN is observed for these layers up to room temperature, its integrated intensity is substantially lower than that obtained from GaN films grown directly on GaAs(001).[6] Therefore, growth of cubic GaN films on 3*C*-SiC-buffered Si(001) is underway in our laboratory.

D. Hexagonal GaN-on-SiC(0001)

In order to fully exploit the advantages of doped SiC substrates for vertically conductive GaN-based light emitting devices, hexagonal GaN layers of high structural perfection must be grown directly on 6*H*-SiC(0001) substrates. However, the major obstacle in nucleating single-crystal GaN directly on SiC is the poor wetting of GaN due to the lower surface energy of 6*H*-SiC(0001).[14] As a consequence, 3D randomly oriented independent GaN islands tend to form upon the onset of nucleation, eventually leading to textured or polycrystalline growth. In addition, there are several possible bonding configurations at the GaN–SiC interface.[15] The growth conditions must hence promote the exclusive formation of N–Si bonds at the interface. The conventional way to overcome these difficulties has been the insertion of an AlN buffer layer. However, even upon doping the AlN layer represents a potential barrier for electrons injected from the SiC substrate.

Smooth and clean 6*H*-SiC(0001) starting surfaces, as evidenced by the clear RHEED pattern and Kikuchi lines in Fig. 7(a), are obtained by heating carefully selected specimen from commercial wafers to 850 °C for 30 min prior to growth. In a wide range of nucleation conditions, 3D polycrystalline GaN islands are formed, as revealed by the spotty and tilted RHEED patterns from the (1$\bar{1}$01) and ($\bar{1}$101) planes of GaN islands in Fig. 7(b). The observed tilt indicates a certain degree of orientational distribution of the GaN island planes, such as (1$\bar{1}$01) and ($\bar{1}$101). Further deposition of GaN on such kind of nucleation layer leads to GaN films of poor crystallinity, as evidenced by the spotty RHEED pattern in Fig. 7(c) and by XRD measurements. Note that no reconstruction features exist in the pattern of Fig. 7(c).

The wetting between 6*H*-SiC and GaN can be strongly improved by, first, nitriding the SiC surface at sufficiently high temperatures of 850 °C and, second, subsequent nucleation of GaN at a low substrate temperature of 620–650 °C and an effective N-to-Ga flux ratio of 1.5–2.0. Under these conditions, Si–N bonds are preferentially formed at the GaN/SiC interface, the interfacial charge is altered, and the local microscopic lattice mismatch is minimized. This recipe thus leads to the onset of quasi-two-dimensional (2D) nucleation of single crystal GaN, as indicated by the elongated RHEED features from nominally 3 ML GaN in Fig. 7(d). Note the absence of any tilt of the GaN nuclei planes. Upon further deposition, the quasi-2D GaN islands quickly coalesce and the growth mode changes to the layer-by-layer mode. This growth mode transition is evidenced by the corresponding transition in the RHEED patterns. In addition, after deposition of 20 ML GaN, clear (2×2) reconstruction features appear along the [11$\bar{2}$0] and [1$\bar{1}$00] azimuths [see Figs. 7(e) and 7(f)]. On such optimized GaN starting layers high-quality GaN films subsequently can be grown, if the (2×2) surface reconstruction is maintained. Stabilization of the (2×2) reconstruction is achieved in a rather narrow growth regime, i.e., a substrate temperature of 750 °C and an effective N-to-Ga flux ratio close to unity.

The specific nucleation and growth conditions in plasma-assisted MBE of GaN-on-6*H*-SiC(0001) deserve some additional comments. In general, higher substrate temperatures are expected to improve the structural perfection of the individual GaN nuclei. However, a low substrate temperature is required to promote the wetting process. Using *in situ* RHEED measurements we found that at substrate temperatures above 850 °C no growth of GaN occurs, probably due to the strongly increased desorption rate of the Ga and N species at elevated temperature from the foreign substrate as compared to a continuous GaN film. As a compromise, we have used substrate temperatures of 620–650 °C for nucleation. Being restricted in selecting the GaN nucleation temperature, the nucleation rate is hence limited to a rather narrow range. In our experiments a low growth rate of nominally less than 0.01 ML/s turned out to be important for the nucleation of single crystal GaN directly on 6*H*-SiC. Higher nucleation rates resulted inevitably in poor crystalline GaN nuclei. The reason for this behavior is assumed to be associated with the strongly reduced migration lengths of the reactive species, in particular Ga, at the rather low substrate temperature.[10,12] As longer times are needed for the species to arrive at their ideal lattice sites, a higher nucleation rate

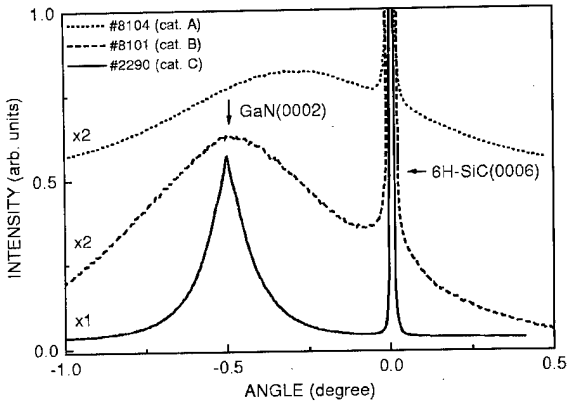


FIG. 8. Double-crystal x-ray rocking curves from sample A (#8104, nucleation and growth under nonoptimized conditions, 0.9 μm thick), sample B (#8101, nucleation optimized and growth nonoptimized, 0.9 μm thick), and sample C (#2290, both nucleation and growth optimized, 0.5 μm thick).

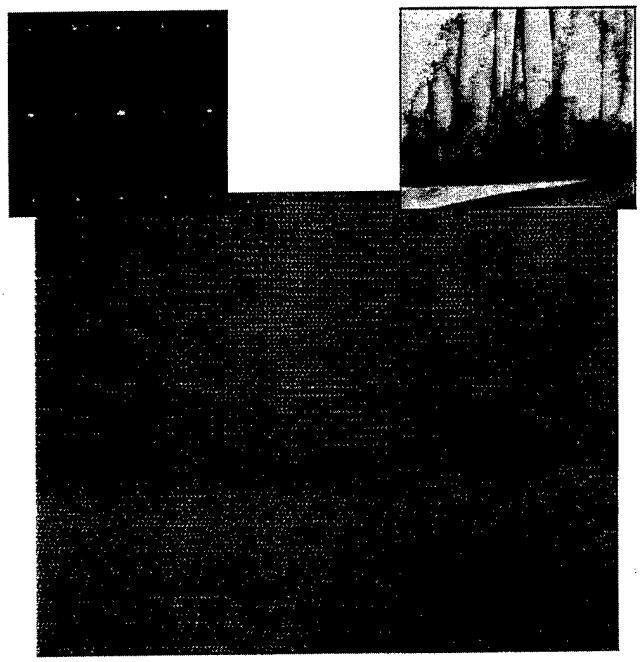


FIG. 9. Cross-sectional TEM images taken from sample C. The HRTEM image is viewed along the $\langle 11\bar{2}0\rangle$ projection of the GaN/SiC interface. The left inset shows the SAD pattern taken from the GaN/SiC interfacial region. The right inset shows the overview TEM image taken along the $\langle 11\bar{2}0\rangle$ $6H$-SiC lattice direction.

would induce disorder, defects, and possibly excess strain in the GaN nuclei. Finally, the effective N-to-Ga flux ratio also plays an important role in the nucleation stage. If this flux ratio is below unity, separated polycrystalline GaN nuclei are formed. However, a flux ratio above 3 results in 3D growth after 2D nucleation. The optimum flux ratio of 1.5–2.0 in our experiments leads to quasi-2D nucleation of GaN-on-$6H$-SiC(0001). To understand this behavior we have to consider not only the wetting of the GaN nuclei to the SiC surface but also the strain relaxation process. As shown in Ref. 12, a higher N-to-Ga flux ratio reduces the Ga diffusion length, thus yielding a high density of small GaN nuclei which is favorable to force GaN nuclei to wet the SiC surface. On the other hand, relaxing as much as possible the 3.4% lattice mismatch at the GaN/$6H$-SiC interface favors the stable 2D growth of GaN in the subsequent stage. A high N-to-Ga flux ratio, yielding a too high density of small GaN nuclei, would lead to the onset of 2D growth of a thin pseudomorphic GaN layer on the $6H$-SiC substrate. The onset of 2D growth of the pseudomorphic GaN layer under these conditions has been observed *in situ* by RHEED. The important results were the high density of threading dislocations formed when the GaN layer thickness reached the critical value (about 5 ML for the 3.4% lattice mismatch) and the subsequent 3D growth mode when growth was continued. The value of 1.5–2.0 for the N-to-Ga flux ratio during GaN nucleation hence compromises the two prerequisites of enforcing wetting and relaxing strain at the GaN/SiC interface.

The results in optimizing both the nucleation and the subsequent growth of GaN-on-$6H$-SiC(0001) separately manifest themselves directly in the improved structural perfection and electronic properties of the as-grown films. As stated before, a sharp (2×2) surface reconstruction stabilized in the narrow growth regime of 750 °C substrate temperature and a unity N-to-Ga flux ratio is essential for continuous GaN growth. In Fig. 8 we show x-ray rocking curves (XRCs) taken from three GaN-on-$6H$-SiC samples grown under different conditions. While the GaN(0002) reflection of sample C exhibits a linewidth of 8 arcmin, samples A and B have much larger values of 60 and 36 arcmin, respectively, and much weaker peak intensities. In addition, the peak position of sample A is shifted to a higher angle, indicating a tilt of the c axis of the GaN film and a considerable amount of strain in the GaN film. These data clearly show that nonoptimized nucleation conditions result in GaN nuclei having a large c-axis tilt and excess inhomogeneous built-in strain (samples of category I). Nonoptimized conditions for the subsequent growth (samples of category II) do not reduce the defect density and the inhomogeneous strain with increasing GaN film thickness. If both the nucleation and the subsequent growth conditions are optimized (samples of category III), the linewidth of the GaN(0002) reflection decreases linearly with increasing film thickness reaching a value of 1 arcmin at 0.8 μm.

The SAD pattern (left inset of Fig. 9) taken from the GaN/SiC interface region of sample C (category III) evidences the undistorted epitaxial relation of $[0001]_{GaN}\|[0001]_{SiC}$, $[11\bar{2}0]\|[11\bar{2}0]_{SiC}$. The nearly symmetric circular SAD features from the GaN layer also indicate that the epilayer is almost free from low-angle grain boundaries. The overview TEM image of the GaN film (right inset of Fig. 9) indicates the absence of columnarlike grains and a low density of stacking faults. Close to the GaN/SiC interface, a high density of defects is generated to relieve the lattice mismatch and to accommodate the large thermal mismatch between GaN and $6H$-SiC. It is noticeable that the defect density strongly decreases with thickness, i.e., by more than one order of magnitude within 200 nm. Finally, mainly threading dislocations are left, and the dislocation

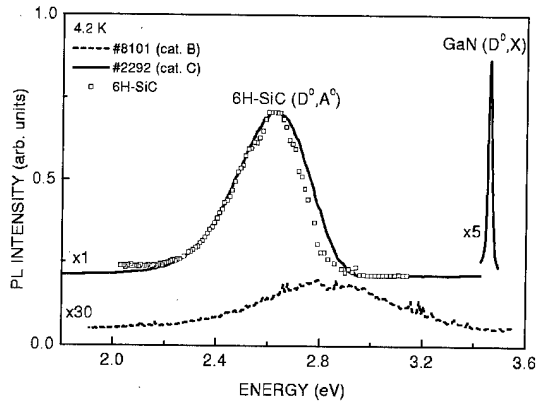


FIG. 10. PL spectra taken at 4.2 K from a sample of category III (solid line), sample of category II (dashed line), and the bare 6*H*-SiC wafer (open squares).

density reaches 5×10^9 cm^{-2} at a thickness of 0.5 μm. This value approaches state-of-the-art GaN films of the same thickness grown with an AlN buffer layer on 6*H*-SiC.[16] The HRTEM lattice image of Fig. 9 demonstrates that the GaN–SiC interface is atomically abrupt.

The impact of the growth conditions on the optical properties of the GaN-on-6*H*-SiC films is shown in Figs. 10 and 11. The low-temperature PL spectrum taken from a sample of category III (Fig. 10) exhibits a sharp but weak peak at the GaN band edge and a broad dominant band centered at about 2.6 eV. The broad PL band is ascribed to the (D^0,A^0) pair transitions in the 6*H*-SiC substrate. Owing to the specific energy band alignment of the wide-gap GaN film on the smaller-gap 6*H*-SiC substrate, photogenerated carriers are readily lost from the GaN to the substrate if the epilayer is thinner than the carrier diffusion length. For comparison, the closed squares in Fig. 10 display the PL spectrum from a piece of bare 6*H*-SiC. Note that both the peak position and the line shape coincide. Therefore, the appearance of this peak should not be confused with the yellow luminescence often observed from GaN layers of inferior quality. In fact,

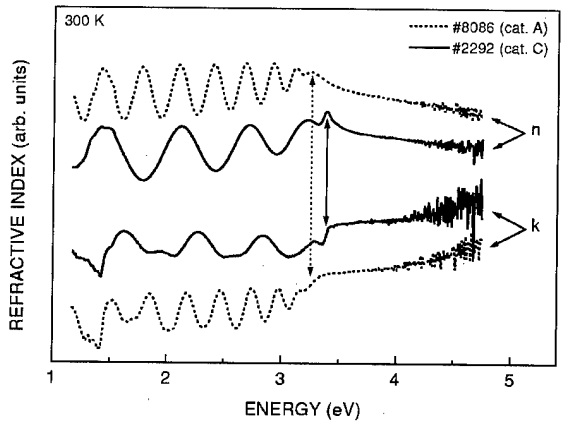


FIG. 11. Spectroscopic ellipsometry spectra taken at 300 K from samples of category I (dotted lines) and samples of category III (solid lines). The arrows indicate the 300 K GaN band-edge PL energy of the respective samples.

the better the quality of the GaN layer, i.e., the lower the background carrier concentration and the higher the carrier mobility, the more nonequilibrium carriers will be lost to the SiC substrate and hence the more intense the SiC luminescence will appear. The PL spectrum taken from a sample of category II, which was grown under nonoptimized conditions, covers almost the whole measured range and the band edge luminescence is absent. The very low integrated intensity clearly shows the existence of a high density of nonradiative recombination centers in this sample.

Finally, the spectroscopic ellipsometry measurements shown in Fig. 11 demonstrate the effect of growth conditions on the energy band structure of the GaN layers. The oscillations on the low-energy side arise from interferences of reflections from the GaN surface and the GaN/SiC interface and allow determination of the layer thickness (0.9 μm for the sample of category I and 0.43 μm for the sample of category III). The pronounced absorption edge features of sample C are in strong contrast to those of sample B, indicating the presence of inhomogeneous strain in this sample which modulates the local band gap energy. The optical data clearly show that under optimized nucleation and growth conditions GaN-on-6*H*-SiC layers with promising electronic properties can be obtained even without the necessity of inserting a buffer layer, such as, e.g., AlN.

IV. CONCLUSIONS

We have shown that the most important condition for the heteroepitaxial growth of GaN films with superior structural perfection is to optimize and control the nucleation stage and the subsequent layer growth independently. Using plasma-assisted MBE the growth rate, the N-to-Ga flux ratio and the substrate temperature can be adjusted carefully. Combined with *in situ* RHEED it allows accurate determination of the actual surface stoichiometry of the growing GaN nuclei and layer, which is crucial for the defect microstructure as well as for the phase purity. In the nucleation stage, sufficient wetting of GaN to the substrate surface, formation of the correct bonding at the interface, and minimization of the local microscopic lattice mismatch are the critical issues. The growth conditions must be adjusted such that the complete coverage of the substrate surface by GaN is very rapid, i.e., within 10 ML GaN. This demand is met by minimizing the size and maximizing the density of the GaN nuclei, i.e., at lower substrate temperature, lower growth rate, and higher N-to-Ga flux ratio. After complete wetting and relaxation of the strain by misfit dislocations, subsequent layer-by-layer growth of the GaN films needs a thorough readjustment of the growth parameters. If the proper conditions for nucleation and for subsequent layer growth as outlined before are fulfilled, the resulting GaN films are of high phase purity and of high structural perfection, the most important requirements for promising device applications.

ACKNOWLEDGMENTS

Part of this work has been sponsored by the Bundesministerium für Bildung, Wissenschaft, Forschung und Tech-

nologie of the Federal Republic of Germany. The authors gratefully acknowledge the technical assistance and the active contributions of B. Jenichen, M. Ramsteiner, and M. Reiche.

[1]O. Brandt, H. Yang, B. Jenichen, Y. Suzuki, L. Däweritz, and K. H. Ploog, Phys. Rev. B **52**, R2253 (1995).
[2]O. Brandt, H. Yang, and K. H. Ploog, Phys. Rev. B **54**, 4432 (1996).
[3]H. Yang, O. Brandt, and K. H. Ploog, Phys. Status Solidi B **194**, 109 (1996).
[4]O. Brandt, H. Yang, and K. H. Ploog, Mater. Sci. Eng., B **43**, 214 (1997).
[5]B. Yang, O. Brandt, A. Trampert, B. Jenichen, and K. H. Ploog, Appl. Surf. Sci. **123–124,** 1 (1998).
[6]B. Yang, A. Trampert, B. Jenichen, O. Brandt, and K. H. Ploog, J. Appl. Phys. **83,** 3800 (1998).
[7]B. Yang, A. Trampert, B. Jenichen, O. Brandt, and K. H. Ploog, Appl. Phys. Lett. (submitted).
[8]B. Yang, O. Brandt, Y. G. Zhang, A. Z. Li, B. Jenichen, G. Paris, and K. H. Ploog, Mater. Sci. Forum **264–268,** 1235 (1998).
[9]H. Yang, O. Brandt, and K. H. Ploog, J. Electron. Mater. **25**, 787 (1996).
[10]H. Yang, O. Brandt, M. Wassermeier, J. Behrend, H. P. Schönherr, and K. H. Ploog, Appl. Phys. Lett. **68**, 244 (1996).
[11]M. Wassermeier, A. Yamada, H. Yang, O. Brandt, J. Behrend, and K. H. Ploog, Surf. Sci. **385**, 178 (1997).
[12]O. Brandt, H. Yang, A. Trampert, M. Wassermeier, and K. H. Ploog, Appl. Phys. Lett. **71**, 473 (1997).
[13]A. Trampert, O. Brandt, H. Yang, and K. H. Ploog, Appl. Phys. Lett. **70**, 583 (1997).
[14]R. F. Davis, T. W. Weeks, Jr., M. D. Bremser, S. Tanaka, R. S. Kern, Z. Sitar, K. S. Ailey, W. G. Perry, and C. Wang, Mater. Res. Soc. Symp. Proc. **395**, 3 (1996).
[15]S. Y. Ren and J. D. Dow, J. Electron. Mater. **26**, 341 (1996).
[16]T. W. Weeks, Jr., M. D. Bremser, K. S. Ailey, E. Carlson, W. G. Perry, and R. F. Davis, Appl. Phys. Lett. **67**, 401 (1995).

Reflection high-energy electron diffraction studies of wurtzite GaN grown by molecular beam epitaxy

O. H. Hughes,[a] D. Korakakis,[b] T. S. Cheng, A. V. Blant, N. J. Jeffs, and C. T. Foxon
Department of Physics, University of Nottingham, Nottingham NG7 2RD, United Kingdom

(Received 21 January 1998; accepted 28 May 1998)

We report a comprehensive reflection high-energy electron diffraction study of the surface structure of GaN as a function of substrate temperature and III–V ratio for growth using elemental gallium and active nitrogen derived from a rf plasma source. An emission spectroscopy analysis of the composition of the nitrogen plasma showed that the neutral atomic species dominated the growth process. The effect of substrate pretreatment is also discussed. It was found that good quality growth accompanied by a reconstructed surface are only obtainable during growth under a slightly Ga-rich regime and after pretreatment by nitridation and the growth of a low temperature buffer layer. Reconstruction mode diagrams are presented both for layers during growth and also for layers which have been cooled after growth. The implications of these plots are discussed in terms of surface vacancy densities. © *1998 American Vacuum Society.* [S0734-211X(98)12104-2]

I. INTRODUCTION

Group III nitride semiconductors such as GaN and AlN show great promise for both light emitting diodes and lasers operating in the blue/ultraviolet (UV) part of the spectrum (due to their large band gaps) and for high-power high-frequency field effect transistors (FETS) (because of their excellent thermal conductivities and breakdown fields). There are, however, several fundamental obstacles in the way of production of high quality material. These difficulties, which are still a long way from being resolved, are in part associated with the relatively poor quality of the bulk epitaxial material available at the present time, but are also caused by the difficulty of controlling the quality of the surfaces and interfaces. A further barrier to progress arises from the fact that nitride semiconductors, unlike materials such as GaAs, are grown at temperatures which are relatively low compared with their melting points. Thus control of the surface stoichiometry during growth is more important with nitrides than it is with other III–V compounds. In the nitride field, successful device results have preceded fundamental understanding, particularly of the growth dynamics of the materials and the nature and stability of their reconstructed and nonreconstructed surfaces.

Group III nitride semiconductors have been grown by a variety of methods, including metalorganic vapor phase epitaxy (MOVPE), hydride vapor phase epitaxy (HVPE), molecular beam epitaxy (MBE), and chemical beam epitaxy CBE). In all cases, the stoichiometry of the material is a key issue and, in the latter two methods, this can be studied *in situ* using reflection high-energy electron diffraction (RHEED). RHEED studies of the growth of cubic (zinc-blende) GaN have shown the importance of controlling the surface reconstruction and hence surface stoichiometry in relation to the properties of the material,[1–3] the last study confirming previous work on cubic phase GaN which had shown that the presence of arsenic changed the growth polytype from hexagonal (wurtzite) to cubic (zinc blende).[4] For the technologically more important case of hexagonal material grown on sapphire or SiC substrates, there have been few detailed studies.[5–8] In all the above work, it was found that substrate temperature played a key role in determining the surface stoichiometry and, in the case of transparent substrates such as sapphire or SiC, this temperature is particularly difficult to measure.

In previous studies of the growth of group III nitrides by MBE, the importance of the choice of nitrogen source has been clearly demonstrated.[6] In particular the presence of nitrogen ions in addition to neutral atomic nitrogen has been shown to exert a profound influence on the surface reconstruction which in turn reflects a change in the electrical and optical properties, possibly due to departures from stoichiometry. In this latter study, deflector magnets were used to reduce the ion flux arriving at the substrate surface.

In this article we present a comprehensive study of the surface structure of GaN grown under various temperature conditions, a range of III–V ratios, and also with and without pregrowth preparation processes. In all cases growth was carried out using elemental gallium and active nitrogen derived from a rf plasma source. We also report on spectroscopic studies of the rf plasma source from which we deduce that the active component is atomic nitrogen, with very little evidence for the presence of molecular species and ions.

II. EXPERIMENTAL DETAILS

This work was carried out in a Varian modular GENII MBE reactor equipped with two rf plasma sources to provide active nitrogen — a liquid nitrogen cooled CARS25 and a water cooled HD25 (both manufactured by Oxford Applied Research Ltd.). The nitrogen gas supply was derived from liquid nitrogen and purified using an active metal getter system (UltraPure PF series). The nitrogen flow rate was controlled using a metal-sealed mass flow controller (MKS) and

[a]Electronic mail: Owain.Hughes@nottingham.ac.uk
[b]Joint post with the Department of Electrical & Electronic Engineering, University of Nottingham.

was typically in the range 1–4 sccm. All the remaining materials (namely, Ga, Al, In, Si, and Mg) were derived from conventional elemental evaporation sources.

The substrate temperature was monitored using an optical pyrometer but, because both GaN and sapphire are transparent in the wavelength range available to us, the temperatures indicated are largely relative. The measured temperature reflects that of the heater rather than that of the wafer itself.

The MBE reactor was evacuated by a combination of ion and cryo pumps, but during growth, only the cryopump was used. The system pressure during growth was typically $1-4\times10^{-5}$ Torr.

RHEED measurements were carried out using a VG LEG110 electron gun at a beam energy of 10–12 keV and a conventional phosphor. Diffraction patterns were recorded with a high sensitivity CCTV camera using a video recorder for studying growth dynamics and a video capture card for direct computer input of static images. To investigate the surface reconstruction, RHEED patterns were recorded for incident electron beams in both the $[1\bar{1}00]$ and $[11\bar{2}0]$ azimuths. These azimuths refer to the axes of the GaN film. The axes of the epitaxial layer are rotated 30° with respect to those of the sapphire substrate.

The emission spectrum of each plasma source was studied using an Oriel 0.25 m monochromator with a Si diode detector, with significant sensitivity from 300 to 1100 nm. Light was conducted to the monochromator entrance slit by means of an UV-compatible optic fiber and, to improve the signal to noise ratio, beam modulation at 75 Hz with lock-in detection was used. For the present experiments, a slit width of 1.24 mm was used giving a resolution of 4 nm. For the CARS25 source, emission spectra had to be viewed through the PBN-lined quartz crucible. In the HD25 source, however, there is direct line-of-sight into the plasma region. This difference is fundamental and will be discussed below.

III. EMISSION SPECTROSCOPY

Figure 1 shows a series of spectra at different rf power levels for the HD25 source at a constant nitrogen flow rate of 3 sccm. In Fig. 1, the strong emissions correspond to atomic nitrogen (e.g., 869 and 745 nm), and the weaker features are associated with molecular and ionized molecular species. As can be seen from Fig. 2, the intensities of the two dominant atomic lines increase monotonically with rf power, whereas the intensities of the molecular and ionized molecular lines show hardly any change. Figure 2 also demonstrates that there is a saturation and slight decrease with nitrogen flow rate for the 869 nm line, which is not apparent in the case of the 745 nm line. There is, however, no evidence for saturation with increasing rf power. The slight decrease in the intensity of the 869 nm line which is observed at high nitrogen flow rates may be indicative of increased recombination due to the high pressure within the plasma volume. If this were the case, however, one might expect the 745 nm line to show similar behavior, but it is likely that a very small decrease in the intensity would not be detectable in such a weak line.

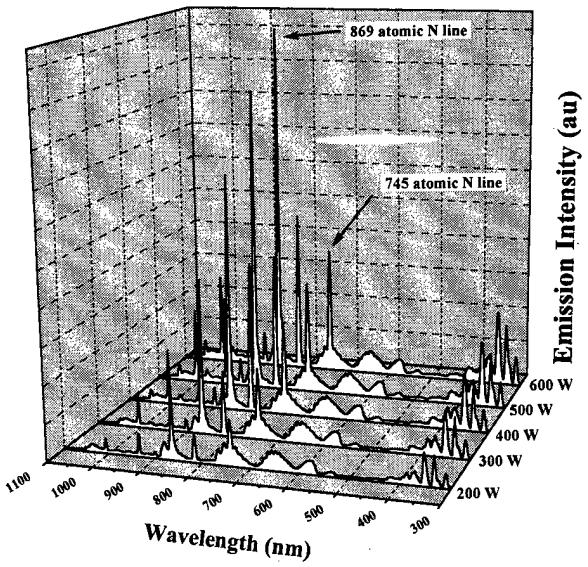


FIG. 1. Emission spectra of the HD25 nitrogen plasma source as a function of rf power.

Previous studies of the CARS source[9] showed an apparent lack of molecular and ionized molecular species, but this is an artifact of the measurement technique due to absorption of short wavelength radiation by the PBN-lined quartz crucible. It is common to find comparative studies of rf plasma sources in the literature. Such studies should be treated with caution since the absorption of short wavelength light by the crucible material eliminates the molecular and ionized molecular nitrogen lines in the emission spectrum, thus making the plasma sources appear to be artificially rich in atomic nitrogen.

IV. RHEED STUDIES

Using the CARS25 rf source with nitrogen flow rates of approximately 3 sccm and a rf power of 500 W, typical growth rates of approximately 0.3 μm/h were achieved. With the new HD25 source, however, under the same conditions, with the same source of gaseous nitrogen and with the same

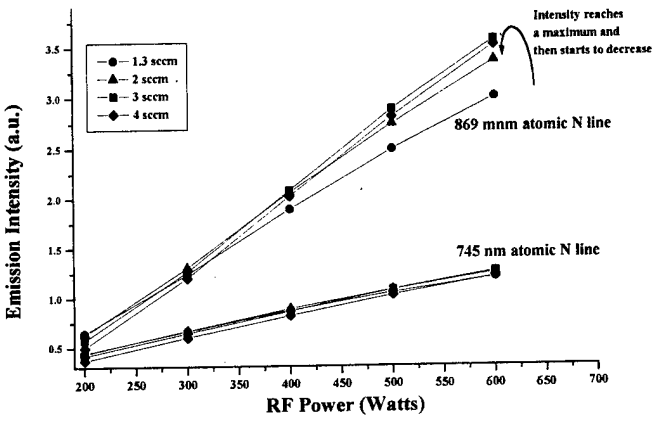


FIG. 2. Variation in intensity of the 745 and 869 nm atomic lines vs rf power for a range of nitrogen gas flow rates.

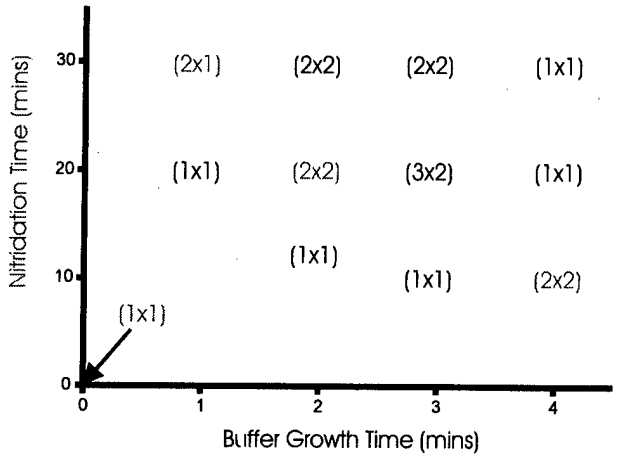


FIG. 3. Phase diagram showing the reconstructions observed during epitaxial growth of GaN films as a function of nitridation time and buffer layer thickness.

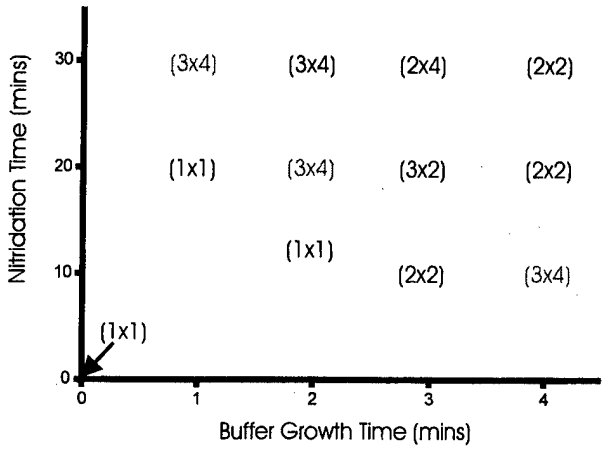


FIG. 4. Phase diagram showing the reconstructions observed upon cooling the epitaxial GaN films to 550 °C as a function of nitridation time and buffer layer thickness.

source–substrate separation, growth rates of 0.8 μm/h were achieved. This growth rate could be increased to 1 μm/h at 600 W rf power.

We have studied the growth of GaN on both GaAs(111) *B* and sapphire (0001) substrates, both with and without pretreatment. Sapphire is a polarized material and, whereas we give the growth direction as (0001), we have no way of knowing whether this is so, or whether growth was in fact in the $(000\bar{1})$ direction. The authors hope to very soon have access to convergent beam electron diffraction equipment which will enable this problem to be resolved unambiguously.

Nominal growth temperatures in the range 500–700 °C for GaAs substrates and 500–850 °C for sapphire were used. Both Ga-rich and N-rich growth regimes were investigated but it was found that good quality growth accompanied by clear RHEED patterns could only be obtained by growing just on the Ga-rich side of stoichiometric conditions. All the results reported below were obtained using such growth conditions. Without any pregrowth treatment, we observed only (1×1) unreconstructed surfaces for both plasma sources and both substrates. However, after pregrowth preparation we observed quite different behavior.

The effects of nitridation and the use of low-temperature buffer layers prior to the growth of the epitaxial films have been studied in a series experiments for growth on sapphire substrates. This study was carried out using the CARS25 source. All the samples were doped with Si to prevent electrostatic charging of the substrate surface. During the growth studies, nitridation times varying from 0 to 30 min and buffer layer thicknesses from 0 to 13 nm were used. The sapphire substrates were nitrided by exposure to the nitrogen plasma at 850 °C and the buffer layers were grown at a nominal temperature of 650 °C. The epitaxial GaN films were grown at 850 °C at a growth rate of 0.3 μm/h at a constant III/V ratio and at a fixed plasma power of 525 W and a nitrogen flow rate of 4 sccm.

Figure 3 shows a plot of the various reconstructions observed during growth and Fig. 4 shows the corresponding data obtained after cooling the films to a nominal temperature of 550 °C. Films grown with thin buffer layers and/or after short nitridation periods exhibit an unreconstructed (1×1) surface. Those based on thick buffer layers and more extended nitridation periods show evidence of surface reconstruction either during growth or subsequently during cooling. Surface reconstructions of (2×2), (3×2), and (3×4) were observed under different conditions as shown in Figs. 3 and 4. Figure 5 shows an example of a typical RHEED pattern from a reconstructed surface.

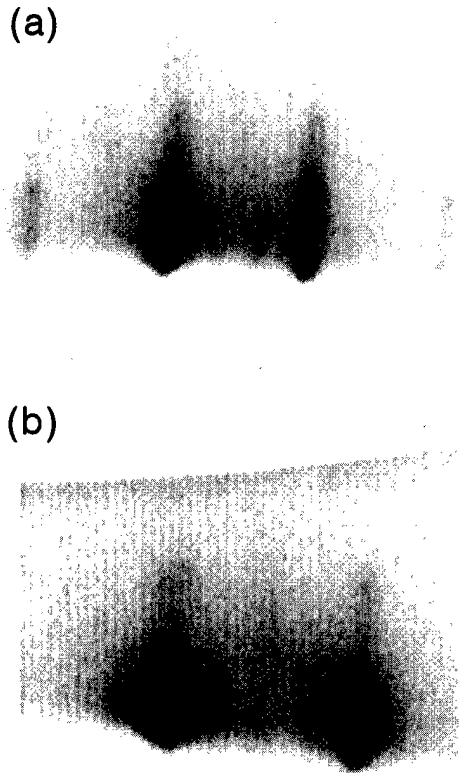


FIG. 5. Typical RHEED patterns taken along (a) the $[1\bar{1}00]$ and (b) the $[11\bar{2}0]$ azimuths of GaN.

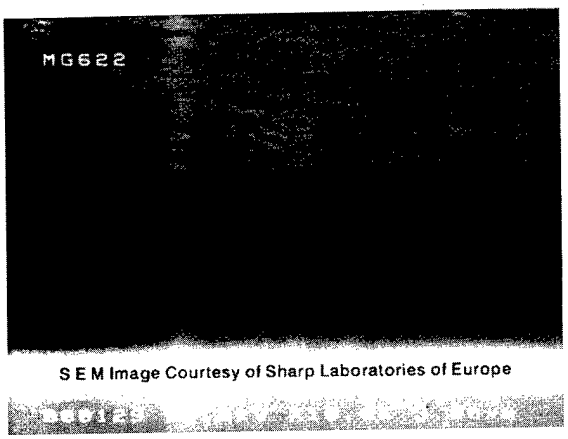


FIG. 6. High resolution scanning electron microscopy photograph of a 4 μm thick GaN epitaxial layer grown on sapphire at a growth rate of 0.8 μm/h.

Surface reconstructions have also been observed with thick buffer layers and extended nitridation times for films grown using a two-step process, showing that the important steps are independent of the details of the epitaxial growth but that the details of growth initiation are of paramount importance.

With the new HD25 source, using the same nominal temperatures for nitridation, buffer layer growth, and epitaxy, similar films have been produced at higher growth rates. Following the growth of the buffer layer, annealing at 750 °C induces a (2×2) reconstruction which persists as long as the temperature is kept constant. Heating to 850 °C, however, produces a gradual change back to an unreconstructed (1×1) surface.

Thicker films grown with the new HD25 source at higher growth rates show much improved electrical and optical properties, which will be reported elsewhere, but the surface morphologies are similar to those we have reported before, as shown in Fig. 6.

A consistent and reproducible pattern of growth behavior is revealed by the reconstructions listed above, but there is still no real explanation of what they signify in terms of atomic ordering. It has been suggested[10] that the various reconstructions of the GaN surface can be attributed to arrangements of Ga vacancies. The authors, and indeed other workers,[6] have observed that a reconstructed GaN surface can be returned to the unreconstructed state by flooding the surface with excess nitrogen. This fact would seem to suggest that, whereas Ga vacancies may exist on the surface, it is nitrogen vacancies which dominate the structure.

Further weight is given to this argument by the results of the cooling experiment reported above. The GaN layers were grown at temperatures that were well below the congruent evaporation temperature of the material. This contrasts with the growth of other III–V compounds such as GaAs which is grown well above its congruent point. Thus, during the growth of GaN, one would expect, on decreasing the substrate temperature, to move into a more Ga-rich regime. Comparison of Figs. 3 and 4 shows that, when a GaN surface is growing with a reconstructed surface at high temperature, a decrease in temperature to 550 °C produces a change to a reconstruction with lower symmetry and, when growing with an unreconstructed surface, it may change to a reconstructed state. This implies that driving the surface into a more Ga-rich regime lowers the symmetry of the growing surface. It therefore seems more likely that the surface reconstruction of the GaN surface is controlled by the existence of nitrogen vacancies.

V. DISCUSSION

Previous studies of the GaN growth process[6] have suggested that surface reconstruction can only be obtained with electron cyclotron resonance (ECR) sources by removal of the ions using magnetic deflection. In this study we have shown that with both the CARS25 and HD25 rf sources, using appropriate growth conditions, we can obtain reconstructed surfaces. In this present study the important parameters in determining the surface reconstruction are the extent of nitridation of the sapphire surface prior to growth and the thickness of the low-temperature buffer layer deposited before the deposition of the epitaxial film. This present study also indicates that it is possible to change the reconstruction by cooling the films after growth and shows that, so far, no complete picture of the surface phase diagrams has been obtained. There is evidence in the literature[7] that the best quality GaN is produced by a growth regime which corresponds to the borderline between the (1×1) unreconstructed surface and the (2×2) surface. The data shown in Figs. 3 and 4 may be of assistance in setting up the optimum pregrowth conditions for improved quality material.

Further, more detailed, studies are continuing and will be presented in the future. After submitting this article the authors became aware of recent detailed scanning tunneling microscopy studies by Smith *et al.*[11] and by Feenstra *et al.*[12] that address reconstruction issues of GaN prepared by cleaving. The surface structures may correspond to those we observe by RHEED.

ACKNOWLEDGMENTS

This work was supported by the European Commission via Brite Euram Contract No. BRPR-CT96-0334 "ANISET." The authors wish to thank Dr. S. E. Hooper and his colleagues at Sharp Laboratories of Europe, Oxford, for the scanning electron microscopy photograph, Fig. 6, published with their permission.

[1]O. Brandt, H. Yang, B. Jenichen, Y. Suzuki, L. Däweritz, and K. H. Ploog, Phys. Rev. B **52**, R2253 (1995).
[2]H. Yang, O. Brandt, M. Wassermeier, J. Behrend, and K. H. Ploog, Appl. Phys. Lett. **68**, 244 (1996).
[3]G. Feuillet, H. Hamaguchi, K. Ohta, P. L. Hacke, K. Okumura, and S. Yoshida, Appl. Phys. Lett. **70**, 1025 (1997).
[4]T. S. Cheng, L. C. Jenkins, S. E. Hooper, C. T. Foxon, J. W. Orton, and D. E. Lacklison, Appl. Phys. Lett. **66**, 1509 (1995).
[5]A. Kikuchi, H. Hoshi, and K. Kishino, Jpn. J. Appl. Phys., Part 1 **34**, 1153 (1995).

[6]K. Iwata, H. Asahi, S. J. Yu, K. Asami, H. Fujita, M. Fushida, and S. Gonda, Jpn. J. Appl. Phys., Part 2 **35**, L289 (1996).
[7]P. Hacke, G. Feuillet, H. Okumura, and S. Yoshida, Appl. Phys. Lett. **69**, 2507 (1996).
[8]P. Hacke, G. Feuillet, H. Okumura, and S. Yoshida, J. Cryst. Growth **175**, 94 (1997).
[9]A. V. Blant and T. S. Cheng (private communication).
[10]J. Neugebauer, T. Zywietz, M. Scheffler, J. E. Northrup, and C. G. Van de Walle, Phys. Rev. Lett. **80**, 3097 (1998).
[11]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
[12]A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, J. Vac. Sci. Technol. B, these proceedings.

Reconstructions of GaN(0001) and (000$\bar{1}$) surfaces: Ga-rich metallic structures

A. R. Smith and R. M. Feenstra[a)]
Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

D. W. Greve
Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

M. S. Shin and M. Skowronski
Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

J. Neugebauer
Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany

J. E. Northrup
Xerox Palo Alto Research Center, Palo Alto, California 94304

(Received 21 January 1998; accepted 11 May 1998)

Reconstructions of GaN(0001) and (000$\bar{1}$) surfaces are studied by scanning tunneling microscopy and spectroscopy, by electron diffraction, by Auger electron spectroscopy, and using first-principles theory. Attention is focused on Ga-rich reconstructions for each surface, which are found to have a metallic character involving significant overlap between Ga valence electrons. The electron counting rule is thus violated for these surfaces, but they nonetheless form minimum energy structures. © *1998 American Vacuum Society.* [S0734-211X(98)05404-3]

I. INTRODUCTION

Much effort in the past five years or so has been devoted to the study of gallium nitride, since its relatively large band gap (3.4 eV) makes it ideal for both optical applications in the blue-to-ultraviolet spectral range and in high power/high temperature electronic applications. Surface science studies of this material are just beginning, with recent reports of surface reconstructions for both cubic (zinc blende) and hexagonal (wurtzite) material.[1–9] A common theme regarding the growth of these surfaces (in the absence of hydrogen) is that stable growth occurs under *metal-rich* or *near-metal-rich* conditions, suggesting that GaN surfaces are stabilized by Ga atoms.[5–13] This behavior can be viewed as arising from the small size of nitrogen compared to gallium, so that reconstructions on the GaN surface are possible which involve purely Ga layers with Ga–Ga separations small enough to produce highly dispersive metallic bands, thereby producing low total energies (a similar situation has been reported recently for Sb-rich GaSb surfaces[14]). Such metallic surfaces violate the simple electron-counting rule,[15] but of course this rule is only meant to give a rough guide to the existence of minimum energy structures, and the highly dispersive metallic bands discussed here provide an alternative means of lowering the energy of a system.

In this article we discuss details of the structural and electronic properties of two reconstructions for wurtzite GaN: the 1×1 structure of the GaN(000$\bar{1}$) surface (or N face), and a pseudo-1×1 structure, denoted by "1×1," of the GaN(0001) surface (or Ga face). On the basis of scanning tunneling microscopy (STM) measurements and first-principles theory, the former reconstruction has recently been shown to consist of a single monolayer of Ga bonded to the outermost N atoms of the N-terminated bilayer on the GaN(000$\bar{1}$) face.[6] Not surprisingly, this structure, pictured in Fig. 1(a), is found to be metallic; theoretical and experimental evidence for this metallicity is presented in Sec. III A, including scanning tunneling spectroscopy (STS) measurements. This 1×1 arrangement of the GaN(000$\bar{1}$) surface is *not* the most Ga-rich structure possible on this surface—adding additional Ga adatoms produces a 3×3 arrangement [shown in Fig. 1(b)] and also 6×6 and $c(6\times12)$ arrangements.

The second structure discussed in this work, the "1×1" arrangement of GaN(0001), is less well understood than the GaN(000$\bar{1}$)1×1. Experimental evidence will be presented in Sec. III B based on STM, reflection high energy electron diffraction (RHEED), low energy electron diffraction (LEED), and Auger electron spectroscopy (AES) that this surface contains at least 2 monolayers (ML) (1 ML=1.14×10^{15} atoms/cm^2) of Ga residing on top of a Ga-terminated bilayer of the GaN(0001) surface. These Ga adlayers are found to have a structure well characterized by a discommensuration-fluid phase, similar to that observed on Au(111) and Pt(111).[16–18] Model calculations confirm that an incommensurate structure of Ga, containing about 7×7 unit cells of Ga in a 6×6 region of the GaN lattice, is energetically quite reasonable, although a structural model based on first-principles calculations has not yet been obtained. This "1×1" structure of GaN(0001) *is* the most Ga-rich structure found on this surface, and it is highly metallic, as

a)Electronic mail: feenstra@andrew.cmu.edu

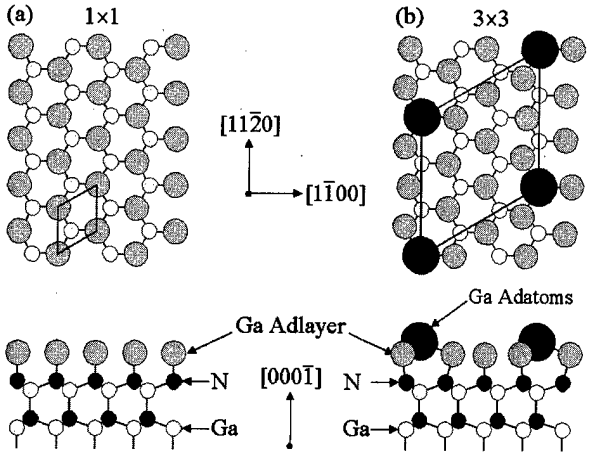


FIG. 1. Model structures determined for the (a) 1×1 Ga adlayer structure and (b) 3×3 adatom-on-adlayer structure of GaN(000$\bar{1}$). The Ga adlayer is under tensile stress since the Ga atoms are stretched further apart compared to their spacing in bulk Ga (3.19 Å compared ~2.7 Å). For the 3×3 structure, the adlayer atoms are able to get closer together by moving in the in-plane (lateral) direction away from the Ga adatoms by 0.51 Å, thus relieving the stress. All other lateral or vertical displacements of the adlayer atoms are less than 0.1 Å.

revealed by STS. Various other reconstructions containing less Ga have also been observed, and will be discussed in more detail elsewhere.[19]

II. EXPERIMENTAL DETAILS

The studies of GaN surfaces presented here are performed using a combination molecular beam epitaxy (MBE)/surface analysis system. Base pressures of both the MBE growth chamber and the analysis chamber are in the 10^{-11} Torr range. GaN films are grown using a standard Knudsen cell for the Ga and a rf plasma source to activate the N_2 (both built by SVT Associates). *In situ* surface analysis capabilities include RHEED, LEED, AES, and STM. GaN surfaces prepared in the MBE chamber are transferred under ultrahigh vacuum (UHV) conditions directly into the adjoining analysis chamber for investigation.

We have developed procedures for preparing both the GaN(0001) and GaN(000$\bar{1}$) faces of wurtzite GaN. Details of the growth of these two structurally inequivalent faces as well as the preparation of the individual reconstructions are discussed elsewhere.[6–9] Briefly, the (000$\bar{1}$) face is grown by nucleating the GaN directly on a solvent-cleaned and plasma-nitrided sapphire substrate. The 1×1 is then prepared by annealing the as-grown film at 800 °C, which removes excess Ga atoms. The (0001) face is grown by performing homoepitaxy on a metalorganic chemical vapor deposition (MOCVD)-grown GaN/sapphire substrate where the substrate is briefly nitrided prior to the homoepitaxial growth. The "1×1" is typically observed after termination of the homoepitaxy under Ga-rich conditions. The "1×1" can also be prepared by annealing the (0001) face at 750 °C, which removes Ga atoms, and subsequently redepositing ~2 ML of Ga, followed by a quick anneal to 700 °C.

III. RESULTS AND DISCUSSION

A. GaN(000$\bar{1}$)1×1 surface

The structure of the GaN(000$\bar{1}$)1×1 reconstruction is shown in Fig. 1(a). As discussed elsewhere, first-principles total energy calculations have been performed for this structure, along with calculations for a variety of other models for surfaces having either (0001) or (000$\bar{1}$) polarity.[6] The calculations are performed with the Ga 3*d* electrons included as valence electrons, and with a plane wave cutoff of 60 Ry. This approach has been employed in studies of the GaN(10$\bar{1}$0) surface,[1] the *c*-plane surfaces of AlN,[2] and for surfaces of cubic GaN.[3] The GaN(000$\bar{1}$)1×1 Ga adlayer model is the only 1×1 structure, of either polarity, which we have found that can account for a stable 1×1 symmetry structure in equilibrium conditions.[6] A similar result has been found for the *c*-plane AlN surfaces.[2]

The relative stability of this 1×1 adlayer structure arises in part from the strong Ga–Ga interaction within the adlayer itself. To see this, consider the following hypothetical reaction. Starting from a GaN(000$\bar{1}$)1×1 N-terminated bilayer and a bulk Ga reservoir, form the 1×1 Ga adlayer by removing atoms from the Ga reservoir and forming bonds with the N atoms on the GaN surface. The energy required to remove the Ga from the reservoir, thereby creating free Ga atoms, is the experimental cohesive energy of Ga (2.8 eV/atom). The energy released by forming the Ga–N bond is the bond strength of the Ga–N bond, 2.2 eV/atom. If these were the only two bonding mechanisms involved in the reaction, the reaction would be endothermic by 0.6 eV/atom. What is still missing from the analysis is the bonding within the Ga adlayer itself. The energy reduction due to the bonding of the Ga within the adlayer may be determined by a direct calculation of the formation energy of a free-standing hexagonal monolayer with a lattice constant of 3.19 Å. This calculation gives a formation energy of 1.0 eV/atom relative to bulk Ga, and so the cohesive energy of the monolayer is $-2.8+1.0=-1.8$ eV/atom. Thus the net reduction in energy in the formation of the Ga adlayer on GaN(000$\bar{1}$) is −1.2 eV/atom. The key point to be made here is that the bonding of the Ga within the adlayer is as important to the stability of the structure as the formation of the Ga–N bond itself. We also note that this estimate of the energy difference between the N-terminated bilayer and the Ga adlayer (1.2 eV/1×1) is almost identical to that determined by our direct calculation reported earlier.[6] Now a similar analysis may be performed for the GaAs(1$\bar{1}\bar{1}$) surface. However, in that case the Ga–Ga separation within an adlayer is much larger (4.0 Å). For such a Ga–Ga separation, calculations show that the cohesive energy of the adlayer is only −0.8 eV/atom. Since the Ga–As bond strength is 1.6 eV, the total reaction energy is $2.8-0.8-1.6=0.4$ eV/atom, and so the reaction is endothermic. It is therefore clear that the reduced Ga–Ga separation possible on the surfaces of GaN plays an important role in establishing the stability of Ga adlayer structures. A similar situation arises for the GaN(001) surface.[3]

The surface electronic structure has been calculated for

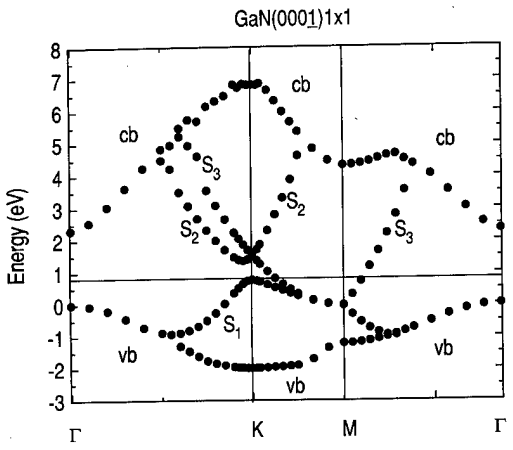


FIG. 2. Band structure for the GaN(000$\bar{1}$)1×1 Ga adlayer based on local density functional calculations. Energies are plotted relative to the VBM. The Fermi level is located 0.75 eV above the VBM. The plot shows the valence and conduction band edges and three surface states: S_1, S_2, and S_3. The computed bulk band gap of GaN is less than the experimental value (3.4 eV).

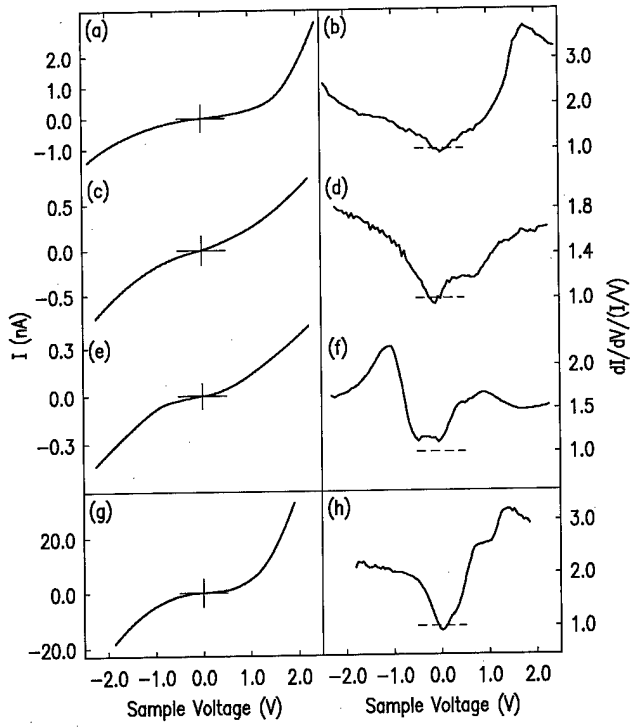


FIG. 3. Averaged tunneling spectroscopy results from three separate experiments (three different tips) on the GaN(000$\bar{1}$)1×1 surface (a)–(f) and a single experiment on the GaN(0001) "1×1" surface (g) and (h). $I-V$ curves are shown on the left with the corresponding normalized conductance curves shown on the right. Crossmarks represent the origins for the $I-V$ curves, while dashed lines indicate where the normalized conductance=1.

the GaN(000$\bar{1}$)1×1 Ga adlayer model and is shown in Fig. 2. This system is metallic, and the Fermi energy is located about 0.75 eV above the valence band maximum (VBM). There exist three highly dispersive surface states inside the band gap. These states are derived from the three p-state orbitals of the Ga adlayer atoms. The band labeled S_1 is fully occupied and has a predominantly p_z character with respect to the Ga atom. A remaining one-quarter electron per cell occupies the bottom of the S_3 band, which exhibits a p_x, p_y, p_z character with respect to the Ga adlayer atom. The S_2 band exhibits a minimum located about 0.6 eV above the Fermi level near the K point of the Brillouin zone. Such a minimum in $E(\mathbf{k})$ gives a step function contribution to the density of states, and it is therefore possible that the onset of tunneling for a bias voltage larger than 0.6 V could give rise to structure in the tunneling $I-V$ spectrum near this energy.

Experimental evidence for the metallicity of the GaN(000$\bar{1}$)1×1 surface has been obtained from STS measurements. The STM probe tip is positioned over a well-ordered region of the 1×1 surface, and then the tip–sample separation is held fixed while the tip–sample voltage (V) is varied and the tunnel current (I) is measured. Results are shown in Figs. 3(a)–3(f), where three representative spectra acquired using three different probe tips are displayed. The $I-V$ curves are shown on the left with the derived $(dI/dV)/(I/V)$ (normalized conductance) curves shown on the right. As is evident from the data, the three spectra are significantly different from each another. This largely represents differences between probe tips, none of which were well characterized for the purpose of spectroscopy.[20] However, the three spectra do have one very important feature in common, namely, that they all have a minimum in the normalized conductance at zero voltage which is very nearly equal to unity (as indicated in Fig. 3 by dashed lines). Such a feature is a defining characteristic of a metallic surface.[21] For a semiconducting surface, this minimum in the normalized conductance will be near zero. Thus, despite the variation in probe tips, the STS measurements clearly show that the 1×1 surface is metallic. In addition, we find that this 1×1 surface can be routinely imaged at tip–sample biases as low as 0.1 eV, also indicating its metallic character.

B. GaN(0001) "1×1" surface

We now turn to a discussion of the most Ga-rich reconstruction of the GaN(0001) surface, prepared as described in Sec. II. The diffraction patterns of this Ga-rich surface show mainly 1×1 streaks (RHEED) or spots (LEED), with sidebands in RHEED or satellite spots in LEED as described below. Hence we refer to this structure as "1×1," using the quotation marks to indicate that the symmetry is not truly 1×1. For this surface, the RHEED pattern at the growth temperature shows only 1×1 streaks, as illustrated in Fig. 4(a). However, as the surface is cooled down to <350 °C, distinct sidebands appear on the high wave vector sides of the first-order streaks along the [11$\bar{2}$0] azimuth, as shown in Fig. 4(b). Depending on the Ga coverage, the spacing of the sidebands from the first-order streaks at room temperature is either 0.16±0.01(≃1/6) or 0.08±0.01(≃1/12) of the 1× spacing $k_1=2.28$ Å^{-1}, as illustrated by the two LEED patterns shown in Figs. 4(c) and 4(d). For lack of better terminology, we refer to these structures as "$1+\frac{1}{6}$" and "$1+\frac{1}{12}$," respectively; the precise difference between these structures is not well understood at present. As shown in Fig.

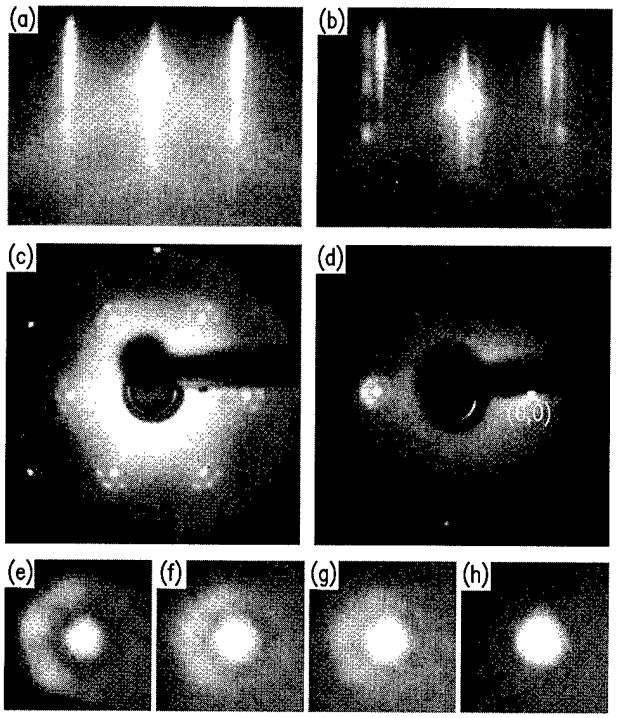


FIG. 4. 1×1 RHEED pattern for GaN(0001) (a) during growth, and (b) after cooling to below 350 °C, where it converts to a $1+\frac{1}{6}$ pattern. The $1+\frac{1}{6}$ LEED pattern ($E_{inc}=100$ eV) is shown in (c). For most "1×1" surfaces (see the text), a $1+\frac{1}{12}$ pattern is observed below 200 °C, as shown in (d) ($E_{inc}=40$ eV). LEED in the vicinity of the (0, 1) spot ($E_{inc}=40$ eV) at various temperatures: (e) room temperature–100 °C, (f) 100–150 °C, (g) 150–200 °C, and (h) above 350 °C.

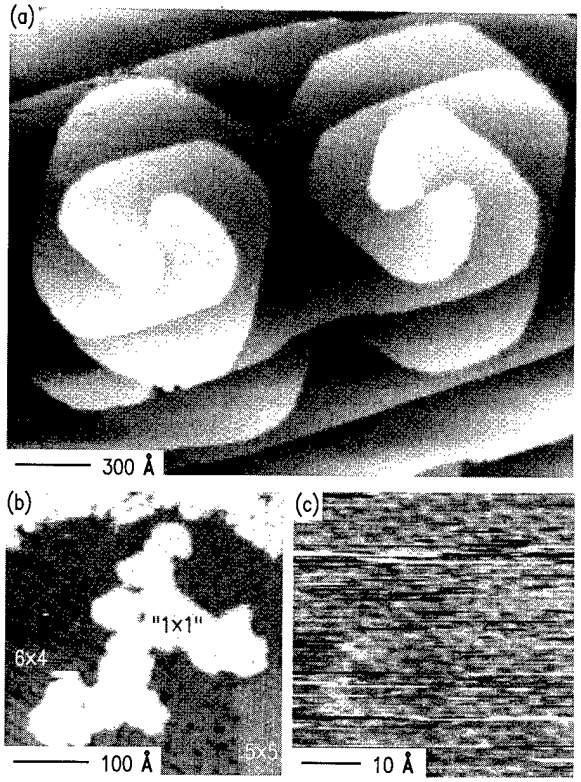


FIG. 5. STM images of (a) GaN(0001) "1×1" surface showing spiral growth, (b) (0001) surface with mixed 5×5, 6×4, and "1×1" reconstructions ("1×1" island height=2.1 Å), and (c) "1×1" reconstruction showing atomic resolution [lateral spacing=GaN lattice constant (3.19 Å) to within <1%]. The tunnel parameters are (a) −2.0 V at 0.1 nA, (b) −2.5 V at 0.075 nA, and (c) −0.25 V at 0.1 nA. The atomic steps seen in (a) are single bilayer high (2.59 Å) (line-by-line background subtraction has been applied to permit viewing of many terraces). The gray scale ranges are (b) 4.0 and (c) 0.27 Å, respectively.

4(c), the $1+\frac{1}{6}$ structure can exist down to room temperature for a narrow range of Ga coverage (just above that needed to form the 6×4, described elsewhere[8,9]), but for all higher coverages, the $1+\frac{1}{6}$ converts to $1+\frac{1}{12}$ as the temperature is reduced to about 200 °C.

The temperature dependence of the "1×1" surface is illustrated in Figs. 4(e)–4(h), focusing on the vicinity of the integral order (0, 1) spot. Between room temperature and about 100 °C, as seen in Fig. 4(e), a modulated ring of intensity with radius $0.08k_1$ is observed around the (0, 1) spot with modulation at 60° intervals. This ring has greater intensity on the high wave vector side of the spot. Similar asymmetric, modulated ring patterns have been seen for Pt(111) and Au(111).[16] As the temperature is increased to about 150 °C, the ring modulation decreases slightly [Fig. 4(f)]. As the surface temperature increases further to around 200 °C, the ring modulation decreases further [Fig. 4(g)]. It is also seen that the radius of the ring appears to have decreased slightly to about $0.07k_1$. As the temperature is raised past 200 °C, the pattern converts to $1+\frac{1}{6}$ (although not observed in this particular LEED experiment, the conversion from $1+\frac{1}{12}$ to $1+\frac{1}{6}$ in this temperature range has been observed consistently in RHEED experiments). Above 350 °C, one sees only the (0, 1) integral order LEED spot [Fig. 4(h)]. This sequence of phase transitions is reversible. Thus we find that the ring modulation decreases with increasing temperature. At the same time, the ring radius decreases slightly from $0.08k_1$ to $0.07k_1$ with increasing temperature until about 200 °C, at which point it increases by a discrete amount to $0.16k_1$. Identical diffraction patterns having the same temperature dependence have been reproducibly observed on numerous "1×1" surfaces prepared on grown films with various morphologies. Thus, the "1×1" patterns do *not* correlate with or depend on faceting or periodic step arrangements on the surface; instead, they suggest an incommensurate surface structure. Moreover, the modulated ring structure and its temperature dependence indicate that this incommensurate structure possesses a considerably dynamic, fluidlike character, even at room temperature. Thus, we infer that the "1×1" surface at room temperature is best characterized by a discommensuration-fluid phase, similar to that seen for Au(111) and Pt(111) at elevated temperatures ($T>0.64T_m$ for Au and $T>0.65T_m$ for Pt).[17] We note that since the melting point of bulk Ga (29.8 °C) is very near room temperature, such a structural phase for a Ga-rich surface is most reasonable. Furthermore, we also infer that, as temperature increases, the discommensuration-fluid phase converts to a disordered, fluid phase.

STM images acquired at room temperature for the GaN(0001) "1×1" surface are shown in Fig. 5. Since we

have not observed any difference between the $1+\frac{1}{6}$ and $1+\frac{1}{12}$ surfaces in the STM studies, we shall refer to them collectively as "1×1" here. Generally the "1×1" surface appears featureless (i.e., no corrugation) in the images, although small-scale images with a sharp tip do reveal atomic corrugation. Figure 5(a) shows a large-scale view of a surface which was imaged directly following Ga-rich growth without any further surface processing. It shows a typical spiral growth morphology where two dislocations, each with a screw component of their Burgers vector of $c[0001]$, are seen intersecting the surface and producing atomic steps. This surface was completely covered by the "1×1" arrangement and had a relatively high Ga coverage (at least 2 ML based on the Auger spectroscopy measurements discussed below). In contrast, the surface of Fig. 5(b) had a Ga coverage of only ~1 ML and was prepared by the annealing, redeposition, and reannealing procedure. For this lower Ga coverage, the surface contains islands of "1×1" surrounded by areas of 5×5 and 6×4 reconstruction. The precise structures of these latter two reconstructions are not known at present, although the 5×5 arrangement is thought to contain a combination of Ga adatoms, N adatoms, and possibly Ga vacancies.[19] Evidence suggesting a relatively high Ga coverage for the "1×1" reconstruction is also contained within the STM image of Fig. 5(b). The height of the "1×1" island above the surrounding 6×4 and 5×5 regions is 2.1 Å. Electronic effects can of course influence this height, but typically by only a few tenths of Å. The 5×5 and 6×4 regions are believed to contain adatoms with height (from theory) of 1.7 Å above the Ga atoms in the outermost GaN bilayer. Thus, we would estimate a thickness of the "1×1" Ga layer of 3.8 Å, corresponding to 1.8 ML. While this estimate is somewhat crude, it does suggest that the "1×1" reconstruction contains around 2 ML of Ga atoms.

For either the $1+\frac{1}{6}$ or $1+\frac{1}{12}$ surfaces, high resolution images reveal atomic corrugation, as seen in Fig. 5(c). However, the signal to noise ratio for these images is typically ~4× smaller than that found on the GaN(000$\bar{1}$)1×1 surface. Such a weak atomic corrugation is consistent with a highly metallic surface. Indeed, STS measurements reveal the surface metallicity, with a minimum in the normalized conductance at zero voltage very close to 1, as shown in Figs. 3(g) and 3(h). Careful measurements of the lateral period of the atomic corrugation, using tips which were calibrated on the GaN(000$\bar{1}$)3×3 reconstruction,[22] reveal that the period is identical to a 1× spacing (3.19 Å) to within <1%.

Auger spectroscopy measurements with an incident electron energy of 3 keV have been performed on the (000$\bar{1}$) "1×1" surfaces, as well as on all other (0001) and (000$\bar{1}$) reconstructions which we have studied, as a routine probe of Ga coverage. Experimental measurements for the (000$\bar{1}$) surfaces are plotted in Fig. 6(a), and those for the (0001) surfaces are plotted in Fig. 6(b). For almost all of the surface reconstructions, the ratio of intensities of the Ga (1055 eV) to N (379 eV) lines is in the range 0.6–0.9, with the exception being the "1×1" surface where this ratio is significantly higher (1.1–1.4). To interpret these measured intensity ratios, we perform model computations by summing intensity contributions from individual atomic layers over a sufficient number of layers extending into the surface to obtain convergence of the Ga/N ratio. We utilize Auger sensitivity factors of 0.12 for Ga and 0.33 for N (taken from Ref. 23). We choose electron escape depths of 14.0 and 9.7 Å for the 1055 and 379 eV electrons, respectively. These escape depths are chosen such that the Ga/N ratio for the GaN(000$\bar{1}$)1×1 reconstruction [indicated by the dashed line in Fig. 6(a)] corresponds to exactly one Ga adlayer located at a height of 1.99 Å above the nitrogen atoms of the last GaN bilayer, which is the known structure of this surface.[6] Using these values, Ga/N Auger intensity ratios are then computed for surfaces of either polarity having 0, 1, 2, and 3 layers of Ga sitting on top of the bulk-terminated bilayers. In the model computations for the (0001) surface, we assume a first additional layer of Ga 2.5 Å above the Ga-terminated bilayer, and successive Ga monolayers at 2.1 Å intervals, with all values based on theoretical results. For the GaN(000$\bar{1}$) surface, successive Ga monolayers after the first monolayer are also spaced at 2.1 Å intervals. The results of the model computations are given by the scales on the right-hand side of Fig. 6. As evidence for the success of the modeling, we note from Fig. 6(a) that the mixed $6\times6/c(6\times12)$ surface corresponds to a Ga coverage of 1.45 ML while the measured Ga coverage for the $c(6\times12)$ was (after correction for sticking coefficient) 1.44±0.02 ML, as reported previously[7] [this surface, while containing a few isolated patches of 6×6, was predominantly $c(6\times12)$, as observed by sweeping the RHEED beam laterally across the surface]. Thus the

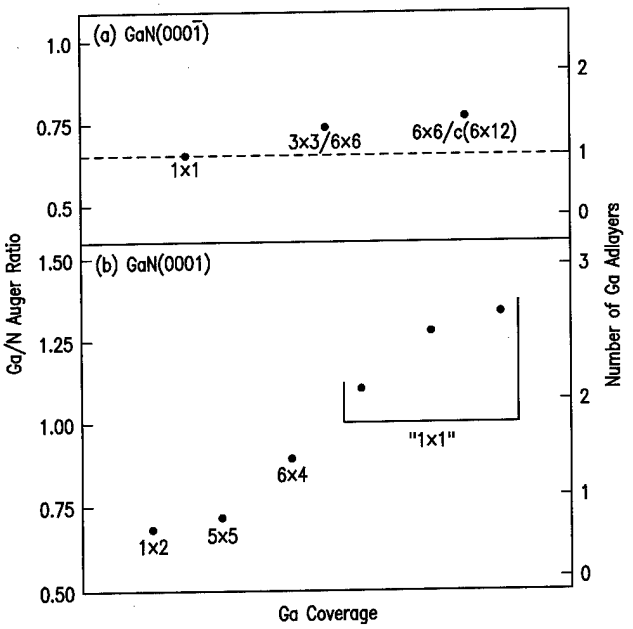


FIG. 6. Plots of Ga/N Auger intensity ratios for (a) GaN(000$\bar{1}$) reconstructions and (b) GaN(0001) reconstructions. The scales on the right are based on model calculations and represent the number of Ga monolayers sitting on top of the bulk-terminated bilayer for each polarity. Dashed line in (a) shows fit of 1×1 Auger ratio to 1 Ga adlayer on right-hand scale.

agreement between the Auger data and the model calculations is quite good.[24]

Consider now the results for the (0001) surface, shown in Fig. 6(b). As discussed elsewhere,[9] the ordered 5×5 is formed by depositing 1/2 ML Ga onto the annealed 1×2 surface and then briefly annealing that surface. Since the 1×2 is known to be disordered,[19] it is not unreasonable that the Ga/N Auger ratios of these two are similar. The 6×4 surface is similarly produced by depositing 1/2 ML Ga onto the 5×5 and briefly annealing that surface. Repeating this deposition and annealing cycle one or two more times results in a "1×1" surface. As seen from Fig. 6(b), this sequence of Ga deposition steps is in good agreement with the increase in Ga coverage from one reconstruction to the next, as deduced from the computed scale on the right-hand side. Based on these Auger results, it is quite clear that the (0001) "1×1" surfaces contain 2–3 additional monolayers of Ga above the Ga-terminated bilayer.

With all of the above experimental data on the "1×1" structure, let us now discuss possible structural models. For the GaN(0001) surface, the most stable structure we have theoretically obtained in the Ga-rich limit is the 2×2 Ga adatom model. In this model, the Ga adatom resides in the *T*4 site. Any proposed model for the high Ga coverage "1×1" phase should be more stable than the 2×2 *T*4 Ga adatom model in Ga-rich conditions. All of the true 1×1 structures that we have examined up to now, and which contain one additional monolayer or bilayer of Ga, are unstable with respect to this 2×2 adatom structure. This result is, of course, consistent with the apparent incommensurate, fluid-like nature of the "1×1" inferred from the diffraction analysis. Calculations performed for free-standing Ga monolayers or bilayers indicate that there is a driving force for a reduction in the in-plane Ga–Ga separation. Given this, and the experimental information discussed already, we consider a laterally contracted bilayer model for the "1×1" consisting of a Ga bilayer where the in-plane separation of the Ga atoms in the layers is contracted to a smaller value. We note that such a laterally contracted structure is not unreasonable for this system, since the GaN lattice constant of 3.19 Å is substantially greater than the typical Ga–Ga spacing of 2.7 Å in bulk Ga, so that a 1×1 arrangement of Ga is under considerable tensile strain.[6] We have performed total energy calculations for a free-standing Ga bilayer and have determined the formation energy $\Omega(a)=E(a)-2\mu_{\mathrm{Ga(bulk)}}$ as a function of a, the hexagonal lattice constant. The minimum in $\Omega(a)$ occurs for $a=2.7$ Å where Ω is equal to 0.46 eV/pair. Thus a free-standing hexagonal Ga bilayer is less stable than bulk Ga by about 0.23 eV/atom. In the contraction from $a=3.19$ to $a=2.7$ Å, the energy/pair of the bilayer is reduced by about $\Delta\Omega=0.68$ eV/pair.

We may employ these results to estimate the surface energy for a structure consisting of a 7×7 bilayer in approximate registry with a 6×6 GaN(0001) substrate. Such a structure would contain the equivalent of 2.7 Ga layers above the Ga-terminated bilayer, in agreement with that estimated from the Auger analysis. The estimated change in surface energy, relative to a 1×1 bilayer structure, may be broken down into three terms. The first term, E_1, is the cost of adding 13=7×7–6×6 additional pairs of Ga atoms to each 6×6 unit cell. The second term, E_2, is the energy benefit of the reduced lattice constant of the bilayer. The third term, E_3, is the energy cost of the imperfect registration of the incommensurate overlayer with the GaN(0001) substrate. E_1 is approximately $13\times\Omega(a=2.7)=6.0$ eV. E_2 is approximately $36\times\Delta\Omega=-24.5$ eV. From calculations for bilayers having different registrations with respect to the substrate, we estimate E_3 to be approximately 3.2 eV. The net effect is a reduction in surface energy of 0.43 eV/1×1 in the Ga-rich limit. This is close to the energy difference between the 2×2 adatom model and the best 1×1 bilayer model, 0.39 eV/1×1 in the Ga-rich limit. Thus it is plausible that such a laterally contracted bilayer structure could be stable under very Ga-rich conditions.

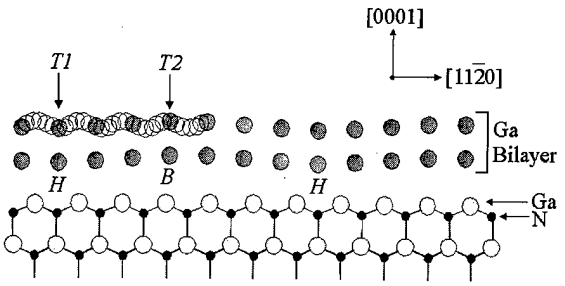


FIG. 7. Side view of possible structural model for the "1×1" surface (at a given instant in time) consisting of 2.7 ML of Ga sitting on top of the Ga-terminated bilayer. The open circles represent the various possible positions of first-layer Ga atoms plotted with respect to each of several GaN unit cells, and illustrate the time-averaged height of the first-layer Ga atoms and thus the 1×1 contour which the STM tip will follow. At a given instant in time, however, this incommensurate structure will manifest itself in diffraction as satellites surrounding the integral order peaks.

A schematic view of our proposed structure for the "1×1" surface is shown in Fig. 7. We consider the Ga bilayer (shown in dark gray circles), with uniform lateral spacing of the atoms of about 2.7 Å. In Fig. 7, the first layer atoms are positioned directly atop the second layer atoms. However, the energy difference between the top and hollow site registrations computed for free-standing Ga bilayers is very small. We expect that such a layer would be slightly buckled on the GaN surface since Ga atoms residing above hollow sites of the GaN surface (e.g., *T*4 sites above fourth layer nitrogen atoms), indicated by *H* in Fig. 7, would be slightly displaced towards the GaN, while Ga atoms residing in between such hollow sites, such as at position *B* in Fig. 7, would be slightly displaced away from the GaN. Such a model at this point is analogous to that used for the Au(111) surface,[17] except that we further assume, based on the diffraction results, that the structure is dynamic, with the Ga atoms moving around rapidly (this would probably imply the presence of vacancies or domain boundaries in the structure to allow the Ga bilayer space for such movement). The model shown in Fig. 7 is thus a picture of the structure at a given instant in time. Let us then consider what the appearance of this dynamic structure would be in STM images. A

surprising aspect of the STM results is the observation of precisely $1\times$ periodicity with $a=3.19$ Å, which appears to be inconsistent with the $1+\frac{1}{6}$ or $1+\frac{1}{12}$ inverse periods seen in diffraction. However, these STM measurements may be reconciled by taking into account the dynamic, fluidlike nature of the Ga bilayer. Consider a sharp STM tip as it scans over this structure. We assume that the time scale for the Ga bilayer motion is much shorter than the time the tip spends at each sampling point in the image. Hence, during the time that the tip is sitting over a given point on the surface, it senses a time average of the vertical positions of the first-layer Ga atoms as they move beneath the tip. This time average will include all possible translations of the incommensurate structure, and is illustrated conveniently by plotting the various possible positions of top Ga bilayer atoms with respect to each of several unit cells of the GaN lattice. These positions are indicated in Fig. 7 by the open circles. Thus, with the tip at position $T1$ over the hollow site H, the time-averaged height is relatively small (i.e., a corrugation minimum). Alternatively, with the tip at position $T2$ over an in-between site B, the time-averaged height is relatively large (i.e., a corrugation maximum). Thus, the STM image will appear to have a true 1×1 periodicity, as seen by the resulting contour of circles, arising from the periodicity of the top bilayer of GaN.

The diffraction patterns and their temperature dependence are accounted for in this model by the different orientational relationships for the surface discommensurations. For discussion purposes, we can define the discommensurations in our case as being associated with the B sites in Fig. 7, where the binding site of the Ga adatoms is in between two hollow sites. Below 100 °C, the data indicate that the discommensurations have a preferred spacing, and that they are aligned along particular crystal directions but, for 100–200 °C, they begin to lose their orientational ordering. Above 200 °C, the system converts into the $1+\frac{1}{6}$ structure, which may indicate a sudden change in the spacing of the discommensurations. Finally, above 350 °C, the system becomes further disordered, and the surface is then characterized by a completely disordered, fluid phase.

IV. CONCLUSIONS

In conclusion, we have investigated the reconstructions which occur on GaN(0001) and (000$\bar{1}$) surfaces. We emphasize that the energetically stable structures exhibit partially occupied surface states and are in direct violation of the electron counting rule (ECR). The ECR asserts that a semiconductor surface is stable only if all anion dangling bonds are doubly occupied and all cation dangling bonds are empty. It is further assumed that all cation dangling bonds are high in energy (close to the conduction band) and all anion dangling bonds are low in energy (close to the valence band). The calculated band structure shown in Fig. 2 for the GaN(000$\bar{1}$)1×1 surface shows that both assumptions are violated: while the Ga adlayer structure consists solely of cations, the occupied surface states ($S1$) are close in energy to the valence band maximum. The small lattice constant gives rise to strong Ga–Ga bonding even without bringing surface atoms together and forming, e.g., dimers, as commonly observed on other semiconductor surfaces. The strong Ga–Ga bonding thus not only stabilizes the 1×1 structure as discussed above, but it also significantly increases the dispersion of the cation surface states. In fact, the energetically lowest surface states are close to the valence band, and occupying these bands gives rise to energetically stable structures.

We have focused in particular on the 1×1 and "1×1" structures, respectively, which are both metallic in nature based on both experiment and theory. The "1×1" structure exhibits satellite peaks in the diffraction patterns below 350 °C, suggesting an incommensurate surface structure. The STM measurements, on the other hand, reveal a lateral atomic periodicity consistent with the surface GaN lattice constant. This apparent discrepancy is resolved by modeling the surface as a dynamic, fluidlike, discommensurate Ga bilayer structure with an increased surface atom density, and where mobile defects enable the motion to occur. Auger spectroscopy measurements reveal that this "1×1" structure is the most metal-rich structure out of all possible (0001) or (000$\bar{1}$) structures. Modeling of the Auger Ga/N peak intensity ratios as well as STM measurements of "1×1" island step heights also suggest a structure consisting of at least 2 ML of Ga on top of the Ga-terminated bilayer.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research under Contract No. N00014-96-1-0214.

[1]J. E. Northrup and J. Neugebauer, Phys. Rev. B **53**, 10 477 (1996).
[2]J. E. Northrup, R. Di Felice, and J. Neugebauer, Phys. Rev. B **55**, 13 878 (1997).
[3]J. Neugebauer, T. Zywietz, M. Scheffler, J. E. Northrup, and C. G. Van de Walle (unpublished).
[4]J. Fritsch, O. F. Sankey, K. E. Schmidt, and J. B. Page (unpublished).
[5]M. Wassermeier, A. Yamada, H. Yang, O. Brandt, J. Behrend, and K. H. Ploog, Surf. Sci. **385**, 178 (1997).
[6]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
[7]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. Northrup, Appl. Phys. A (to be published).
[8]A. R. Smith, V. Ramachandran, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, J. Vac Sci. Technol. A (to be published).
[9]A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, Appl. Phys. Lett. (submitted).
[10]P. Hacke, G. Feuillet, H. Okumura, and S. Yoshida, Appl. Phys. Lett. **69**, 2507 (1996).
[11]Z. Yu, S. L. Buczkowski, N. C. Giles, T. H. Myers, and M. R. Richards-Babb, Appl. Phys. Lett. **69**, 2731 (1996).
[12]L. T. Romano and T. H. Myers, Appl. Phys. Lett. **71**, 3486 (1997).
[13]E. J. Tarsa, B. Heying, X. H. Wu, P. Fini, S. P. DenBaars, and J. S. Speck, J. Appl. Phys. **82**, 5472 (1997).
[14]L. J. Whitman, P. M. Thibado, S. C. Erwin, B. R. Bennett, and B. V. Shanabrook, Phys. Rev. Lett. **79**, 693 (1997).
[15]W. Harrison, J. Vac. Sci. Technol. **16**, 1492 (1979); M. D. Pashley, Phys. Rev. B **40**, 10 481 (1989); J. E. Northrup and S. Froyen, *ibid.* **50**, 2015 (1994).
[16]A. R. Sandy, S. G. J. Mochrie, D. M. Zehner, G. Grübel, K. G. Huang, and D. Gibbs, Phys. Rev. Lett. **68**, 2192 (1992).
[17]D. L. Abernathy, D. Gibbs, G. Grübel, K. G. Huang, S. G. J. Mochrie, A. R. Sandy, and D. M. Zehner, Surf. Sci. **283**, 260 (1993).

[18]U. Harten, A. M. Lahee, J. P. Toennies, and Ch. Wöll, Phys. Rev. Lett. **54**, 2619 (1985).
[19]A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup (unpublished).
[20]Proper treatment of STM tips in order to achieve reproducible spectroscopic features has been discussed in detail elsewhere; see, for example, R. M. Feenstra, Phys. Rev. B **50**, 4561 (1994).
[21]R. M. Feenstra, J. A. Stroscio, and A. P. Fein, Surf. Sci. **181**, 295 (1987).
[22]The lateral calibration of the STM depends on the length of the probe tip, due to the bending motion of the tube scanner.
[23]L. E. Davis, N. C. MacDonald, P. W. Palmberg, G. E. Riach, and R. E. Weber, *Handbook of Auger Electron Spectroscopy*, 2nd ed. (Physical Electronics Division, Eden Prairie, MN, 1978), p. 13.
[24]We note that the Ga/N Auger ratio depends on the details of the given reconstruction. For example, the adatoms making up the GaN(000$\bar{1}$)3 $\times$3 reconstruction are known from theoretical calculation to be at a height of only 0.9 Å above the first Ga adlayer. If this value is used in the Auger model calculation, the computed Ga/N ratio for the 3$\times$3 is 0.70. Using a height of 2.1 Å above the first Ga adlayer yields a Ga/N ratio of 0.74. But of course the higher order reconstructions, such as 6 $\times$6 and $c(6\times12)$, may contain adatoms at heights greater than 0.9 Å above the first Ga adlayer but somewhat less that 2.1 Å. In fact, the calculated Ga/N ratios using both 0.9 and 2.1 Å adatom heights above the first Ga adlayer bracket the experimentally measured values shown in Fig. 6(a).

Soft x-ray emission studies of the bulk electronic structure of AlN, GaN, and $Al_{0.5}Ga_{0.5}N$

Kevin E. Smith,[a] Laurent C. Duda, Cristian B. Stagarescu,[b] and James Downes
Department of Physics, Boston University, Boston, Massachusetts 02215

Dimitris Korakakis, Raj Singh, and Theodore D. Moustakas
Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts 02215

Jinghua Guo and Joseph Nordgren
Department of Physics, Uppsala University, Uppsala S75121, Sweden

(Received 20 January 1998; accepted 20 May 1998)

The electronic structure of wurtzite GaN, $Al_{0.5}Ga_{0.5}N$, and AlN has been studied using synchrotron radiation excited soft x-ray emission spectroscopy. In particular, the elementally resolved partial densities of states has been measured and found to agree well with calculations. The shift in energy of the valence band maximum as x varies from 0 to 1 in $Al_xGa_{1-x}N$ was measured by recording N K-emission spectra, and found to be linear. Furthermore, N K-emission spectra revealed resonantlike hybridization of N $2p$ and Ga $3d$ states at 19 eV below the GaN valence band maximum. The spectral intensity of this feature is proportional to Ga content. © *1998 American Vacuum Society.* [S0734-211X(98)08804-0]

I. INTRODUCTION

Due to optoelectronic and high temperature device applications, GaN, InN, AlN, and related alloys have rapidly developed into an important class of wide band-gap semiconductor material.[1] Despite their importance, there is a significant lack of experimental data concerning the basic electronic structure of such materials. A fundamental understanding of electronic structure issues is required if these materials are to achieve their full technological potential, and only recently have the intrinsic band structure, the density of states, and the electronic properties of surfaces come under scrutiny. The standard experimental probe of valence band electronic structure is photoemission spectroscopy, and it has been successfully applied recently to study surface and bulk states in wurtzite GaN,[2] bulk states in cubic GaN,[3] GaN/AlN/InN heterojunction offsets,[4–7] and metal overlayer growth on GaN.[8,9] However, one major problem with photoemission is the necessity to prepare and maintain atomically clean surfaces; this has been extensively discussed by Bermudez.[10] A further drawback of photoemission is that, while it is the only spectroscopy that measures band dispersion (when performed in an angle resolved mode),[11,12] it does not measure the partial density of states (PDOS). Rather, angle integrated photoemission measures the joint density of states.[11,12] The surface sensitivity of photoemission comes from the small inelastic mean free path of low energy electrons. Photons of equivalent energies have a much greater escape depth (typically 1000 Å), and thus soft x-ray emission (SXE) spectroscopy is able to directly measure bulk electronic structure. Due to the dominant atomic dipole selection rules that govern the x-ray emission process, SXE measures the elementally resolved PDOS.[13]

We present here results of a study of the bulk electronic structure of $Al_xGa_{1-x}N$ ($x=0$, 0.5, 1) using synchrotron radiation excited SXE. (Results for GaN are discussed more fully in an earlier publication.[14]) The occupied valence band PDOS for N, Ga, and Al have been studied by measuring K and L edge emission spectra. We find that the measured PDOS agree well with theory.[15] X-ray emission has been used for the first time as a tool to monitor motion of the valence band maximum (VBM) in $Al_xGa_{1-x}N$, and we find the VBM varies linearly with Al concentration, x.

This article is structured as follows. In Sec. II we discuss the experimental details, while in Sec. III we briefly review the principles of SXE spectroscopy. Section IV presents our measurements for $Al_xGa_{1-x}N$ and discusses our results in the context of recent band structure measurements.

II. EXPERIMENTAL DETAILS

Our samples were thin film wurtzite $Al_xGa_{1-x}N$ grown using electron cyclotron resonance assisted molecular beam epitaxy on sapphire substrates. The samples were n type, and the growth procedure has been reported elsewhere.[16] The thickness of the samples was approximately 0.5×10^{-6} m (as determined by scanning electron microscopy). Experiments were performed at the undulator beamline BW3 at HASYLAB/DESY in Hamburg, Germany, and on beamline X1B at the National Synchrotron Light Source, Brookhaven National Laboratory. Emission spectra were recorded using a Nordgren-type grazing-incidence grating spectrometer using a 5 m, 1200 lines/mm grating in first order of diffraction at a resolution of approximately 0.8 eV.[17,18] The acquisition time for individual spectra was approximately 90 min. The base pressure in the experimental system was 1.0×10^{-8} Torr. This vacuum is quite adequate since SXE is primarily a bulk probe, and surface phenomena were not under investigation.

[a]Author to whom correspondence should be addressed; electronic mail: ksmith@bu.edu
[b]On leave from the Institute of Microtechnology, Bucharest, Romania.

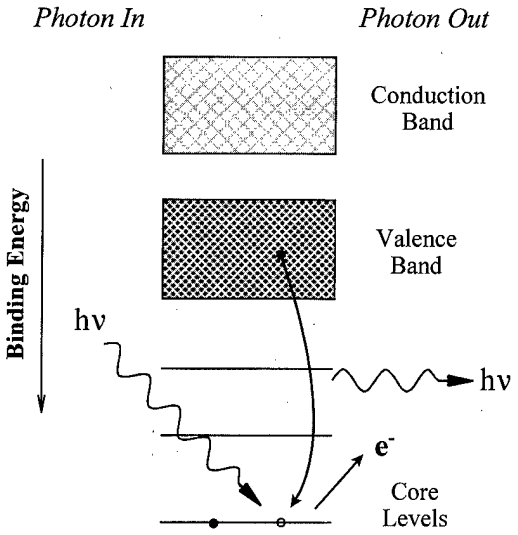


FIG. 1. Schematic illustration of the x-ray emission process.

III. SOFT X-RAY EMISSION SPECTROSCOPY

As mentioned in the introduction, SXE is a bulk spectroscopy that measures the PDOS of occupied valence states. The basic SXE process is illustrated in Fig. 1. Monochromatic light from a synchrotron radiation source is incident on the sample and excites a core electron. The resultant core hole is filled by electrons making transitions from higher lying states. Most core holes are filled via Auger processes. However, a small fraction decay radiatively with the emission of a photon. The radiative decay process is governed by dipole selection rules. Thus if photons resulting from transitions from the valence band to the core state are measured using a suitably high resolution spectrometer, the spectrum reflects the angular momentum resolved density of states, or PDOS, if the final state rule is valid.[19] (The final state rule states that the spectral features that appear in x-ray spectroscopy correspond to the theoretical DOS calculated with the final state potential.) Furthermore, since the probability of an interatomic decay is much smaller than that for an intra-atomic event, the PDOS is in fact elementally specific in a multi-element system. By using monochromatic synchrotron radiation to excite the core hole at threshold, this elemental specificity is further enhanced.[20,21]

We compare our SXE spectra of $Al_xGa_{1-x}N$ with the *partial* density of states taken from Xu and Ching.[15] The procedures used for setting a binding energy scale are discussed fully in our earlier study of pure GaN.[14] Briefly, we first calibrate our spectrometer detector by using the elastic scatting of incident photons at various energies. The accuracy of this method depends on the combined instrumental resolution of the monochromator and the spectrometer as well as on the relative visibility (number of counts) of the elastic peak. (The monochromator calibration is accurate to within 5%, and since we are using relative energy differences in the spectra, we use the monochrometer setting as the photon energy. Clearly the *absolute* energy scale, which is not of primary interest here, is less accurate.) This allows us to plot

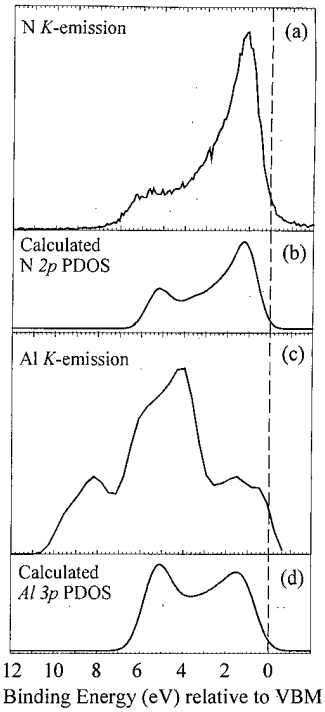


FIG. 2. Comparison of experimental N and Al K-emission spectra from pure AlN with a broadened PDOS from Ref. 15. The energy scale is relative to the experimentally determined N $2p$ valence band maximum (VBM).

the SXE spectrum as a function of photon energy, i.e., the emission spectrum is plotted on the same energy scale as the absorption spectrum. For situations where the elastic scattering was too weak to be visible in our spectrometer, we instead record inner shell transitions and calibrate against known spin orbit splittings.[14] To put the N emission spectra on a *binding energy* scale, we next determine the experimental VBM by extrapolating the leading edge of the emission spectrum to zero.

IV. RESULTS AND DISCUSSION

Figure 2 presents SXE spectra for both the Al and N valence band PDOS in pure AlN. Also included in Fig. 2 are the results of a first principles orthogonalized linear combination of atomic orbitals (OLCAOs) calculation in the local density approximation of the N $2p$ and Al $3p$ PDOS in AlN.[15] The theoretical PDOS was convoluted by both Lorentzians to simulate core-hole lifetime broadening and by Gaussians to simulate the instrumental broadening. While the agreement between the spectral shape and band width for the N $2p$ states is quite good [Figs. 2(a) and 2(b)], the measured Al $3p$ contribution to the valence band is significantly broader than predicted [Figs. 2(c) and 2(d)]. Low signal strength required that the spectrum in Fig. 2(c) be taken with a resolution of 1.5 eV. However, this poor resolution does not account for the difference in bandwidth. A similarly larger than predicted bandwidth has been observed for AlN by photoemission spectroscopy.[4–6]

Considering the N $2p$ states, we have studied these systematically in the $Al_xGa_{1-x}N$ alloy system by recording N

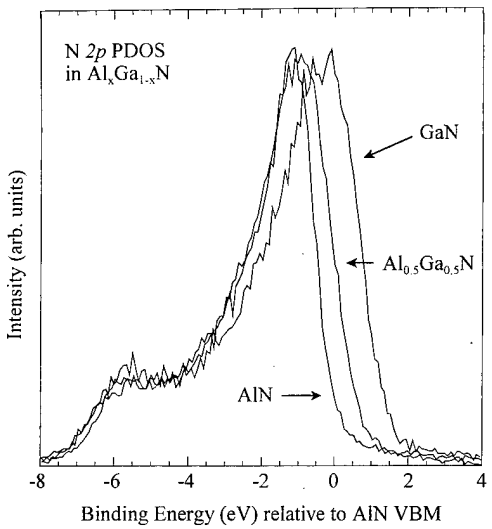


FIG. 3. N K-emission spectra from $Al_xGa_{1-x}N$ ($x=0$, 0.5, 1.0) Only the valence band emission is shown. The spectra are not shifted in energy, but plotted relative to the binding energy of the VBM of pure AlN ($x=1$).

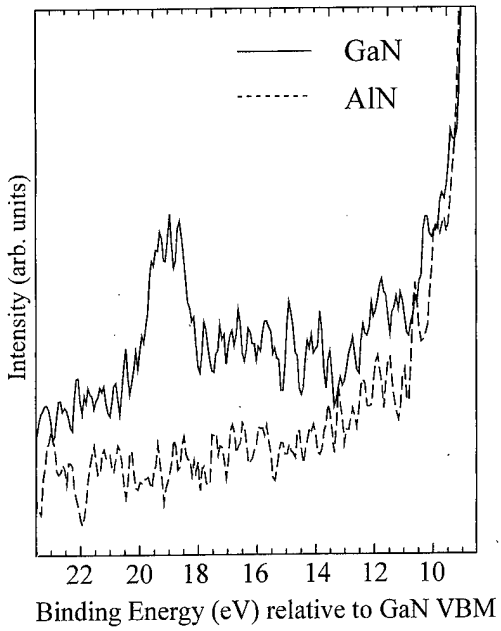


FIG. 4. N K-emission spectra from pure AlN and pure GaN. This is the same data as in Fig. 3, but the emission below the valence band is shown. Note the emission at 19 eV below the GaN VBM. This peak disappears as the Ga is replaced by Al. It originates from N $2p$ states hybridized with Ga $3d$ states.

K-emission spectra for $x=0$, 0.5, and 1. Figure 3 presents SXE spectra from $Al_xGa_{1-x}N$. The incident photon energy was set above the adsorption edge, nominally at 405 eV. These spectra are referenced to the AlN VBM as determined from the $x=1$ spectrum. A significant narrowing and peak shift downward of the $Al_xGa_{1-x}N$ valence band with increasing Al concentration is clearly visible. The binding energy of the bottom of the valence band shifts upwards by approximately 0.25 eV, while the top of the valence band moves linearly upwards by 1.4 eV. Theoretically, the behavior of the gap in such an alloy system may be described by the following formula:[22]

$$E_g(x)=\bar{E}_g+\Delta E_g(x-\tfrac{1}{2})-bx(1-x).$$

Here, $\bar{E}_g$ is the average gap, ΔE_g is the difference between the gaps of the end members of the alloy system, and b is the bowing parameter. Values of b ranging from approximately +1.0 eV[23,24] through $b=0$ eV (Ref. 25) to approximately $b=-0.8$ eV (Ref. 26) have been reported. A recent optical absorption study of high quality $Al_xGa_{1-x}N$ films found $b=0$ eV.[27] Our data indicate that the VBM varies linearly with x, and that the band gap opens symmetrically. The latter conclusion comes from the observation that the difference in the bulk band gaps of AlN and GaN is 2.8 eV, and our linear VBM motion is 1.4 eV. If the gap indeed opens symmetrically, then our measurement of a linear motion of the VBM is consistent with a value of $b=0$ eV.

In our earlier study of pure GaN we attributed a very weak feature observed at 19 eV below the VBM in the N K emission to hybridization between N $2p$ and Ga $3d$ states.[14] A possible alternative explanation would be emission due to N $2s-1s$ transitions which would occur in AlN as well as GaN, but such dipole-forbidden transitions are extremely unlikely. However, the poor signal-to-noise ratio in our original measurement called for a re-examination that could conclusively identify this emission as being due to hybrid states.

Figure 4 shows the N K emission below the valence band minimum for two samples: pure AlN and pure GaN. (Note that here the spectra are plotted relative to the GaN VBM.) The emission at 19 eV is clearly visible in the GaN sample and totally absent in AlN. Pure AlN has no Ga $3d$ states with which the N $2p$ states can hybridize, and thus the feature disappears. Moreover, the spectral width of this peak is quite narrow (resonantlike) whereas the N $2s$ states of GaN are known from photoemission experiments to be quite broad in energy.[3] These findings prove conclusively that this is indeed a hybridization phenomenon and not due to dipole-forbidden $2s-1s$ transitions.

V. CONCLUSIONS

We have studied the electronic structure of the $Al_xGa_{1-x}N$ alloy system using soft x-ray emission spectroscopy. Comparison to an *ab initio* calculation of the valence band PDOS shows good agreement. The valence band narrows with increasing Al content, but the bottom of the valence band remains essentially fixed. We estimate from these spectra that the bowing parameter b for $Al_xGa_{1-x}N$ is zero. Finally, a sharp emission feature at 19 eV below the GaN valence band maximum was proved to originate from N $2p$ states hybridizing resonantly with Ga $3d$ states.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. DMR-9504948. Our x-ray emission spectrometer is funded by the U.S. Army Research Office under Grant No. DAAH04-95-0014. Some experiments were performed at the NSLS which is supported by the U.S.

Department of Energy, Divisions of Materials and Chemical Sciences. The Uppsala program is supported in part by the Swedish Natural Research Council (NFR) and the Göran Gustafsson Foundation. TDM acknowledges the support of DoD/ARPA under Grant No. MDA972-96-3-0014. LCD gratefully acknowledges a grant from NFR for a postdoctoral fellowship at Boston University.

[1]*Diamond, Silicon Carbide and Nitride Wide Bandgap Semiconductors*, Mater. Res. Soc. Symp. Proc., edited by C. H. Carter, Jr. (Materials Research Society, Pittsburgh, PA, 1994), Vol. 339.
[2]S. S. Dhesi, C. B. Stagarescu, K. E. Smith, D. Doppalapudi, R. Singh, and T. D. Moustakas, Phys. Rev. B **56**, 10271 (1997).
[3]S. A. Ding, G. Neuhold, J. H. Weaver, P. Haberle, K. Horn, O. Brandt, H. Yang, and K. Ploog, J. Vac. Sci. Technol. A **14**, 819 (1996).
[4]G. Martin, A. Botchkarev, A. Rockett, and H. Morkoc, Appl. Phys. Lett. **68**, 2541 (1996).
[5]G. Martin, S. Strite, A. Botchkarev, A. Agarwal, A. Rockett, W. R. L. Lambrecht, B. Segall, and H. Morkoc, J. Electron. Mater. **24**, 225 (1995).
[6]G. Martin, S. Strite, A. Botchkarev, A. Agarwal, A. Rockett, H. Morkoc, W. R. L. Lambrecht, and B. Segall, Appl. Phys. Lett. **65**, 610 (1994).
[7]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **68**, 2879 (1996).
[8]V. M. Bermudez, T. M. Jung, K. Doverspike, and A. E. Wickenden, J. Appl. Phys. **79**, 110 (1996).
[9]V. M. Bermudez, R. Kaplan, M. A. Khan, and J. N. Kuznia, Phys. Rev. B **48**, 2436 (1993).
[10]V. M. Bermudez, Appl. Surf. Sci. **126**, 69 (1998).
[11]*Angle Resolved Photoemission*, edited by S. D. Kevan (Elsevier, Amsterdam, 1991).
[12]K. E. Smith and S. D. Kevan, Solid State Chem. **21**, 49 (1991).
[13]L. G. Paratt, Rev. Mod. Phys. **31**, 616 (1959).
[14]C. B. Stagarescu, L.-C. Duda, K. E. Smith, J. H. Guo, J. Nordgren, R. Singh, and T. D. Moustakas, Phys. Rev. B **54**, 17335 (1996).
[15]Y.-N. Xu and W. Y. Ching, Phys. Rev. B **48**, 4335 (1993).
[16]T. Lei, M. Fanciulli, R. J. Molnar, T. D. Moustakas, R. J. Graham, and J. Scanlon, Appl. Phys. Lett. **59**, 944 (1991).
[17]J. Nordgren, G. Bray, S. Cramm, R. Nyholm, J.-E. Rubensson, and N. Wassdahl, Rev. Sci. Instrum. **60**, 1690 (1989).
[18]J. Nordgren and R. Nyholm, Nucl. Instrum. Methods Phys. Res. A **246**, 242 (1986).
[19]U. v. Barth and G. Grossmann, Phys. Rev. B **25**, 5150 (1982).
[20]J. Nordgren, J. Phys. (Paris) Colloq. **9**, 693 (1987).
[21]J. Nordgren and N. Wassdahl, Phys. Scr. **T31**, 103 (1989).
[22]E. A. Albanesi, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **48**, 17841 (1993).
[23]M. R. H. Khan, Y. Koide, H. Itoh, N. Sawaki, and I. Akasaki, Solid State Commun. **60**, 509 (1986).
[24]Y. Koide, H. Itoh, M. R. H. Khan, K. Hiramatsu, N. Sawaki, and I. Akasaki, J. Appl. Phys. **61**, 4540 (1987).
[25]M. A. Kahn, R. A. Skogman, R. G. Schulze, and M. Gershenzon, Appl. Phys. Lett. **43**, 492 (1983).
[26]S. Yoshida, S. Misawa, and S. Gonda, J. Appl. Phys. **53**, 6844 (1982).
[27]D. K. Wickenden, C. B. Bargeron, W. A. Bryden, J. Miragliotta, and T. J. Kistenmacher, Appl. Phys. Lett. **65**, 2024 (1994).

Formation of an Sb–N compound during nitridation of InSb (001) substrates using atomic nitrogen

L. Haworth,[a] J. Lu, P. Hill, D. I. Westwood, and J. E. Macdonald
Department of Physics and Astronomy, University of Wales, P.O. Box 913, Cardiff, United Kingdom CF2 3YB

N. Hartmann, A. Schneider, and D. R. T. Zahn
Halbleiterphysik, TU-Chemnitz, Germany D-09107

(Received 21 January 1998; accepted 20 May 1998)

The effect of atomic nitrogen, generated by a radio frequency plasma source, on clean InSb(001) at 275 °C has been studied using x-ray photoelectron spectroscopy (XPS) and resonant Raman scattering (RRS). Chemically shifted XPS features of the Sb 3*d* region revealed the formation of a reacted Sb species. This reacted Sb was unambiguously identified as mainly Sb–N by comparison with results from as deposited and nitrided, thick elemental Sb layers on InSb. The Sb 3*d* feature due to this Sb–N species was found to have a chemical shift of 1.65±0.10 eV to higher binding energy compared with the InSb peak, while for the elemental Sb the shift was only 0.45±0.10 eV in the same direction. Although not obvious from the XPS data the RRS spectra of a much longer nitridation at 275 °C showed the presence of crystalline elemental Sb. Annealing studies of elemental Sb and nitrided Sb layers showed the Sb–N species to be significantly less volatile than elemental Sb. © *1998 American Vacuum Society.* [S0734-211X(98)08704-6]

I. INTRODUCTION

With GaN and related materials becoming increasingly important in the fields of visible optoelectronic devices[1] and high power electronics[2] there is considerable interest in the details of their growth processes. This is fueled by the fact that, despite the success of GaN devices, the substrates that are used (primarily sapphire and silicon carbide) are largely mismatched to GaN and hence the epitaxial material contains many defects. As a result many different substrate materials have been examined[3] and one strand of research that has proved interesting is growth onto cubic substrates, predominantly III-V materials, which can result in growth of the cubic phase GaN rather than the usual wurtzite phase.[4–6] On GaAs (001), the phase that is produced has been found to be heavily dependent on the very first stages of the growth process and particularly the initial nitridation of the substrate.[5] Initial expectations of nitridation of III-V substrate materials such as GaAs and InSb predicted a simple anion exchange reaction with N replacing As (or Sb)[6]

$$\mathrm{InSb}_{(s)} + \mathrm{N}_{(g)} \Leftrightarrow \mathrm{InN}_{(s)} + \mathrm{Sb}_{(?)}\,. \tag{1}$$

Here, the question mark denotes the uncertain volatility of the reacted Sb, which will depend upon substrate temperature. However , in the case of nitridation of GaAs, evidence has been found, using both hydrazoic acid as a precursor and radio frequency plasma produced atomic nitrogen, that an As–N compound is also formed.[7,8] In this study nitridation of clean InSb (001) surfaces with atomic nitrogen has been examined using both x-ray photoemission spectroscopy (XPS) and resonant Raman scattering (RRS). The study of the interaction of atomic N with InSb is of interest both fundamentally, due to its link to the present technological interests in group III nitrides, and in its own right as a precursor to the growth of GaN on InSb. The lattice constants of cubic GaN and InSb differ by a factor of $\sim\sqrt{2}$ offering the tantalizing possibility of achieving virtually lattice matched growth if the GaN rotates by 45° with respect to the InSb substrate beneath it, as occurs for Al on GaAs (001).[9] Rotational epitaxy has also been reported for CdTe growth on GaAs (001), where the CdTe is (111) or (100) oriented as a result of different growth conditions and partly stemming from the interfacial chemical interactions.[10,11] Of course, complete success in the InSb system would not be expected to simultaneously solve the substrate problem for GaN due to the low melting point (~500 °C) of InSb.

II. EXPERIMENTAL DETAILS

All XPS experiments were carried out at Cardiff in a home-made molecular beam epitaxy (MBE) style growth chamber joined, under ultrahigh vacuum (UHV), to a conventional VG ESCA-LAB system. In addition to two Knudsen cells on the source flange, an Oxford Applied Research CARS25 radio frequency (rf) plasma source (aperture 25×0.5 mm holes) is attached for the production of atomic nitrogen or hydrogen. The chamber also has the capability of performing reflection high-energy electron diffraction (RHEED) analysis. A more detailed description of the chamber and its capabilities can be found in Ref. 8.

Solvent rinsed InSb samples were first cleaned under an atomic hydrogen flux for 15 min at 275 °C using a forward rf power of 400 W and a chamber pressure of (5.0±0.5) $\times 10^{-5}$ mbar. An optical emission detector (OED), based on an optical fiber fed to a silicon photodiode, read 1.0±0.1 V. After this exposure the RHEED showed an In rich (4×2) reconstruction and XPS scans of the O(*KVV*) and C 1*s* regions showed the absence of carbon and oxygen.

[a]Electronic mail: Haworthl@Cardifif.ac.uk

During nitridation samples were maintained at 275 °C and exposed to atomic nitrogen, produced in a plasma of forward rf power 450 W and nitrogen flow rate 1.0 sccm, corresponding to a chamber pressure of $(5.0\pm0.5)\times10^{-5}$ mbar with an OED value of 3.2 ± 0.1 V. Nitridation was performed sequentially for times from 30 s to 3.5 h and samples were scanned using XPS between exposures. XPS measurements were taken at normal emission using the Mg anode (photon energy 1253.6 eV) of the Sb 3*d* and 4*d* and In 3*d* and 4*d* photoelectron peaks with a 5 eV pass energy and the N 1*s* peak with a 20 eV pass energy. The 3*d* peaks were chosen for further study as they are more surface sensitive than the 4*d* peaks, due to their smaller electron escape depth, and are easier to fit, due to their larger spin orbit splitting. The Sb $3d_{3/2}$ peaks were used for fitting in preference to the Sb $3d_{5/2}$ peak, due to the latter's proximity to the O 1*s* peak position. Sb deposition on the cleaned InSb substrate was carried out at room temperature using a cell temperature of 440 °C corresponding to a deposition rate of 0.5 Å s^{-1}.

Raman experiments were carried out at Chemnitz in a specially designed UHV chamber which allows light scattered from the sample to be collected through a view port and focused onto the entrance slit of a Dilor XY spectrometer with multichannel detection. A more detailed description of this system and its capabilities may be found in Ref. 12. The atomic nitrogen and hydrogen were provided by an rf plasma source (SVT RF 4.5) operated at a power of 450 W. For nitridation the N_2 flux was 2.5 sccm leading to a chamber pressure of 5×10^{-4} mbar during nitridation. Resonant Raman scattering was used as this leads to a massive enhancement of the scattering cross section and so enhances thin layer detection when the excitation energy is chosen in the vicinity of the fundamental band gap or another critical point in the band structure of the material. Therefore, for the nitridation of InSb the 520.8 nm (2.38 eV) emission line of a Kr^+ ion laser was used for excitation with a power of 35 mW to obtain resonance conditions for the InSb $E_1+\Delta_1$ gap, while for the Sb layer deposition and nitridation experiments the 514.5 nm (2.41 eV) emission line of an Ar^+ ion laser with a power of 40 mW was used to maximize resonance conditions for elemental Sb electron interband transition.

Raman spectra were recorded in the backscattering geometry with the polarization of the incident light parallel to the [110] direction of the (001) oriented substrate. For the InSb nitridation experiment the scattered light was also parallel polarized in the [110] direction. In this configuration scattering by the longitudinal optical (LO) phonon of InSb is allowed according to the symmetry selection rules, whereas scattering by the transverse optical (TO) phonon is forbidden. For the Sb layer deposition, at a rate of ≤0.009 Å s^{-1} for 4 h, and subsequent nitridation the spectra were recorded with the scattered light detected unpolarized to maximize the signal. However, detection of light with parallel polarization is strongly favored by the monochromator.

XPS can be used to reveal both the atomic content of a sample and information about chemical states by examining chemical shifts in the binding energies of the photoelectron peaks. For the In and Sb 3*d* photoelectron peaks the escape depths (the distance at which only 1/*e* of the initial electrons remain unscattered inelastically) are ~19 and 17 Å,[13] respectively, making it a relatively surface sensitive technique.

RRS on the other hand measures the scattered light intensity giving the eigenfrequencies of the elementary excitations of the crystal structure. Therefore, it gives information as to the identity of crystalline materials, their phase, orientation and crystalline perfection and stresses within the crystal lattice by evaluation of the phonon frequencies, half-widths, line shapes and intensities in the Raman spectrum. However, if a phase is not crystalline in nature or has a resonant frequency far different from that of the exciting line, its phonon features will be very broad and/or of relatively low intensity making it difficult to detect. The information depth from the exciting radiation is much greater here (~9 nm) than for XPS due to the combination of the laser lines and materials used.[14]

In general, the XPS highlights the near surface reaction, showing overlayer formation and attenuation of subsurface signals. Whereas any overlayer formed is virtually transparent with Raman spectroscopy, unless its modes are resonant with the laser line used, this technique gives the subsurface story and gives more information about changes in or near to the substrate.

III. RESULTS

A. Nitridation of InSb: "XPS studies"

1. RHEED results

The (4×2) reconstruction observed from the InSb surface after hydrogen cleaning converts to unreconstructed bulk stripes on lighting the nitrogen plasma. During the subsequent nitridation the bulk features are modulated to spots of decreasing intensity, while the diffuse background around them increases in intensity. This is indicative of disorder in the surface layer but also implies the preservation of bulk spacing. Most of these changes in pattern occur within the first 2 min of nitridation, after which only one bulk spot in a high intensity background was visible.

2. XPS results: Sb 3d

Even for the shortest nitridation time of 30 s, a reacted Sb feature can be seen to appear on the higher binding energy side of the InSb feature. As nitridation time increases this grows so that it becomes of approximately equal intensity to the InSb feature after 10 min and then quickly surpasses it.

To better describe the peak shape changes, the Sb $3d_{3/2}$ peaks were curve-fitted using a combination of a linear background and Voigt peak profiles. It was found that the clean InSb feature could be fitted with a single Voigt function of Lorentzian and Gaussian widths of 0.6 ± 0.1 and 0.63 ± 0.05 eV, respectively. With this as a guide, the following scheme was used to fit the reacted Sb feature: (a) following nitridation the Sb $3d_{3/2}$ feature was assumed to be composed of only two Voigt functions; (b) the Lorentzian widths of both contributions were constrained to the clean InSb value; (c) the Gaussian width of the InSb contribution was constrained

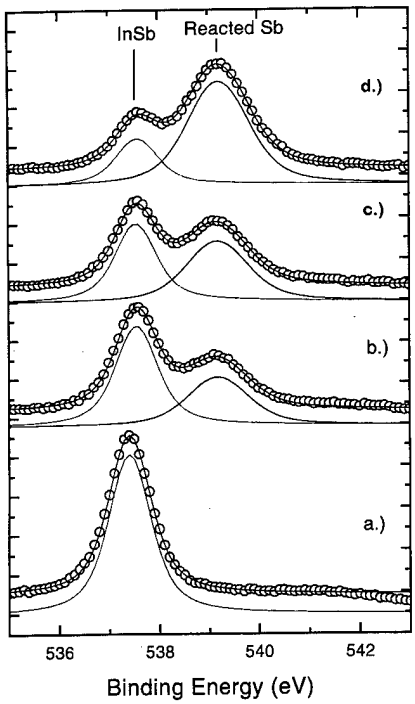


FIG. 1. Fitted XPS scans of the Sb $3d_{3/2}$ spectral region for (a) a clean surface and following nitridation at 275 °C, (b) for 1 min, (c) for 5 min, and (d) for 35 min. The background has been subtracted, the fit to the curve passes through the data points, while the individual components are shifted down for clarity.

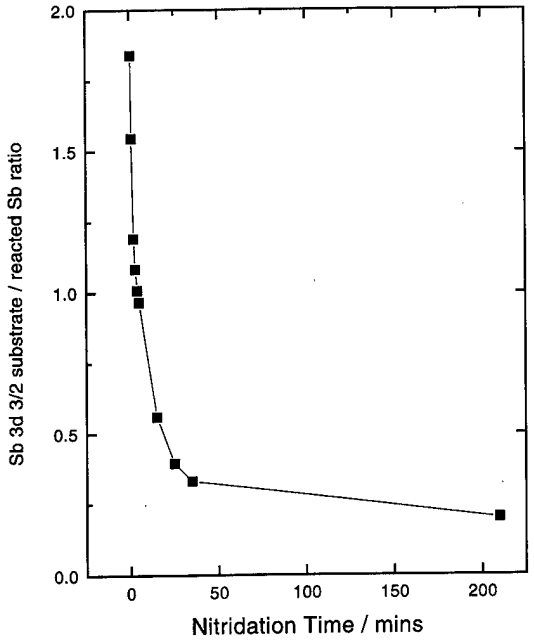


FIG. 2. Ratios of the areas under the fitted contributions (Sb $3d_{3/2}$ substrate/reacted) following background subtraction, as a function of nitridation time at 275 °C.

to the clean value; (d) the Gaussian width of the reacted Sb feature was allowed to broaden giving information on the crystalline quality of the phase. These fits resulted in a reacted Sb feature of Gaussian width 1.38 ± 0.10 eV and chemical shift of 1.65 ± 0.10 eV to the higher binding energy side of the substrate peak (Fig. 1). A graph (Fig. 2) showing the substrate to reacted peak area ratio against nitridation time was then plotted. This shows the reaction rate decreasing with nitridation time.

3. XPS results: In 3d and N 1s

As the time of nitridation increased the In $3d_{3/2}$ feature broadened but no separate reacted peak could be discerned. Although the $3d_{3/2}$ peak could be fitted with a single peak of Lorentzian width 0.6 eV but with increasing Gaussian width for longer nitridation times, to allow for the formation of the chemically shifted InN, it was considered more informative to fit the feature with two peaks: a substrate and a reacted component. In order to achieve this, the area under the substrate peak was fixed in each case with a In/Sb ratio of 0.81 ± 0.05, a value obtained from repeated fitted scans of InSb clean surfaces, from the relevant Sb $3d_{3/2}$ peak. The Gaussian value of the substrate peak was also fixed at 0.52 ± 0.02 eV, again from clean substrate values, and the Lorentzian width was 0.6 eV. The reacted peak could then be fitted, always offset to higher binding energy than the substrate and with a varying Gaussian. The Gaussian width was found to increase with increasing nitridation time, from 0.6 to 1.0 eV, however a large variation in the peak offset of InN, from 0.1 to 0.4 eV, could be obtained for the same goodness of fit.

The N 1*s* peak appeared as a single symmetric peak of increasing intensity with increasing nitridation time, of fixed full width half maximum (FWHM) 1.49 ± 0.05 eV.

4. Identification of the reacted Sb species

In order to investigate the validity of the straight anion substitution described by Eq. (1), elemental Sb was deposited on InSb and then nitrided. XPS was then used to compare the chemical shift of the elemental Sb with respect to the reacted Sb after nitridation in order to elucidate the chemical state of the reacted Sb from nitrided InSb discussed above.

To ascertain the peak shift of elemental Sb relative to the InSb position, Sb was deposited on InSb in several monolayer stages to a total thickness of about 10 nm, calculated from the deposition rate and assuming two-dimensional (2D) growth. As Sb was deposited the Sb $3d_{3/2}$ peak's FWHM initially widened from its value of 1.0 eV for the clean surface by about 0.3 eV after deposition of ~6 Å, then decreased again to the vicinity of its clean value. Accompanying these changes was a final peak shift of about 0.45 ± 0.05 eV to higher binding energy. These results are consistent with peaks from elemental Sb being chemically shifted to higher binding energy. The relatively small chemical shift and consequent peak overlap produces the apparent broadening of the substrate peak as the elemental peak increases in intensity and the gradual shift of the combined peak to higher binding energy. As deposition continues a decrease in the combined peak width is then seen as the substrate is

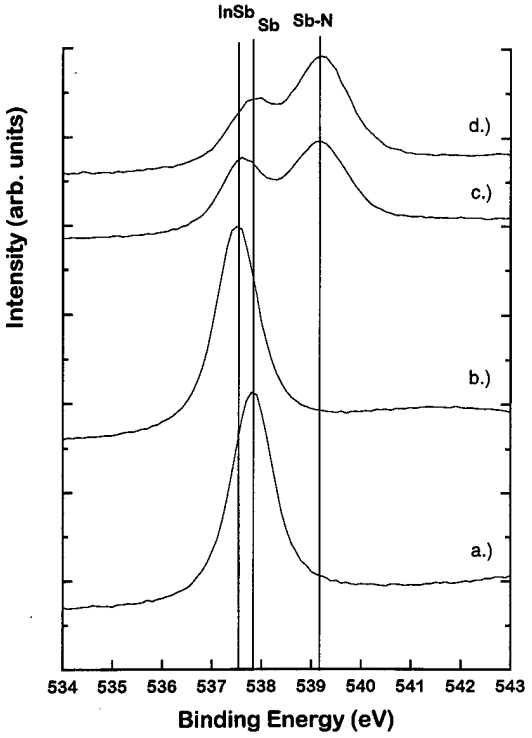


FIG. 3. XPS scans showing the chemical shifts of the peaks in the Sb $3d_{3/2}$ spectral region for (a) a thick elemental Sb layer, (b) clean InSb, (c) following 5 min nitridation of InSb at 275 °C, and (d) following 5 min nitridation of elemental Sb at 275 °C.

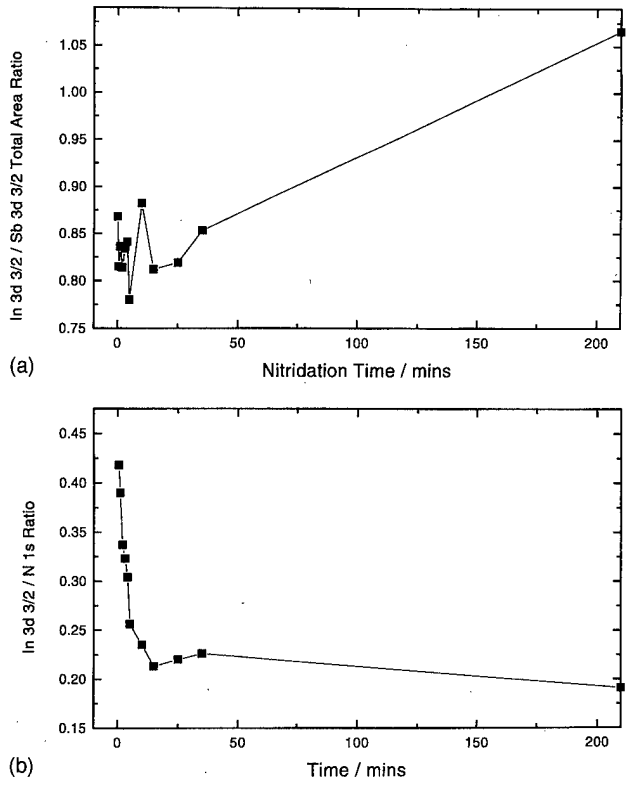


FIG. 4. (a) Ratios of the total areas under curves (In $3d_{3/2}$/N $1s$), following background subtraction, as a function of nitridation time at 275 °C. (b) Ratios of the total areas under curves (In $3d_{3/2}$/Sb $3d_{3/2}$), following background subtraction, as a function of nitridation time at 275 °C.

buried and the peak continues shifting to the elemental peak position. However, on nitriding this Sb layer for 5 min at 275 °C a distinct feature developed 1.35 eV to higher binding energy from the elemental Sb position. This is in agreement with the reacted peak position of nitrided InSb and reveals that a species distinct from elemental Sb is formed on nitridation. The absolute differentiation between the elemental Sb and Sb–N species chemical shift is best seen in Fig. 3 which shows both the clean InSb and the Sb/InSb layer together with the spectra taken after the nitridation of both for 5 min at 275 °C.

As this is the first report of a stable Sb–N species it was of interest to attempt to find out more about its nature. Therefore nitridation of InSb was further examined using resonant Raman scattering.

5. Quantification of overlayer thicknesses

The large amount of sample manipulation from the growth chamber to the ESCALAB analysis chamber, required in the process of performing the nitridation experiments, inevitably causes uncertainties in absolute intensity measurements. Therefore, it was considered much more reliable to compare relative intensities of different peaks obtained under identical conditions without sample manipulation. To avoid errors associated with curve fitting especially in the In $3d$ and N $1s$ cases, ratios of the total areas under the curves following background subtraction as a function of nitridation time have been considered. However it should be noted that the In $3d$ and N $1s$ peaks contain two contributions and the Sb $3d$ peak has possibly three contributions. It was found that the In $3d_{3/2}$/Sb $3d_{3/2}$ ratio remains constant for nitridation times of up to 35 min, at a value of around that of the clean substrate but increases slightly for the 3.5 h ratio, showing a small loss of near-surface Sb [Fig. 4(a)]. The In $3d_{3/2}$/N $1s$ ratio with nitridation time shows a decrease in the reaction rate so the reaction slows almost to a halt after ~15 min [Fig. 4(b)].

However, to actually quantify the extent of nitridation it is necessary to calculate the areas under specific peaks, which for XPS may be obtained using

$$I = K \int_{t_1}^{t_2} \exp\left(\frac{-l}{\lambda}\right) dt, \tag{2}$$

where I is the intensity of the electrons emitted from the material at a certain kinetic energy, t_1 and t_2 are the thickness limits of the layer, l is the distance below the surface (for normal emission), λ is the electron escape depth for the particular materials and electron kinetic energy considered, and K is a constant taking into account: (a) the number density of atoms in the material, (b) sample area, (c) differential photoionization cross section, (d) x-ray flux and detector efficiency.[15] By considering the intensity ratios of contributions from the same element and core level (see, for example,

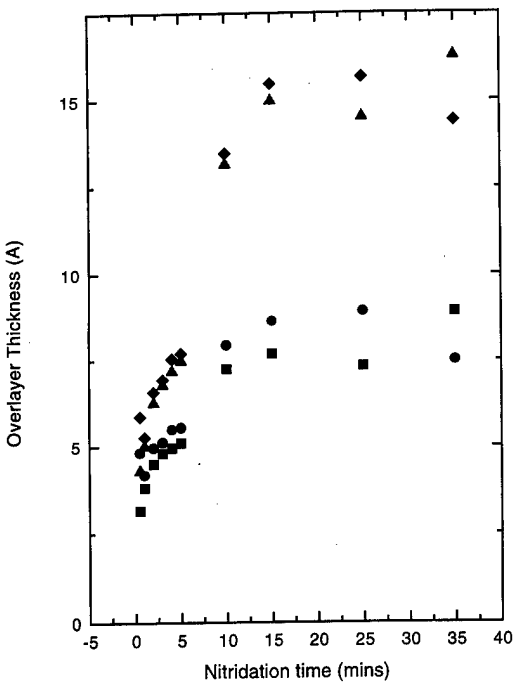


FIG. 5. Estimated film thickness as a function of nitridation time at 275 °C. Calculated assuming the overlayer is made of two distinct layer contributions: Sb–N layer thickness (■) calculated assuming Sb–N is at the surface and (●) assuming it is at the buried interface, InN layer thickness (▲) calculated assuming InN is at the surface and (◆) assuming it is at the buried substrate interface.

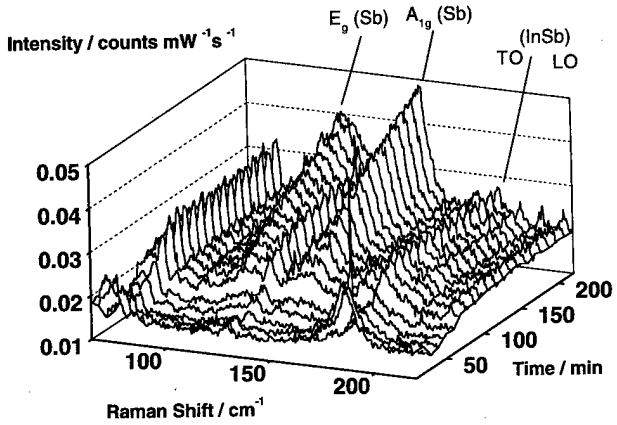


FIG. 6. Evolution of Raman spectra during nitridation of InSb (001) for 3.5 h at 275 °C. (λ_{ex}=520.8 nm P=35 mW) Incident and scattered polarization parallel to [110] direction of the substrate.

InSb and Sb–N contributions to the Sb 3*d* signal) Eq. (2) is much simplified and the thickness of the overlayer can be calculated.

Using the curve fits that assume the Sb overlayer contains just Sb–N and the In overlayer is entirely InN, the substrate to reacted component ratios can then be used to calculate overlayer thicknesses. This is complicated by the fact that there are then two contributions to the overlayer so its composition is unclear. In order to calculate the number densities required it was assumed that the InN was in the cubic form and the number density for the Sb–N species was then calculated by reducing the InSb value by the same percentage as it reduced to form InN. Electron escape depths of 19.1 and 17.1 Å were used for the In $3d_{3/2}$ and Sb $3d_{3/2}$ contributions, respectively.[13] Consideration of uncertainties in the parameters in the equation leads to an estimated error of ±25% inherent in the equation assuming the layers are uniform and abrupt and consist of only the species considered. The calculations were then performed twice, to account for the unknown physical location of the Sb–N with respect to the InN, for the extreme cases of Sb–N being located on the surface or at the substrate overlayer interface (Fig. 5).

Figure 5 shows that the InN and Sb–N layer thicknesses appear to stop increasing after about 15 min nitridation at a total overlayer thickness of ~23 Å, the InN and Sb–N thickness are both then between 8 and 15 Å. It can also be seen that layer thicknesses of a few monolayers are quickly formed even after just 30 s at this temperature and the substantial quantity of Sb–N formed, especially if amorphous, could easily disrupt subsequent growth performed at this temperature. Measurements on a subsequent 3.5 h nitridation at 275 °C gave a total overlayer thickness of ~55 Å implying that although the reaction rate decreases with time it does not stop completely.

Atomic force microscopy (AFM) images taken after 5 min and 3.5 h nitridation show no significant increase in roughness in comparison to the clean substrate surface (mean roughness 0.6±0.2 nm), adding credence to the assumption of uniform layer formation.

B. Nitridation of InSb: "Raman scattering studies"

The resonance of the exciting laser line with features in the band structure of both hexagonal elemental Sb and cubic InSb provides them with a large scattering cross section enabling the order of monolayers of each to be detected.[16] InN conversely has a low scattering cross section at the laser line used and has proved notoriously hard to detect[17] due to the difficulty of producing a sufficiently thick crystalline sample. As nothing is known about the Sb–N crystal structure, its cross section with the laser line used is not known, nor could a suitable line be chosen specifically to achieve resonance with it.

Figure 6 shows a 3D plot of the evolution of the Raman spectra during the continuous measurement of nitridation of InSb (001) for 3.5 h at 275 °C. Each curve shows the sum of the data over the preceding 10 min. The first curve shows the starting surface displaying the InSb LO phonon mode at 186 cm^{-1}, shifted by ~5 cm^{-1} due to the elevated temperature. As nitridation progresses this peak gets broader with a shoulder appearing to the lower shift at 177 cm^{-1}; this is attributed to the forbidden TO mode of InSb being observed as the surface roughens. Further evidence of surface roughening may be seen in the increasing intensity of the peaks attributed to scattered light from the nitrogen plasma at ~70 and ~130 cm^{-1}. The peaks at 112 and 146 cm^{-1} are attributed to the E_g and A_{1g} modes of elemental Sb, respectively, and these are seen to increase throughout the nitridation. Their positions and relatively small FWHMs show that the Sb is

crystalline. In addition their sustained existence at this temperature, which is close to the reported total desorption temperature for elemental Sb of 280 °C,[14] provides indirect evidence of the formation of a protective InN/Sb–N overlayer.

Vibration modes from InN would be expected at ~500–600 cm^{-1},[17] depending on its crystalline structure but is not seen here probably due to the combination of: a lack of resonance with the line used, the presence of plasma lines, the small amount of InN present and the fact that it may be of a disordered nature. Similar reasons are likely to account for the lack of features attributable to Sb–N.

Even experiments performed after nitridation on the sample at room temperature using an excitation wavelength of 647.1 nm, for which the corresponding photon energy (1.92 eV) is close to the band gap of InN and should therefore lead to resonant enhancement, did not reveal the presence of an InN layer.

C. Annealing study of Sb and Sb–N

One property of the new Sb–N compound mentioned previously, and relatively easy to examine, is its volatility. In addition to distinguishing between the species, it is important in the context of growth, where the presence of large quantities of highly disordered Sb–N would disrupt subsequent epitaxial growth, shown to be a problem after nitridation at this temperature where quick disruption of surface structure was shown by the RHEED data.

XPS annealing studies were performed on both the 10 nm elemental Sb layer and the 10 nm elemental Sb layer that was subsequently nitrided. On annealing the elemental Sb layer for 30 min at 350 °C the elemental Sb $3d_{3/2}$ peak was found to shift back to the position of the clean InSb, showing almost complete evaporation of a 10 nm layer. However, on annealing the nitrided Sb layer at 350 °C for 30 min, the elemental Sb position remained unshifted and the Sb–N feature also remained unaltered. The elemental Sb must therefore be protected by the reacted layer enabling it to remain unevaporated at higher temperatures. However, on annealing at 400 °C for 30 min the lower binding energy peak shifted back to the InSb peak position and increased in intensity, showing the evaporation of the elemental Sb, while the Sb–N peak remained unshifted but decreased in intensity. This process was continued during an anneal at 450 °C for 30 min, but even after this anneal some Sb–N remained. Therefore the Sb–N protects the elemental Sb to some extent but the elemental component is still more volatile than the Sb–N when both coexist. The relative stability of the Sb–N at temperatures approaching the melting point of the substrate (for example, ~527 °C) is a concern when considering suitable GaN/InSb (001) growth temperatures.

These results may be compared to a similar experiment on InSb in which InSb was nitrided for 35 min at 275 °C, then annealed at 275 and 350 °C, each for 30 min. This resulted in an unchanged Sb $3d_{3/2}$ peak shape. A small reduction of the relative Sb–N intensity was seen after a 30 min anneal at 400 °C, but a significant reduction in relative Sb–N intensity was not observed until after a 30 min anneal at 450 °C. This shows that the Sb–N formed by nitridation of the substrate and an elemental Sb layer have similar volatility and chemical shift.

A Raman scattering study of annealing was also performed on a nitrided Sb layer. A thick Sb layer (of less than critical thickness ~13.5 nm, the thickness at room temperature where elemental Sb has been shown to undergo spontaneous crystallization on ZnSe[18]) was deposited on hydrogen-cleaned InSb. RRS showed that as it was heated to 275 °C the layer crystallized. It was then nitrided at 275 °C for 3.5 h; during which the intensity of both Sb modes increased while that of the InSb remained constant. This increase was probably due to interference enhanced Raman scattering in the elemental Sb layer upon overlayer formation.[14] On annealing the nitrided layer at 350 °C for 60 min no change was visible in any of the peaks but on increasing the temperature to 450 °C and annealing for 90 min the Sb modes were seen to decrease in intensity rapidly to almost zero, while the InSb peak increased in intensity slightly. This, in broad agreement with the XPS results, again indicates a more stable Sb–N species is formed by the nitridation which in turn stabilizes the elemental Sb coexisting with it.

IV. DISCUSSION/CONCLUSIONS

The results of the nitridation of InSb at this fairly low temperature in many ways appear to mirror the more well researched process of nitriding GaAs. The reaction rate of GaAs (001) at 300 °C like InSb (001) at 275 °C was seen to decrease with increasing nitridation time.[8,19] AFM images of both nitrided surfaces showed roughness comparable to the clean substrate. Overlayer thicknesses calculated from the XPS results revealed combined GaN/As thicknesses ~1/2 that of the combined InN/Sb–N thickness after any nitridation time. This difference can be explained as being largely due to the much lower volatility of the reacted Sb (Sb–N) compared with that of the reacted As (As–N). In addition the reaction rate may be slightly higher for InSb. However, details of neither the Sb–N nor As–N species stoichiometries are yet known. Further work must be carried out, for example, to find the InN In 3*d*/N 1*s* ratio by growing a thick InN layer in order to subsequently deduce the Sb–N Sb 3*d*/N 1*s* ratio.

Although the XPS chemical shift results were conclusive that the reacted Sb reported after nitridation was mainly due to an Sb–N species, RRS only detected a crystalline elemental Sb component. Due to the stringent conditions for detecting low levels of material by RRS these results are not contradictory. However, in order to address this issue, the Sb $3d_{3/2}$ peaks from the XPS spectra were refitted to ascertain how much elemental Sb could be accommodated without reducing the goodness of fit. The Lorentzian width (0.6 eV) and Gaussian widths (0.63 eV) of the elemental and substrate contributions were fixed and the elemental Sb contribution was given a peak shift of 0.45 eV to higher binding energy, values obtained from fits of the clean InSb (001) and thick Sb layers. The peak position and Gaussian width of the Sb–N contribution were allowed to vary to allow insertion of

the extra peak. It was found that an elemental Sb peak of maximum area ~1/25–1/6 that of the Sb–N peak could always be included in the fit. This ratio is somewhat dependent on peak shape, so that it is easier to fit more elemental Sb when the InSb and Sb–N contributions are of comparable sizes. Therefore it is likely that the maximum elemental Sb contribution is at the lower end of the range given above. It should also be noted that direct comparison with the Raman scattering results is difficult because of the different experimental arrangements and time scales of the experiments. Hence, the first 40 min of nitridation in RRS showed the presence of only a small amount of elemental Sb and this experiment continued over a 3 h period, whereas the XPS was conducted mainly over 0–40 min. However, the Raman nitridation results did indirectly show the presence of an Sb–N species through their action as a protective capping layer preventing evaporation of the thick elemental Sb overlayer component.

These nitridation results have many implications for proposed rotated GaN growth on InSb. It has been found that nitridation occurs to some extent even when the shutter to the plasma source is closed[8] and also that plasma sources require a time to achieve a stable flux of active species during striking and tuning,[20] so a certain amount of nitridation is inevitable and a knowledge of the products of nitridation will aid in refining the growth process. The low maximum congruent temperature of InSb (~325 °C) limits the initial growth temperature on an untreated surface. However, epitaxial growth performed after even a few minutes nitridation at a temperature just below this may be hindered by the stable, thick but disordered layers of InN and Sb–N formed on the surface. This problem might be resolved by annealing to desorb the Sb component of the overlayer but leaving a more ordered InN to protect the substrate stoichiometry at the elevated temperature before growth starts. It may be the case that as little nitridation as possible is the key to successful epitaxial GaN growth and a two stage growth, where the first few monolayers are grown at a temperature just below the noncongruent temperature before continuing growth at an elevated temperature is the answer. However, in the case of CdTe/GaAs[10,11] it has been shown that some interface reactivity will be needed to achieve rotational epitaxy. The issue is clearly complicated and still must be addressed.

In summary, nitridation of InSb at 275 °C has been investigated using XPS and Raman scattering for times up to 3.5 h. The XPS studies reveal the reacted Sb to be mainly chemically shifted by 1.65 ± 0.1 eV to higher binding energy from the substrate peak. By deposition and nitridation of a thick elemental Sb layer this peak was able to be differentiated from elemental Sb, which was found to be shifted by only 0.45 ± 0.10 eV in the same direction, showing that an Sb–N species is formed on nitridation. The Raman scattering experiments, however, revealed a crystalline elemental Sb component to be also present during nitridation, which is not immediately obvious from the XPS data. The lack of scattering from InN and Sb–N in the RRS spectra suggests these species are disordered. Annealing studies show that the Sb–N species is significantly less volatile than elemental Sb and only becomes unstable at temperatures approaching the substrate melting point. The formation of relatively thick, and probably amorphous, layers during nitridation at these low temperatures has implications for the design of a suitable growth process to aim for rotational epitaxial GaN growth on InSb.

[1]S. Nakamura, *Proceedings of the International Conference on the Physics of Semiconductors* (World Scientific, Singapore, 1996), Vol. 1, p. 11.
[2]M. A. Khan, A. R. Bhattarai, J. N. Kuznia, and D. T. Olsen, Appl. Phys. Lett. **63**, 1214 (1993).
[3]S. Strite and H. Morkoc, J. Vac. Sci. Technol. B **10**, 1237 (1992).
[4]H. Yang, O. Brandt, and K. Ploog, J. Electron. Mater. **25**, 787 (1996).
[5]A. Kikuchi, H. Hoshi, and K. Kishino, Jpn. J. Appl. Phys., Part 1 **33**, 688 (1994).
[6]L. A. Delouise, J. Vac. Sci. Technol. A **10**, 1637 (1992).
[7]Y. Bu and M. C. Yin, Surf. Sci. **317**, 152 (1994).
[8]P. Hill, D. I. Westwood, L. Haworth, J. Lu, and J. E. Macdonald, J. Vac. Sci. Technol. B **15**, 1133 (1997).
[9]M. A. Al-Khafaji, D. Cherns, C. J. Rossouw, and D. A. Woolf, Philos. Mag. B **65**, 385 (1992).
[10]S. Tatarenko, J. Cibert, Y. Gobil, G. Feuillet, K. Saminadayar, A. C. Chami, and E. Ligeon, Appl. Surf. Sci. **41/42**, 470 (1989).
[11]A. Bourret, P. Fuoss, G. Feuillet, and S. Tatarenko, Phys. Rev. Lett. **70**, 311 (1993).
[12]V. Wagner, D. Drews, N. Esser, D. R. T. Zahn, and W. Richter, J. Appl. Phys. **75**, 7330 (1994).
[13]H. Grant and W. Monch, Surf. Sci. **105**, 217 (1981).
[14]D. Drews, J. Sahm, W. Richter, and D. R. T. Zahn, J. Appl. Phys. **78**, 4060 (1995).
[15]T. E. Madey, J. T. Yates, and L. E. Erickson, Chem. Phys. Lett. **19**, 487 (1973).
[16]W. Richter, in *Springer Tracts of Modern Physics* (Springer, Berlin, 1976).
[17]H. Kwon, Y. Lee, O. Miki, H. Yamano, and A. Yoshida, Appl. Phys. Lett. **69**, 937 (1996).
[18]A. Schneider, D. Drews, D. R. T. Zahn, D. Wolffframm, and D. A. Evans, J. Cryst. Growth **159**, 732 (1996).
[19]P. Hill, D.I. Westwood, L. Haworth, J. Lu, and J.E. Macdonald, Appl. Surf. Sci. **123/124**, 126 (1998).
[20]M. Sato, Jpn. J. Appl. Phys., Part 1 **34**, 1080 (1995).

Influence of growth conditions, inversion domains, and atomic hydrogen on growth of (0001) GaN by molecular beam epitaxy

T. H. Myers[a] and L. S. Hirsch
Department of Physics, West Virginia University, Morgantown, West Virginia 26506

L. T. Romano[b]
Xerox Palo Alto Research Center, Palo Alto, California 94304

M. R. Richards-Babb
Department of Chemistry, West Virginia University, Morgantown, West Virginia 26506

(Received 21 January 1998; accepted 11 May 1998)

Growth of GaN by rf-plasma molecular beam epitaxy leads to different surface morphologies for nitrogen-stable growth versus gallium-stable growth. Nitrogen-stable growth produces a granular surface morphology with many samples having a significant density of pyramidal hillocks. In contrast, gallium-stable growth results in a flat surface morphology. The hillocks were directly linked to the presence of inversion domains which originated in the nucleation layer. Nitrogen-stable growth and growth under atomic hydrogen enhanced the growth rate of inversion domains with respect to the surrounding matrix, while growth under Ga-stable conditions resulted in a more nearly equal growth rate. Evidence is presented suggesting that hydrogen may stabilize the surface of growing GaN. © *1998 American Vacuum Society.* [S0734-211X(98)05504-8]

I. INTRODUCTION

The potential applications of blue and ultraviolet optoelectronic devices based on GaN have been recognized for many years.[1] Recent advances in epitaxial GaN growth by metal organic chemical vapor deposition (MOCVD) have lead to the first demonstration of a blue laser based on GaN.[2] Rapid progress in this direction is also being accomplished by molecular beam epitaxy (MBE) growth using active nitrogen species.[3–5] Several issues remain to be resolved. As noted by Tarsa *et al.*,[6] it is becoming clear that certain aspects of growth phenomenon relating to GaN are universal in nature, particularly the predominant effects of the ratio of Ga to nitrogen during growth. For example, growth of GaN under Ga-stable conditions has resulted in improved structural, electrical, and optical properties with smoother surface morphologies when compared to N-stable growth.[6–10]

Heteroepitaxial GaN layers grown by any technique contain a high dislocation density as well as other classes of defects such as nanopipes, voids, and inversion domain boundaries (IDBs). Inversion domains (IDs), consisting of regions of GaN with the opposite polarity to the primary matrix, have been observed to some extent in GaN grown by all techniques.[11–17] However, there is some indication that IDBs are both more common and more stable during growth of GaN by molecular beam epitaxy.[7,15] We present the results of a study of IDBs in GaN grown by rf-plasma-assisted MBE detailing how different growth conditions determine the effect IDs have on surface morphology.

One of the obvious differences between MOCVD and MBE growth of GaN is the presence of hydrogen, primarily as a component of the molecule supplying the nitrogen. Most studies have either ignored the effect of hydrogen on the growth, or have only considered its effects in compensating *p*-type dopants such as Mg.[18] Hydrogen may, however, be altering the growth kinetics as well. We also report the effects of the presence of atomic hydrogen on the growth kinetics of GaN. We present evidence that atomic hydrogen can have a significant effect on the growth kinetics of GaN when the growth is limited by the amount of active nitrogen present.

II. EXPERIMENT

The GaN layers for this study were grown at West Virginia University (WVU) by MBE in a custom system. A standard MBE source (EPI-40M) provided the Ga flux. A cryogenically cooled rf-plasma source (Oxford Applied Research CARS-25) operating at either 500 or 600 W was used to produce the active nitrogen flux. All layers reported here were grown with a nitrogen flow rate of 6 sccm, controlled by a mass flow controller, resulting in a system background pressure of 6×10^{-5} Torr during growth.

Atomic hydrogen was produced using a thermal cracker (EPI). Typically, 1×10^{-6} Torr beam equivalent pressure (BEP) of hydrogen was passed through the thermal source operating at 9.5 A, although the hydrogen flux was varied for several samples. Literature supplied with the source indicated a dissociation efficiency of about 10% for this operating condition. Therefore, the sample was exposed to both atomic and molecular hydrogen during growth.

Determination of substrate temperature and growth rate were important in this study. A spring-loaded type K thermocouple was in intimate contact with the back of the molybdenum sample block with the substrate mounted with indium. Prior to this sequence of growths, an additional thermocouple was attached to the front of a mounted sub-

[a]Electronic mail: tmyers@wvu.edu
[b]Electronic mail: romano@parc.xerox.com

strate and a calibration curve determined between the front and back thermocouples. This calibration was routinely checked by using the melting points of various metals, oxide desorption from GaAs, and the use of an optical pyrometer for the higher temperatures. Temperature determination was reproducible to ±5 °C.

Relative measurements of sample growth rates were performed by measuring interference effects in reflectance from the growing layer using 680 nm light from a semiconductor laser. These measurements were converted to an absolute growth rate by using total sample thickness and the growth time. The thickness at the sample's center was determined from interference fringes in optical transmittance measured using a Cary-14 spectrophotometer. This determination of total thickness was found to agree with values determined by transmission electron diffraction (TEM).

All samples were grown on *c*-plane sapphire substrates (Union Carbide Crystal Products). Prior to growth, the substrates were degreased and etched in a phosphoric/sulfuric (1:3) acid mixture heated to 80 °C. Based on our earlier study,[9] buffer layers were grown by heating the substrate to 730 °C under an atomic hydrogen flux for 20 min and then cooling to 630 °C for the growth of a 200 Å thick GaN buffer layer under a Ga flux of 5.0×10^{-7} Torr (BEP) with a 6 sccm nitrogen flow at 600 W. Buffer layer growth was initiated by simultaneous exposure to the Ga and N flux. This nucleation layer was then annealed at 730 °C for 20 min under nitrogen flux, cooled to the growth temperature, and growth was resumed. These conditions represent buffer layer growth under highly Ga-stable conditions. However, after the 730 °C anneal, examination of the buffer layers by atomic force microscopy (AFM) indicated continuous films with no evidence of Ga condensation. This procedure led exclusively to the nucleation and growth of (000$\bar{1}$)-oriented (or N-terminated) GaN as determined using the polarity-indicating etch described by Seelmann-Eggebert *et al.*[19] As discussed in a later section, this determination is consistent with other observations.

Atomic force microscopy and TEM were used to study surface morphology and its relation to microstructure in a series of GaN layers grown by MBE on (0001) sapphire. Detailed descriptions of other layer characterizations are given elsewhere.[7,20,21] Cross-section TEM (XTEM) studies were performed with a JEOL 3010 microscope operated at 300 kV on samples that were prepared by polishing and then ion milling to electron transparency. TEM images were taken under various diffraction conditions including multiple dark-field imaging with $g=+/-(0002)$ along either the $\langle11\bar{2}0\rangle$ or $\langle10\bar{1}0\rangle$ axis in order to reveal inversion domains.[11] AFM was performed in air using a Digital Instruments Nanoscope II.

III. GROWTH CONDITIONS AND SURFACE MORPHOLOGY

As reported by our group[7] as well as others,[3,6,8,9] growth under Ga-stabilized conditions versus N-stabilized conditions leads to a drastically different surface morphology. When varying the Ga flux with all other growth conditions fixed, two different regimes are indicated by reflection high-energy electron diffraction (RHEED), as shown in Fig. 1. For a higher Ga flux, a RHEED pattern was observed consisting of well-defined, truncated streaks similar to Fig. 1(a). Such a pattern indicates a surface that is locally smooth on an atomic scale, and is indicative of two-dimensional (2D) growth. Lowering the Ga flux lead yielded a transition to a spot pattern such as shown by Fig. 1(b). This transition to a pattern indicative of three-dimensional (3D) growth implies a surface that is rough on an atomic scale. For a short duration of 3D growth, the transition was reversible and 2D growth could be recovered by increasing the Ga flux. The transition was fairly abrupt, occurring over a small change in Ga flux, about 0.5×10^{-7} Torr BEP.

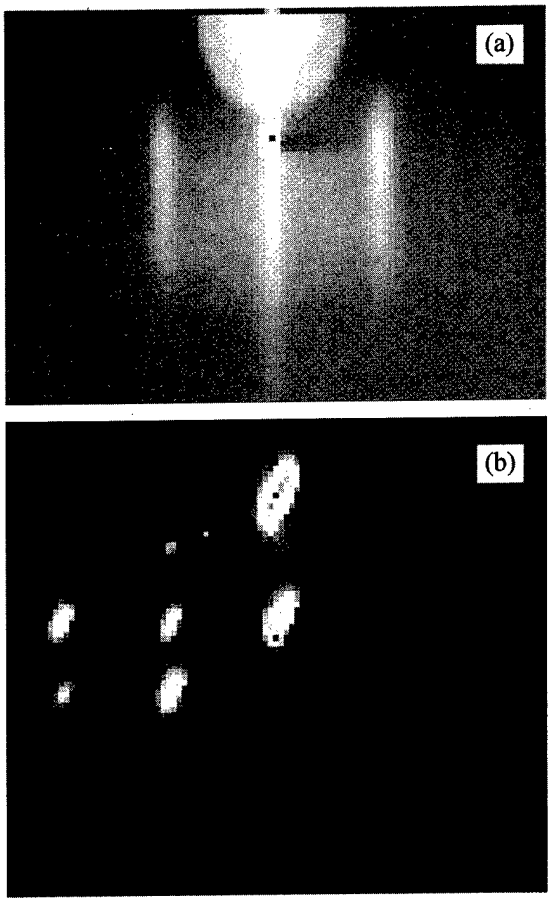


FIG. 1. RHEED along [2110] of GaN grown under (a) Ga-stable conditions and (b) N-stable conditions.

Measurement of the GaN growth rate as a function of Ga flux indicates that the RHEED pattern change can be attributed to a transition from Ga-stable to N-stable growth. In Fig. 2, the growth rate is plotted as a function of Ga beam equivalent pressure. The nitrogen flow rate was held constant at 6 sccm, 600 W. For an active nitrogen flux in excess or equal to the Ga flux, one would expect a linear increase in the growth rate with increasing Ga flux since any excess nitrogen will desorb. When the surface ratio of Ga to active nitrogen is increased beyond unity, then either Ga condensation will occur if the desorption rate for excess Ga is low or

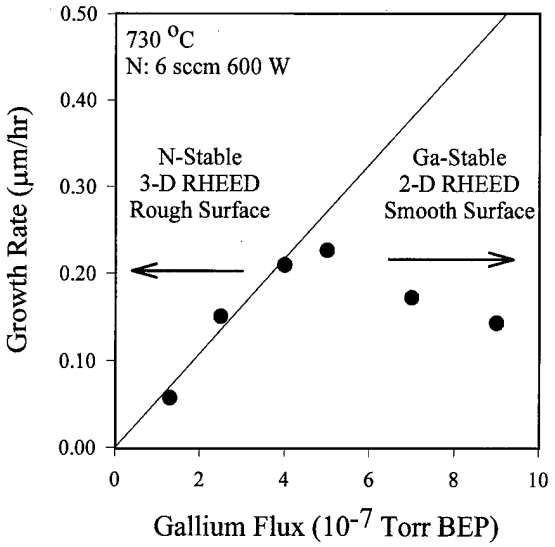


FIG. 2. Growth rate of GaN for various values of Ga flux.

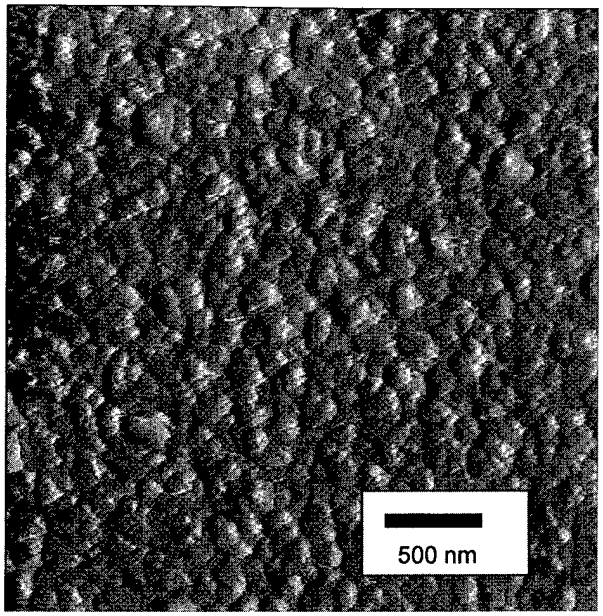


FIG. 3. AFM micrograph of GaN grown under N-stable conditions. Surface features had a rms surface roughness of about 30 nm.

the growth rate will become fairly constant with increasing Ga flux if the excess Ga desorption rate is high. A linear increase in growth rate was observed in going between 1.3 to close to 4.0×10^{-7} Torr BEP Ga. A least-squares fit to these points, including a fixed point of zero growth at zero flux, is represented by the solid line in Fig. 1. The growth rate did not increase significantly above this Ga flux, indicating that the growth was now limited by the amount of active nitrogen available. As indicated in Fig. 1, we actually observed a decreasing growth rate for increasing Ga flux, similar to that reported by Held *et al.*[10] for GaN growth with ammonia-based MBE. This region of constant (or decreasing) growth rate represents Ga-stable growth, whereas the linearly increasing region represents the nitrogen-stable case. A series of experiments where the Ga flux was varied and growth rate measured *in situ* indicated that the 2D-to-3D transition and plateau in growth rate occurred at the same Ga flux, to within 0.5×10^{-7} Torr BEP. For the conditions shown in Fig. 2, this occurred for a Ga flux around 5×10^{-7} Torr BEP.

As we reported earlier, both AFM[7] and x-ray surface scattering[21] measurements indicate a very smooth surface morphology for GaN grown under Ga-stable conditions, with a rms surface roughness less than 1 nm. Features less than 1 nm in height are difficult to resolve with our AFM. In contrast, samples grown under N-stable conditions exhibited a highly textured surface morphology, with rms surface roughness of 10–20 nm. Such samples displayed a coarse surface morphology dominated by grainlike features as indicated by Fig. 3. Smaller Ga-to-N flux ratios appeared to produce a finer-scale morphology. On many samples, a high density of triangular-shaped pyramidal hillocks was also observed. As discussed in Sec. IV, the hillock features were associated with inversion domains.

A recent study by Tarsa *et al.*[6] produced essentially identical results for homoepitaxial growth of GaN on MOCVD-grown GaN using similar growth conditions. Of interest here is that the MOCVD GaN used by Tarsa *et al.* is believed to have a (0001)-oriented (or Ga-terminated) growth surface, whereas ours is (000$\bar{1}$). The similarity in morphology and growth modes indicates that the fundamental mechanisms governing growth kinetics may be independent of orientation. Our results, coupled with that of Tarsa *et al.*, suggest that N-stable conditions result in low surface adatom mobility resulting in "quenched" growth and leading to statistical roughening of the surface.[22] In contrast, Ga-stable conditions appear to promote surface adatom mobility, resulting in a larger surface diffusion length and leading to two-dimensional growth.

IV. INVERSION DOMAINS AND SURFACE MORPHOLOGY

Samples grown either under N-stable conditions or under an atomic hydrogen flux often exhibit a highly textured surface morphology consisting of pyramidal hillocks. Figure 4 is an AFM micrograph illustrative of such a surface completely dominated by the triangular hillocks, while Fig. 5 is from a sample exhibiting isolated hillocks. The hillocks have triangular cross sections that range from 0.2 to 0.5 μm on a side for 1 μm thick samples, and between 150 and 250 nm in height. Increasing the Ga/N ratio (without atomic H) results in smaller size pyramids. These features were not observed for highly Ga-rich conditions.

Figure 6 shows a XTEM image of a hillock from a 1.0 μm thick sample grown under an atomic hydrogen flux with a surface that contained a high density of pyramid-shaped hillocks. The hillocks were ~100 nm high and ~200 nm wide at the base and each contained an ID with a cross section of ~10 nm as shown in Fig. 6. The IDs were found to originate at the film/substrate interface and extend to the film surface with a constant cross-sectional area between 5 and 20 nm. The IDBs were along the {10$\bar{1}$0} planes, similar to IDs

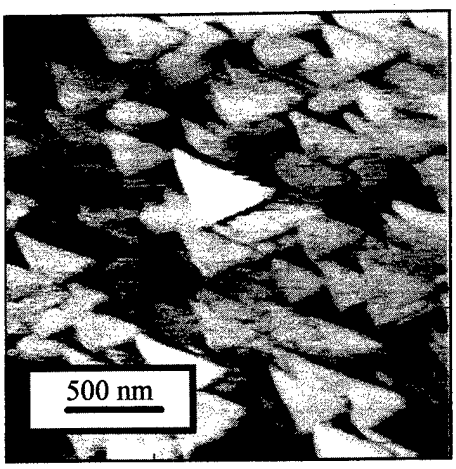


FIG. 4. AFM micrograph of a GaN surface dominated by triangular hillocks. The pyramidal structures were about 100 nm high.

found in other films grown by MBE[7,15] and MOCVD.[12,14] What is amazing is the pronounced effect the small ID has on surface morphology. In terms of sample volume, the ID density in the sample shown in Fig. 4 is less than 5%, yet effects associated with the IDs totally dominate the sample's morphology.

GaN layers grown by MOCVD with a similar surface morphology have been investigated by Daudin *et al.*[12] using ion channeling and convergent beam electron diffraction. They found that the IDs associated with pyramids at the surface of their samples were oriented (0001) (or Ga terminated), while the bulk of the matrix was oriented (000$\bar{1}$) (or N terminated). In contrast, they were also able to grow "flat" samples which were single phase (0001) and contained no IDs. Hillock formation apparently results from the higher growth rate of the ID located at its midpoint. The resultant strain at the boundary may also enhance the growth rate of the opposite phase and dictate the final surface morphology. Perhaps this nonplanar growth region is a source of point defects which degrade the properties of samples grown under N-stable conditions.[8]

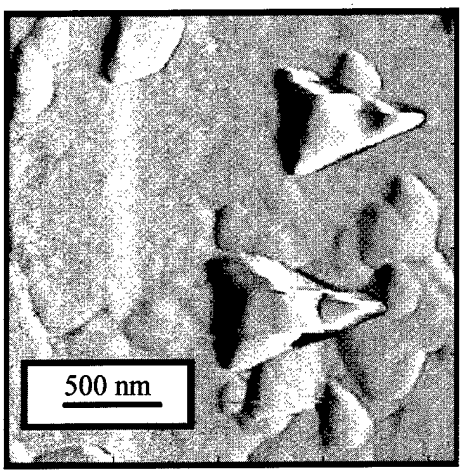


FIG. 5. AFM micrograph of a GaN surface showing isolated hillocks. The hillocks were about 120 nm high. This sample was grown slightly Ga stable, resulting in a smooth (rms roughness about 1 nm) between the hillocks.

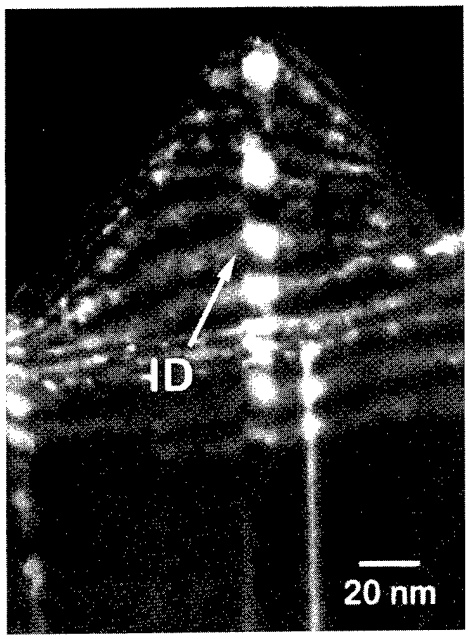


FIG. 6. Dark-field TEM image of a pyramid at the surface taken with $g=(0002)$ near the $(11\bar{2}0)$ zone axis of a film grown under Ga-stable conditions with atomic hydrogen showing the central inversion domain.

We believe our IDs are oriented (0001) similar to the MOCVD results.[12] This is supported by a recent study by Smith *et al.*,[23] who found by scanning tunneling microscope measurements for films grown by rf-plasma MBE using similar techniques, that the polarity of the matrix was N terminated. If our assessment is correct, however, the Ga-terminated surface can have a significantly higher growth rate than a N-terminated surface, in agreement with the speculations of Middleton *et al.*[24] The growth rate differential is also consistent with recent results on MOCVD GaN reported by Liliental-Weber *et al.*,[25] where (000$\bar{1}$)-oriented IDs grow at a slower rate than the (0001)-oriented matrix, leading to oriented "pits" that could be described as inverted hillocks. One implication is that if the bulk matrix is nucleated to be Ga terminated and a N-terminated ID is of small cross section, the possibility exists for overgrowth of the ID under N-stable conditions, removing this defect after growth of a sufficient layer thickness. Thus, the observation that inversion domains are stable during MBE growth of GaN may be due to the relative differential growth rate on (000$\bar{1}$)-oriented material.

Samples grown under excess Ga were found to be free of pyramidal hillocks, even for films containing high densities of IDBs. Again, IDBs were observed to originate at the film/substrate interface and propagate through the growing layer. In sharp contrast to N-stable growth, no significant hillock features were observed at the intersection of the IDB with the surface of the GaN. However, as reported earlier[16] the surface of the ID is found to be 2 nm higher than the surround-

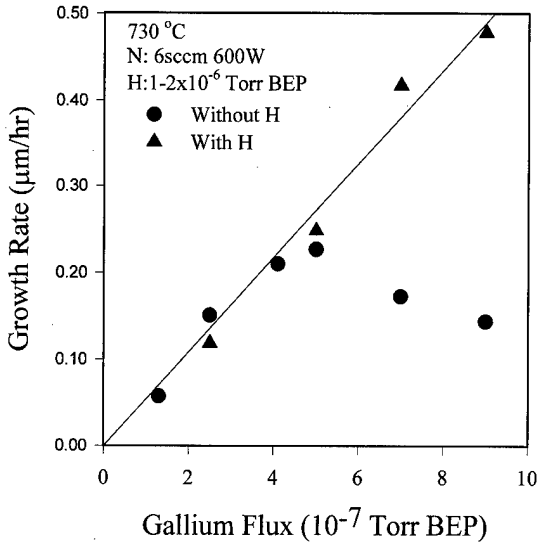


FIG. 7. Effect of atomic hydrogen on the growth rate of GaN for various conditions.

ing matrix, indicating a slight enhancement in the growth rate in the region near the ID. This height difference is significantly less than in the film shown in Fig. 6. Thus, Ga-stable growth appears to suppress the growth-rate differential between the two different polarities of GaN, directly leading to the observed smoother surface morphologies.

The formation of IDBs is not an intrinsic property of MBE growth of GaN, but is apparently related to nucleation conditions. We have grown several samples under N-stable conditions that only exhibit the granular texture indicated by Fig. 3, without the presence of hillocks. The IDs initiate within the buffer layer and are localized laterally with a relatively small cross section, indicating a localized nucleation site. We obtain different densities of IDBs with what we believe are identical nucleation conditions. This could be due to the presence of steps in the sapphire surface as suggested previously,[11] defects in the substrate surface itself from remnant polishing damage, or possibly related to high-energy ions present in the nitrogen flux itself. (With respect to the latter, we have measured a small but finite flux of nitrogen ions with energies >25 eV from our nitrogen source.[26])

V. INFLUENCE OF ATOMIC HYDROGEN

The presence of atomic hydrogen dramatically increases the growth rate of GaN grown under Ga-stable conditions, as shown in Fig. 7. The growth rates indicated by the solid triangles in Fig. 7 were grown under a total (atomic and molecular) hydrogen flux of 1×10^{-6} Torr BEP. The other growth parameters remained the same as for the corresponding filled-circle case. Note that the growth rate was essentially tripled for the most Ga-rich growth. In contrast, growth under nitrogen-stable conditions did not exhibit an enhanced growth rate. A recent study by Daudin and Widmann[12] indicates a similar increase for the growth rate of AlN grown under Al-stable conditions due to the introduction of hydrogen.

The increase in growth rate for the Ga-stable cases was not very sensitive to the overall hydrogen flux. Changing the hydrogen flux from 0.5 to 2.0×10^{-6} Torr BEP gave the same value for the increase in the growth rate. To see if the increased growth rate originated with molecular hydrogen, samples were grown under hydrogen flux with the cracker turned off. The resulting samples exhibited identical growth rates to the GaN grown without hydrogen, indicating that molecular hydrogen is not significantly affecting the growth kinetics. Also, to see if the atomic hydrogen was possibly forming active species with molecular nitrogen, an attempt was made to grow under an atomic hydrogen flux with the rf power turned off on the nitrogen source. The resulting GaN growth rate, if nonzero, was too small to be detected.

The increase in growth rate for Ga-stable conditions is apparently related to the presence of atomic hydrogen. The growth rate enhancement itself is not sensitive to the actual atomic hydrogen overpressure within the limits we investigated, but depends only on its presence. It is well known that hydrogen easily bonds to the surface of other semiconductor systems, such as silicon[27] or diamond.[28] Also, there is evidence that atomic hydrogen alters the growth kinetics in GaAs.[29] One possibility is that the atomic hydrogen becomes loosely bonded to the growing GaN surface. Nitrogen atoms adsorbed on the surface are then attracted by this hydrogen layer, resulting in an increased nitrogen residence time. The longer residence time increases the probability that a Ga atom will diffuse to within an interaction distance of the nitrogen, and thus enhance the growth rate of GaN. Thus, the atomic hydrogen could be increasing the effective active nitrogen concentration. Of interest, the surface morphology for samples grown under atomic hydrogen more nearly resembled N-stable growth. We also observed the same enhanced ID growth rate as for N-stable growth, resulting in the pyramidal structures. The morphology of the background was generally granular, similar to Fig. 3. Surface morphology became smoother only for the most Ga-stable-plus-

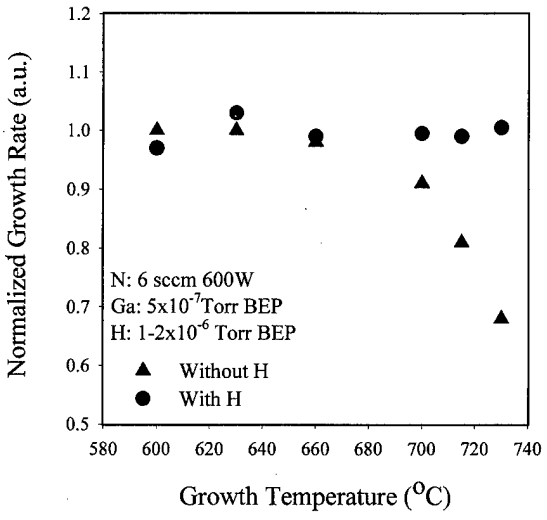


FIG. 8. Influence of atomic hydrogen on the temperature dependence of the growth of GaN.

hydrogen conditions, where rms roughness values of 1–2 nm were measured. These observations, along with the increased growth rate, are consistent with shifting the growth kinetics towards a more nitrogen-stable growth.

The effect of hydrogen on growth can be illustrated another way. Many groups using MBE have reported a trend similar to that shown by the filled triangles in Fig. 8. That is, the growth rate of GaN is fairly constant for fixed flux values as the temperature is increased up to some value, typically, in the range 700–800 °C, where the growth rate rapidly decreases. This temperature depends on the absolute magnitudes of the flux and may be related to the onset of rapid desorption of Ga from the growing surface.[30] The growth rate can be recovered for temperatures up to the limit of our heater by introducing atomic hydrogen, as indicated by the filled circles in Fig. 8. Thus, the atomic hydrogen may be affecting the desorption kinetics of Ga as well. The stabilizing effect of atomic hydrogen for the growing GaN surface deserves additional study.

VI. CONCLUSION

In conclusion, our nucleation and growth conditions produce (0001)-oriented GaN on basal-plane sapphire. For this orientation, N-stable growth produced a granular, 3D surface while Ga-stable growth resulted in a smoother surface morphology indicative of 2D growth. The presence of pyramidal hillocks could be directly linked to the presence of IDs in our GaN layers. Nitrogen-stable growth and growth under atomic hydrogen enhanced the growth rate of IDs with respect to the surrounding matrix, whereas growth under Ga-stable conditions resulted in a more nearly equal growth rate. IDs apparently originated in the initial nucleation layer, and were stable with respect to layer growth due to a differential growth rate. We have demonstrated that the presence of atomic hydrogen can have a significant effect on the growth rate of GaN under Ga-stable conditions. Hydrogen may stabilize the growing GaN surface at higher temperatures.

ACKNOWLEDGMENTS

Work at WVU was supported by DoD/ONR Grant No. N00014-94-1-1149 and ONR Grant No. N00014-96-1-1008. The work at Xerox was supported by DARPA MDA972-96-3-0014.

[1]See, for example, J. Pankove, Mater. Res. Soc. Symp. Proc. **97**, 409 (1987).
[2]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Jpn. J. Appl. Phys., Part 2 **35**, L74 (1996).
[3]T. D. Moustakas, T. Lei, and R. J. Molnar, Physica B **185**, 36 (1993).
[4]M. Smith, G. D. Chen, J. Z. Li, J. Y. Lin, H. X. Jiang, A. Salvador, W. K. Kim, O. Aktas, A. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **67**, 3387 (1995).
[5]M. A. L. Johnson, Z. Yu, C. Boney, W. C. Hughes, J. W. Cook, Jr., J. F. Schetzina, H. Zhao, B. J. Skromme, and J. A. Edmond, Mater. Res. Soc. Symp. Proc. **449**, 215 (1997).
[6]E. J. Tarsa, B. Heying, X. H. Wu, P. Fini, S. P. DenBaars, and J. S. Speck, J. Appl. Phys. **82**, 5472 (1997).
[7]Z. Yu, S. L. Buczkowski, N. C. Giles, T. H. Myers, and M. R. Richards-Babb, Appl. Phys. Lett. **69**, 2731 (1996); S. L. Buczkowski, Z. Yu, M. R. Richards-Babb, N. C. Giles, T. H. Myers, and L. T. Romano, Mater. Res. Soc. Symp. Proc. **449**, 197 (1997).
[8]H. Reichert, R. Averbeck, A. Graber, M. Schienle, U. Straub, and H. Tews, Mater. Res. Soc. Symp. Proc. **449**, 149 (1997).
[9]R. Beresford, K. S. Stevens, Q. Cui, A. Schwartzman, and H. Cheng, Mater. Res. Soc. Symp. Proc. **449**, 361 (1997).
[10]R. Held, D. E. Crawford, A. M. Johnston, A. M. Dabiran, and P. I. Cohen, J. Electron. Mater. **26**, 272 (1997).
[11]L. T. Romano, J. E. Northrup, and M. A. O'Keefe, Appl. Phys. Lett. **69**, 2394 (1996).
[12]B. Daudin, J. L. Rouviere, and M. Arley, Appl. Phys. Lett. **69**, 2480 (1996).
[13]X. H. Wu, L. M. Brown, D. Kapolonek, S. Keller, B. Keller, S. P. DenBaars, and J. S. Speck, J. Appl. Phys. **80**, 3228 (1996).
[14]Z. Liliental-Weber, H. Sohn, N. Newman, and J. Washburn, J. Vac. Sci. Technol. B **13**, 1578 (1995).
[15]L. T. Romano, B. S. Krusor, R. Singh, and T. D. Moustakas, J. Electron. Mater. **26**, 285 (1997).
[16]L. T. Romano and T. H. Myers, Appl. Phys. Lett. **71**, 3486 (1997).
[17]V. Potin, P. Ruterana, and G. Nouet, J. Appl. Phys. **82**, 2176 (1997).
[18]M. S. Brandt, J. W. Ager III, W. Gotz, N. M. Johnson, J. S. Harris, Jr., R. J. Molnar, and T. D. Moustakas, Phys. Rev. B **49**, 14 758 (1994).
[19]M. Seelmann-Eggebert, J. L. Weyher, H. Obloh, H. Zimmermann, A. Rar, and S. Porowski, Appl. Phys. Lett. **71**, 2635 (1997).
[20]M. Richards-Babb, S. L. Buczkowski, Z. Yu, and T. H. Myers, Mater. Res. Soc. Symp. Proc. **395**, 237 (1996).
[21]D. Lederman, Z. Yu, T. H. Myers, and M. R. Richards-Babb, Appl. Phys. Lett. **71**, 368 (1997).
[22]D. L. Smith, *Thin Film Deposition* (McGraw-Hill, New York, 1995), pp. 119–184.
[23]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
[24]P. G. Middleton, C. Trager-Cowan, A. Mohammed, K. P. O'Donnell, W. Van Der Stricht, I. Moerman, and P. Demeeser, Mater. Res. Soc. Symp. Proc. **449**, 471 (1997).
[25]Z. Liliental-Weber, Y. Chen, S. Ruvimov, and J. Washburn, Phys. Rev. Lett. **79**, 2835 (1997).
[26]EPI MBE Products Group (St. Paul, MN), Application Note 97-3 (August, 1997).
[27]See, for example, V. A. Burrows, Y. J. Chabal, G. S. Higashi, K. Raghavachari, and S. B. Christman, Appl. Phys. Lett. **53**, 998 (1988).
[28]See, for example, *The Properties of Natural and Synthetic Diamond*, edited by J. E. Field (Academic, London, 1992), and references therein.
[29]H. Shimomura, Y. Okada, and M. Kawabe, Jpn. J. Appl. Phys., Part 2 **31**, L628 (1992).
[30]C. R. Jones, T. Lei, R. Kaspi, and K. R. Evans, Proceedings of the 1995 Fall MRS Meeting, Symposium AAA (unpublished).

Epitaxial growth and properties of III–V magnetic semiconductor (GaMn)As and its heterostructures

Masaaki Tanaka[a)]
Department of Electronic Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan and PRESTO (Sakigake-21), Japan Science and Technology Corporation, Kawaguchi 332, Japan

(Received 21 January 1998; accepted 26 May 1998)

We present epitaxial growth, structures, magnetization, magneto-transport, and magneto-optic properties of (GaMn)As thin films and (GaMn)As/AlAs superlattice heterostructures grown by molecular beam epitaxy (MBE) on GaAs substrates. $(Ga_{1-x}Mn_x)As$ is a new class of III–V based magnetic alloy semiconductor grown by low temperature MBE, which contains a large amount of Mn atom (Mn concentration x up to 8%) far above the equilibrium solubility of Mn in GaAs. The feasibility of preparing such a III–V based magnetic semiconductor and its heterostructures gives a new degree of freedom in the materials design of III–V systems, offering new opportunities to explore spin-related phenomena as well as potential device applications using both magnetic and electronic/optical functions in III–V semiconductors. © *1998 American Vacuum Society.*
[S0734-211X(98)10604-2]

I. INTRODUCTION

Semiconductors and magnetic materials are two very important materials in the current electronics industry, as well as being two big branches of solid state physics. Most electronic and optical devices are made of semiconductors, while magnetic materials are used for nonvolatile memory, magnetic sensors, optical isolators, etc. However, integration of semiconductors and magnetic materials is very difficult, because the two materials are generally very dissimilar. One of the approaches for the integration of semiconductors and magnetic materials is the epitaxial growth of ferromagnetic thin films on semiconductors.[1] In the past, we have successfully grown ferromagnetic metallic compound (MnGa, MnAs)/III–V semiconductor heterostructures with thermodynamically stable interfaces by molecular beam epitaxy (MBE).[2]

Another more traditional approach is to prepare diluted magnetic semiconductors (DMSs) or semimagnetic semiconductors (SMSCs).[3] DMSs are compound alloy semiconductors containing a large fraction of magnetic ions (Mn^{2+}, Cr^{2+}, Fe^{2+}), and were studied mainly on II–VI based materials.[4] This is because such 2+ magnetic ions are easily incorporated into the host II–VI crystals by replacing group-II cations. In such II–VI based DMSs such as (CdMn)Te, magneto-optic properties were extensively studied, and optical isolators were recently fabricated using their large Faraday effect.[5] However, carrier control is generally difficult in II–VI semiconductors and no ferromagnetic order has been found in II–VI DMSs.

On the other hand, (InMn)As was the only III–V based DMS studied in detail until recently.[6,7] In (InMn)As, carrier control is possible and carrier (hole) induced ferromagnetism was found in p-type samples.[7,8] Since (InMn)As is based on InAs, the lattice constant is much larger than that of GaAs substrates with lattice mismatch of about 7%, and the band gap is narrow (about 0.36 eV). For a wider range of materials design and applications in III–V systems, a new type of magnetic semiconductors based on GaAs, the most widely used compound semiconductor, needs to be prepared. Recently, De Boeck *et al.* have shown that ferromagnetic MnAs clusters can be controllably buried in GaAs by annealing (GaMn)As grown by low temperature MBE (LT-MBE).[9,10] However, there had been no studies on the properties of (GaMn)As and its heterostructures when we started this research.[11–13] In this article, we present and review our study on the epitaxial growth, structures, magnetization, magneto-transport, and magneto-optic properties of (GaMn)As thin films and (GaMn)As/AlAs superlattice heterostructures. (GaMn)As is a new class of III–V based magnetic semiconductor, which could be potentially coupled with current III–V electronics and optoelectronics. The feasibility of growing III–V magnetic heterostructures in particular can offer a unique opportunity to couple spin-related phenomena with well-established band engineering in III–V semiconductors.

II. MOLECULAR BEAM EPITAXIAL GROWTH AND STRUCTURAL CHARACTERIZATIONS

A. MBE growth

The equilibrium solubility of Mn in GaAs is known to be at most 10^{19} cm^{-3} (about 0.002 at mole fraction). Thus we used low temperature MBE growth, in which strong non-equilibrium growth conditions are realized, so that a large number of Mn atoms can be incorporated into GaAs. Here, we briefly describe the MBE growth process of (GaMn)As thin films.[11]

After growing a GaAs buffer layer on semi-insulating GaAs(001) substrates at 580 °C under normal growth conditions, the substrate temperature was cooled to 200–300 °C [for homogeneous (GaMn)As, typically 250 °C], during which the surface reconstruction changed from (2×4) to

a)Electronic mail: masaaki@ee.t.u-tokyo.ac.jp

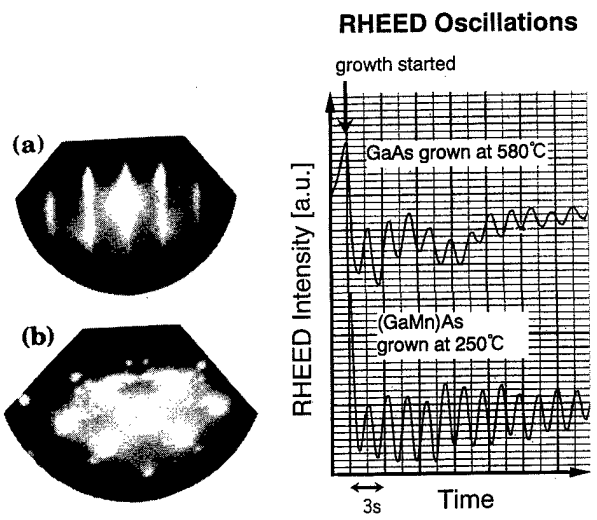


FIG. 1. (a) Typical RHEED pattern during the growth of homogeneous $(Ga_{1-x}Mn_x)As$ with Mn content $x=0.05$ at a substrate temperature of 250 °C. On the right, RHEED oscillations at the initial stage of the (GaMn)As growth at 250 °C are shown, compared with the RHEED oscillations during the growth of the GaAs buffer layer at 580 °C. (b) Typical RHEED pattern of the inhomogeneous structures, which was taken during the growth of $(Ga_{1-x}Mn_x)As$ with nominal Mn content $x=0.10$ at a substrate temperature of 300 °C. Here hexagonal MnAs with a $(\bar{1}101)$ face appeared in the background of zinc-blende (GaMn)As or the GaAs lattice.

$c(4\times4)$. $(Ga_{1-x}Mn_x)As$ thin films with various Mn contents x (0–0.17) were grown on this $c(4\times4)$ surface at various temperatures (200–500 °C) at a growth rate of 0.5–1.0 μm/h. Figure 1(a) shows a typical reflection high-energy electron diffraction (RHEED) pattern during the growth of homogeneous $(Ga_{1-x}Mn_x)As$ with Mn content $x=0.05$ at a substrate temperature of 250 °C. A streaky RHEED pattern with zinc-blende symmetry was seen, suggesting that homogeneous (GaMn)As ternary alloy films were grown. This homogeneous (GaMn)As alloy was obtained when the Mn content x and the growth temperature were relatively low. At 250 °C, homogeneous alloys were obtained up to $x=0.08$. The RHEED pattern of (GaMn)As was (1×1) or (1×2), the latter was observed at higher temperatures (around 250 °C). At the initial stage of the growth of (GaMn)As, RHEED oscillations were seen, as shown on the right-hand side of Fig. 1, indicating that the growth proceeds layer by layer with two-dimensional nucleation.

In contrast, when the nominal Mn content and growth temperature are relatively high, we have seen inhomogeneous structures, clustering of hexagonal MnAs during the growth, as shown in the RHEED pattern of Fig. 1(b) which was taken during the growth of $(Ga_{1-x}Mn_x)As$ with nominal Mn content $x=0.10$ at the substrate temperature of 300 °C. Here the RHEED pattern shows that hexagonal MnAs with a $(\bar{1}101)$ face appeared in the background of zinc-blende (GaMn)As or GaAs.

This result clearly shows the limitation of the solubility of Mn in GaAs during low temperature MBE growth. Figure 2 shows a diagram of the film properties of (Ga, Mn, As) in relation to two growth parameters, growth temperature and

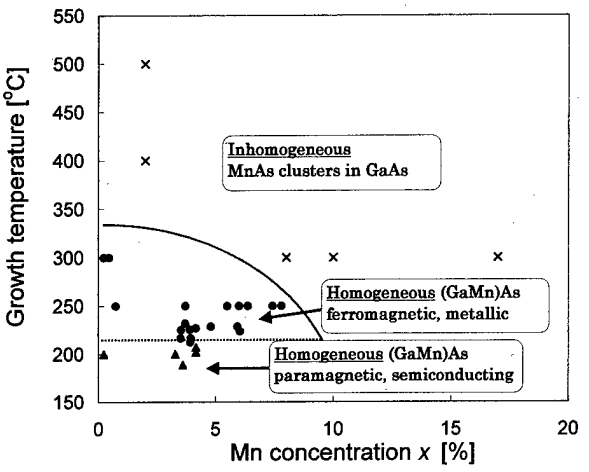


FIG. 2. Diagram of the film properties of (Ga, Mn, As) in relation to two growth parameters, the growth temperature and Mn concentration x. Transport and magnetic properties of the homogeneous (GaMn)As are shown to depend on the growth temperature.

Mn concentration x. In this article, we concentrate on the homogeneous (GaMn)As films. In Fig. 2, we have also shown that transport and magnetic properties of the homogeneous (GaMn)As depend on the growth temperature; (GaMn)As films grown at lower temperatures are paramagnetic and semiconducting (electrical resistivity becomes large with lowering temperature), whereas (GaMn)As films grown at higher temperatures are ferromagnetic and metallic.[14]

B. Structural characterizations

In order to characterize the structure of the homogeneous $(Ga_{1-x}Mn_x)As$ films with the thicknesses of 0.2–1.4 μm grown on GaAs(001) substrates, we have measured x-ray diffraction spectra (both θ–2θ and rocking curves). The linewidth of (GaMn)As peaks was about 150 s, only a little broader than that of GaAs substrates, indicating high epitaxial quality. The values of the lattice constant of (GaMn)As are slightly larger than that of GaAs. Since the (GaMn)As films grown on GaAs are under compressive strain, their lattice spacings are expanded in the vertical direction. We calculated intrinsic lattice constants $a_{(GaMn)As}$ of cubic (GaMn)As assuming that the elastic constants of (GaMn)As are equal to that of GaAs, and they are plotted in Fig. 3 as a function of Mn content x. The values of $a_{(GaMn)As}$ are on a straight line following Vegard's law. The maximum lattice

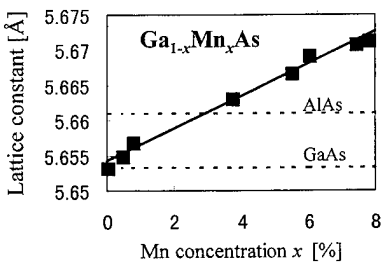


FIG. 3. Intrinsic lattice constants $a_{(GaMn)As}$ of cubic (GaMn)As as a function of Mn content x. The lattice constants of GaAs and AlAs are also shown.

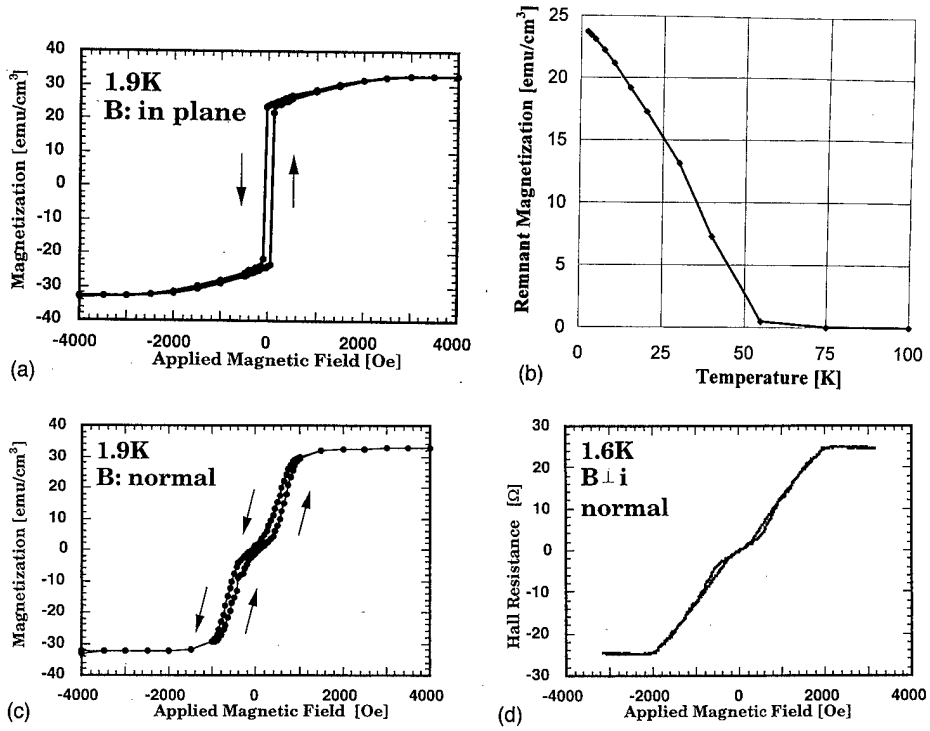


FIG. 4. (a) Magnetization ($M-H$) curve of a 1.4 μm thick $(Ga_{1-x}Mn_x)As$ film with $x=0.074$ measured by a SQUID at 1.9 K, when the magnetic field was applied in plane along the ⟨110⟩ axis. (b) Temperature dependence of remnant magnetization M_r, indicating the Curie temperature of this sample is about 60 K. (c) $M-H$ curve of the same sample measured by a SQUID at 1.9 K, when the magnetic field was applied perpendicular to film plane. (d) Hall resistance of the same sample measured at 1.6 K as a function of the magnetic field perpendicular to the film plane.

mismatch to GaAs is only 0.37% at $x=0.078$, indicating good structural compatibility with well-established GaAs/AlAs systems. Very recently, local structures around Mn in the (GaMn)As films were investigated using the Mn K-edge extended x-ray fine structure (EXAFS) technique.[15] It was found that Mn atoms are substituted for the Ga sites in the zinc-blende GaAs lattice.

In contrast, in the inhomogeneous samples whose RHEED patterns were like Fig. 1(b), hexagonal MnAs peaks were also observed in the x-ray spectra. The crystal orientation of MnAs clusters in GaAs [or possibly (GaMn)As] is mainly $(\bar{1}101)$, in agreement with the RHEED observations.

III. MAGNETIC PROPERTIES OF (GaMn)As

A. Magnetization

First, magnetization measurements were performed on the homogeneous (GaMn)As films using a superconducting quantum interference device (SQUID). Figure 4(a) shows a magnetization ($M-H$) curve of a 1.4 μm thick $(Ga_{1-x}Mn_x)As$ film with $x=0.074$ measured at 1.9 K when the magnetic field was applied in plane along the ⟨110⟩ axis. This sample is p type with a hole concentration of 2×10^{20} cm^{-3}. A squarelike hysteresis was observed at low field, indicating a ferromagnetic order. The value of coercive field H_c was 56 Oe, and remnant magnetization M_r and saturation magnetization M_s are 24 and 35 emu/cm^{-3}, respectively. This M_s value corresponds to 2.3 μ_B per Mn atom, where μ_B is the Bohr magneton. This hysteresis behavior remained when the temperature was increased up to about 60 K, above which remnant magnetization M_r disappeared. Figure 4(b) shows the temperature dependence of M_r, indicating the Curie temperature of this sample is about 60 K, much higher than that (7.5 K) of p-type (InMn)As.[7] This result rules out the possible contribution of MnAs clusters in this film, because the Curie temperature of MnAs is 313 K.

Figure 4(c) shows a $M-H$ curve at 1.9 K when the magnetic field was applied perpendicular to the film plane. A distorted hysteresis was observed, exhibiting much higher saturation field (1500–2000 Oe). This result indicates that the perpendicular direction is a hard axis of magnetization, and that the easy magnetization axis of the present (GaMn)As film lies in the plane. In our experiments, most of the (GaMn)As films showed in-plane magnetization, although a (GaMn)As film with low Mn content ($x=0.005$) showed the presence of perpendicular magnetization.[12] This suggests that the strain in the epitaxial films plays an important role in determining the magnetic anisotropy. When the compressive strain in the (GaMn)As films is very weak for $x\leqslant0.005$, perpendicular magnetization appears, while the compressive strain in the (GaMn)As films with $x\geqslant0.005$ results in in-plane magnetization. It is reported that a (GaMn)As film grown on InGaAs with tensile strain showed clear perpendicular magnetization.[13]

B. Transport properties

Transport measurements were done on photolithographically patterned Hall bars. All the (GaMn)As samples showed p-type conduction with the hole concentrations ranging from low 10^{18} to low 10^{20} cm^{-3}, because the Mn atoms which are incorporated into the Ga sites act as acceptors. In general, the Hall resistance R_{Hall} and the Hall resistivity ρ_{Hall} of magnetic materials is expressed as[16]

$$R_{Hall}d=\rho_{Hall}=R_0B+R_sM, \quad (1)$$

where d is the sample thickness, R_0 the ordinary Hall coefficient, B the magnetic flux density, R_s the extraordinary Hall coefficient, and M the magnetization. The first term represents the ordinary Hall effect caused by Lorenz force, and the second term is the extraordinary Hall effect (EHE) caused by the anisotropic scattering between carriers and local magnetic moments. When the materials are ferromagnetic (at low temperature), one can neglect the first term because the second term is much larger. If we assume R_s is represented as $c\rho$ (c: constant, ρ: the resistivity of the sample) for simplicity, the Hall resistance is proportional to the magnetization perpendicular to the film plane; $R_{Hall}d=\rho_{Hall}=c\rho M$. Hall measurements were done on the previous $(Ga_{1-x}Mn_x)As$ film with $x=0.074$ at 1.6 K, as shown in Fig. 4(d). Good agreement was seen between the $M-H$ curve of Fig. 4(c) and the $R_{Hall}-H$ curve of Fig. 4(d), indicating that the EHE is dominant in the Hall data.

Using the temperature dependence of the EHE, we can estimate the Curie temperature T_C of (GaMn)As. Figure 5(a) shows data of Hall resistance R_{Hall} versus magnetic field perpendicular to the film at various temperatures (31–110 K) on a 50 nm thick $(Ga_{1-x}Mn_x)As$ film with $x=0.039$. One can see ferromagnetic behaviors similar to those seen in Figs. 4(c) and 4(d) at lower temperatures, which turn to paramagnetic behavior with the increase of temperature. Since R_{Hall} is proportional to M, the magnetic susceptibility χ, which follows the Curie–Weiss law as $\chi=C/(T-T_C)$ where C is the Curie constant, can be obtained from the zero-field slopes of Fig. 5(a). Then the temperature dependence of $(R_H^*/\rho)^{-1}$, which should be proportional to χ^{-1}, was plotted as shown in Fig. 5(b), where $R_H^*(=R_{Hall}/B)$ is the Hall coefficient. Thus the Curie temperature of this sample was estimated to be 78 K.

We can also estimate the Curie temperature T_C by using the Hall data under higher magnetic field. By using the simple relation $\rho_{Hall}/\rho=cM$ described above, one can make the Arrott plot $(\rho_{Hall}/\rho)^2$ vs $B/(\rho_{Hall}/\rho)$ at high field regions, and then the temperature dependence of saturation magnetization was obtained. Thus we have estimated T_C of this sample to be 74 K, in good agreement with the estimation by using low field Hall data.

In this way, the Curie temperatures T_C of various $(Ga_{1-x}Mn_x)As$ films were estimated. While the value of T_C tends to increase as x increases up to $x=0.03$, there was no special correlation at $0.03<x<0.08$, where T_C was between 45 and 80 K. However, there is a strong correlation between the low-temperature hole concentration and the appearance

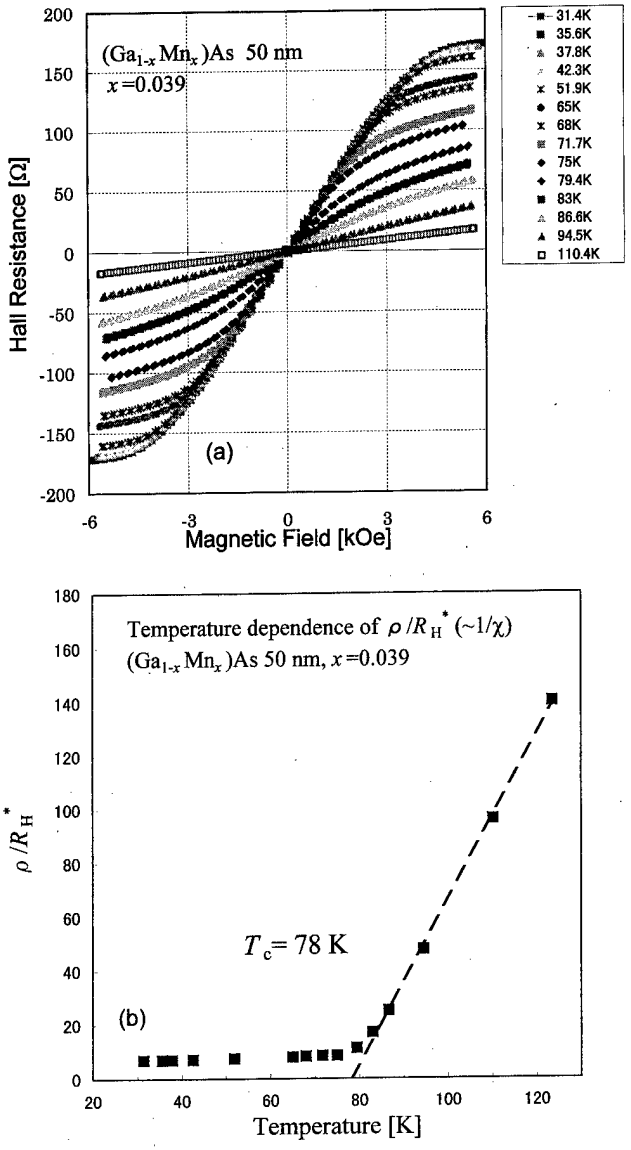


FIG. 5. (a) Hall resistance R_{Hall} as a function of the magnetic field perpendicular to the film at various temperatures (31–120 K) on a 50 nm thick $(Ga_{1-x}Mn_x)As$ film with $x=0.039$. (b) Temperature dependence of $(R_H^*/\rho)^{-1}$, which should be proportional to χ^{-1}.

of ferromagnetic order, as described in Sec. II A. This indicates that the ferromagnetism of (GaMn)As was induced by holes. It is considered that the holes are supplied by Mn acceptors, but are likely to be compensated for by a number of defects such as As antisites caused by low-temperature MBE growth. Therefore, for precisely controlling magnetic properties of (GaMn)As, defect control is important. For this purpose, further systematic study on the growth condition dependence is underway.[14]

C. Magneto-optic properties

In order to investigate magneto-optic properties and band structure of (GaMn)As, we have measured magnetic circular dichroism (MCD) spectra on our samples.[17] MCD measures the optical reflectance difference between σ_+ and σ_- circu-

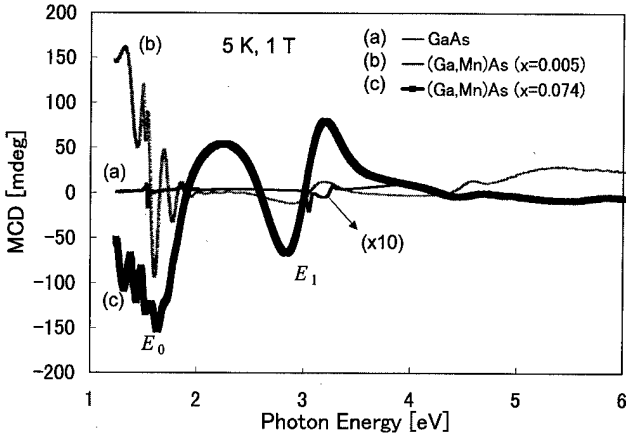


FIG. 6. MCD spectra of (a) a GaAs substrate as a reference and two $(Ga_{1-x}Mn_x)As$ films, (b) $x=0.005$, thickness 1.0 μm, and (c) $x=0.074$, thickness 1.4 μm, both grown on GaAs(001), measured at 5 K under a perpendicular magnetic field of 1 T. In the spectra of (GaMn)As, the MCD was greatly enhanced at E_0 and E_1. Oscillations at around 1.5 eV are due to the interference in the epitaxial thin films.

lar polarizations.[18] The polarization modulation technique was used with a quartz stress modulator operating at 50 kHz in order to sensitively detect band splittings induced by the magnetic field. The direction of the input and reflected light beams deviated by 10°. The magnetic field was applied perpendicular to the sample plane from the surface to the substrate direction. In our experiments, the MCD signal in units of degrees is expressed as

$$\frac{90}{\pi}\frac{R_+-R_-}{R_++R_-}, \tag{2}$$

where R_+ and R_- are the reflectivities for σ_+ and σ_- circular polarizations, respectively.

Figure 6 shows MCD spectra of a GaAs substrate as a reference, and of two $(Ga_{1-x}Mn_x)As$ films; $x=0.005$, thickness 1.0 μm and $x=0.074$, thickness 1.4 μm, both grown on GaAs(001). The measurements were done at 5 K under the perpendicular magnetic field of 1 T. Although the MCD signal of GaAs was very weak, sharp peaks corresponding to the optical transitions at Γ and Λ critical points, which are denoted as E_0 (band gap energy, 1.519 eV) and E_1 (3.04 eV), respectively, are clearly observed. For (GaMn)As, the MCD intensity is greatly enhanced with increasing the Mn content x, particularly at the critical point energies of E_0 and E_1. This indicates that $sp-d$ hybridization is very strong, and that the band structure of (GaMn)As is similar to that of zinc-blende type semiconductors. By excluding interference oscillations which are superimposed with the MCD spectra of (GaMn)As at around 1.5 eV, the transition energy corresponding to the band gap of (GaMn)As with $x=0.005$ can be estimated to be 1.50 eV, which is very close to the band gap of GaAs. Furthermore, from the polarity of the MCD signals, the $p-d$ exchange interaction in (GaMn)As was found to be antiferromagnetic in (GaMn)As,[17] in contrast with the ferromagnetic $p-d$ exchange reported in Mn-doped GaAs with doping concentrations of $10^{17}-10^{18}$ cm^{-3}.[19]

IV. III–V BASED MAGNETIC SEMICONDUCTOR [(GaMn)As]/NONMAGNETIC SEMICONDUCTOR [AlAs] SUPERLATTICES

Since knowledge of III–V based magnetic semiconductors is limited so far, only a few papers on the ultrathin heterostructures containing III–V based magnetic semiconductors[20,21] were published. Those papers, however, did not report on the band structure control, such as the shift of transition energy and the formation of subbands, in the ultrathin layers of III–V based magnetic semiconductors. Based on the previously described fundamental properties of (GaMn)As, which show good compatibility with the existing III–V semiconductors in terms of crystal structure and band structure, we are now able to design and fabricate a new class of magnetic quantum heterostructures.[22] In Sec. IV, we show the MBE growth and properties of III–V based superlattices (SLs), consisting of magnetic [(GaMn)As] and nonmagnetic [AlAs] semiconductors. The SLs are formed with excellent crystal quality and abrupt interfaces, and some of them exhibit a ferromagnetic order at low temperature, depending on the (GaMn)As thickness. The blueshift of the interband transition energy and the formation of the subbands are observed in the MCD spectra.

A. Growth and structure

(GaMn)As/AlAs SLs were grown on semi-insulating (001) GaAs substrates by MBE. After growing a GaAs buffer layer at 600 °C under normal growth conditions, the substrate temperature was cooled to 250 °C. The As_4 flux was kept on throughout the growth. During the cooling process, the RHEED pattern changed from (2×4) to $c(4\times4)$. On this $c(4\times4)$-(001) GaAs surface, a AlAs layer was first grown at a growth rate of 0.3 μm/h, and then a (GaMn)As layer was grown at 0.5 μm/h. This cycle was repeated 20–31 times to grow a SL structure at 250 °C. The growth was interrupted for 10–20 s at each interface. The film thicknesses of (GaMn)As and AlAs were 45–120 and 30–68 Å, respectively. Figure 7 shows typical RHEED patterns with the azimuths of [110] and $[\bar{1}10]$ taken from a (GaMn)As layer and a AlAs layer during growth of the SLs. Note that both RHEED patterns were fairly streaky despite low-temperature growth, and that a clear (1×2) reconstruction was observed on the (GaMn)As surface whereas almost no reconstruction was seen on the AlAs surface. This (1×2) reconstructed (GaMn)As surface and (1×1) unreconstructed AlAs surface were repeatedly observed during the growth of the (GaMn)As/AlAs SLs.

To characterize the structure of the (GaMn)As/AlAs SLs, x-ray diffraction measurements were done on all the samples. Figure 8(a) shows a θ–2θ spectrum of a SL of $[(Ga_{1-x}Mn_x)As$ (45 Å)/AlAs (30 Å)] with $x=0.04$ around the GaAs(002) substrate peak (31.6°). Six satellite peaks are clearly observed on both sides of the GaAs peak. Similar results were obtained for all other SL samples. This indicates that SL structures of excellent crystal quality were formed with abrupt interfaces between (GaMn)As and AlAs. Figure 8(b) shows a cross-sectional transmission electron mi-

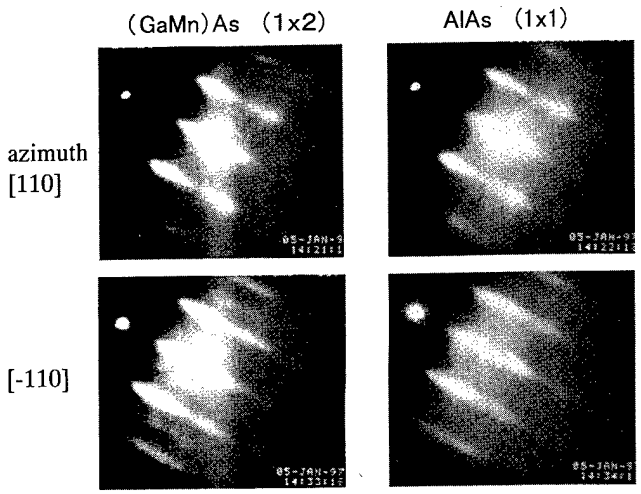


FIG. 7. Typical RHEED patterns with the azimuths of [110] and [$\bar{1}$10] taken from a (GaMn)As layer and a AlAs layer during growth of the SLs. Note that a clear (1×2) reconstruction is observed on the (GaMn)As surface whereas almost no reconstruction is seen on the AlAs surface.

croscopy (TEM) image of a [$(Ga_{1-x}Mn_x)$As (50 Å)/AlAs (30 Å)] SL with $x=0.05$ grown under the same growth conditions, clearly indicating that the superlattice structure was formed with fairly abrupt interfaces.

B. Magnetization and transport properties

First, magnetic properties of the SLs were studied using a SQUID. The closed circles and the lines in Fig. 9 show a $M-H$ (magnetization versus applied magnetic field) curve of a SL of [(GaMn)As (120 Å)/AlAs (68 Å)] with Mn content $x=0.060$ measured at 2.0 K when the magnetic field was applied in plane. Clear hysteresis was observed, indicating the SL is ferromagnetic at low temperature. The coercive field H_c was 388 Oe at 2.0 K, larger than that (~50 Oe) of the 1.4 μm thick (GaMn)As single layer [Fig. 4(a)]. When the magnetic field was applied perpendicular to the film plane, a distorted little hysteresis was observed as shown in Fig. 9 by the broken lines. Similar results were observed for other SL [(GaMn)As (70 Å)/AlAs (30 Å)]. This indicates that the easy axis of magnetization lies in plane, as in the case of (GaMn)As single films grown on GaAs(001). In contrast, for the SL samples with (GaMn)As thickness of 65 Å or less, no ferromagnetic order was found at the temperature range (2.0–300 K) measured in this experiment.

To investigate the magneto-transport properties, we have done Hall and resistivity measurements at various temperatures. The SLs with narrow (≤65 Å) (GaMn)As layers were highly resistive and transport measurements were difficult, probably because carriers are localized due to structural fluctuations such as interface roughness. The fluctuation of the quantum confinement energy due to interface roughness is known to increase dramatically as the (GaMn)As quantum well thickness is reduced. Hence, we measured the transport of the SLs with wider (≥70 Å) (GaMn)As. By analyzing the results in the same manner described in Sec. III B, we derived the temperature dependence of the extraordinary Hall

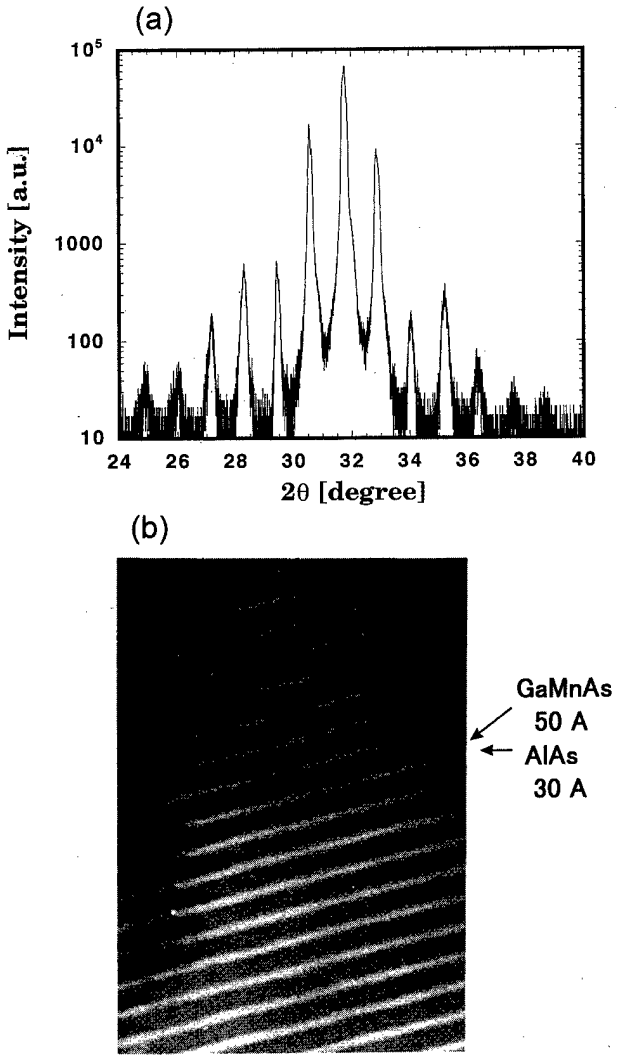


FIG. 8. (a) X-ray θ–2θ spectrum of a SL of [$(Ga_{1-x}Mn_x)$As (45 Å)/AlAs (30 Å)] with $x=0.04$ around the GaAs(002) substrate peak (31.6°). Six satellite peaks are clearly observed on both sides of the GaAs peak. (b) Cross-sectional TEM image of a [$(Ga_{1-x}Mn_x)$As (50 Å)/AlAs (30 Å)] SL with $x=0.05$ grown under the same growth conditions.

effect which is proportional to the magnetization. By plotting the temperature dependence of $(R_H^*/\rho_{\text{sheet}})^{-1}$, which should be proportional to χ^{-1}, and also by making Arrott plots [$(\rho_{\text{Hall}}/\rho_{\text{sheet}})^2 - B/(\rho_{\text{Hall}}/\rho_{\text{sheet}})$ plots], where ρ_{sheet} is the sheet resistivity, we have estimated the Curie temperature of the SLs. The Curie temperatures of the SLs of [(GaMn)As (120 Å)/AlAs (68 Å)] with Mn content $x=0.06$ and of [(GaMn)As (70 Å)/AlAs (30 Å)] with $x=0.049$ were 63 and 45 K, respectively. The experimental fact that the SLs with wider (GaMn)As (≥70 Å) are ferromagnetic and conductive, whereas the SLs with narrow (≤50 Å) (GaMn)As are not ferromagnetic and are highly resistive, strongly suggests that the ferromagnetic order in (GaMn)As is induced by carriers (holes). In the SL of [(GaMn)As (70 Å)/AlAs (30 Å)] whose (GaMn)As width gives the minimum value for the presence

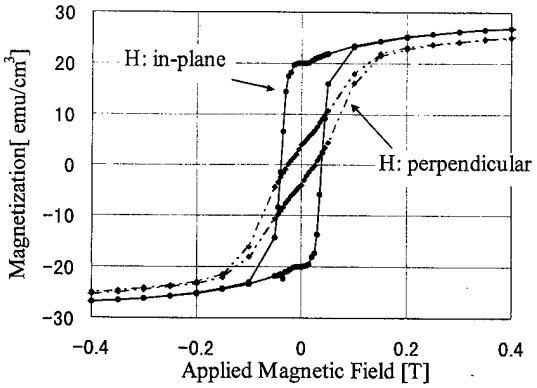


FIG. 9. $M-H$ (magnetization vs applied magnetic field) characteristics of a SL of [(GaMn)As (120 Å)/AlAs (68 Å)] with Mn content $x=0.060$ measured by a SQUID at 2.0 K when the magnetic field was applied in plane (solid lines) and perpendicular to the plane (broken lines).

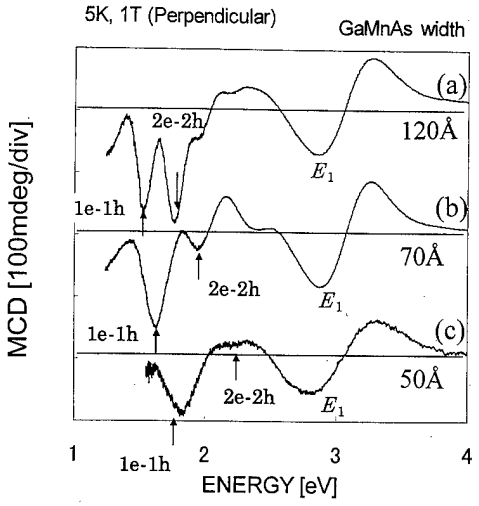


FIG. 10. MCD spectra of various (GaMn)As/AlAs SLs measured at 5 K under a magnetic field of 1 T. The thicknesses of $(Ga_{1-x}Mn_x)As$ and AlAs and the Mn concentration x in the SLs were (a) 120 and 68 Å, $x=0.06$, (b) 70 and 30 Å, $x=0.049$, and (c) 50 and 30 Å, $x=0.047$, respectively. The spectrum of (c) is magnified 10 times. The arrows show calculated transition energies assuming the Kronig–Penny model.

of ferromagnetic order, the hole concentration was estimated to be $1-2\times10^{13}$ cm^{-2} per (GaMn)As layer. The magnetic properties of the SLs compared with a 1.4 μm thick (GaMn)As film are summarized in Table I.

C. Magneto-optic properties: Quantum size effect and formation of subbands

To investigate the magneto-optic properties and the band structures of the (GaMn)As/AlAs SLs, we performed MCD measurements under the same experimental setup described earlier (Sec. III C). MCD spectra of the (GaMn)As/AlAs SLs at 5 K under a magnetic field of 1 T are shown in Fig. 10. The thicknesses of $(Ga_{1-x}Mn_x)As$ and AlAs and the Mn content x in the SLs were 120 and 68 Å, $x=0.06$, 70 and 30 Å, $x=0.049$, and 50 and 30 Å, $x=0.047$ in Figs. 10(a)–10(c), respectively. The spectrum of Fig. 10(c) is magnified 10 times because the MCD of Fig. 10(c) is much weaker than that of Figs. 10(a) and 10(b). The reason is that the SLs of Figs. 10(a) and 10(b) are ferromagnetic, whereas the SL of Fig. 10(c) is paramagnetic at this temperature. It was found that these SLs have large MCD peaks at the critical point energies, which correspond to the transition energies between the electron subbands and hole subbands at the Γ point of (GaMn)As. Large peaks were also seen at the Λ points at the energy of $E_1\sim3$ eV, as in the case of thick (GaMn)As films. As the thickness of (GaMn)As in the SLs decreases, the lowest transition energy, which is $1e-1h$ (the first electron subband to the first heavy-hole subband at Γ), is shifted to higher energy due to quantum confinement effect. Also, a higher transition peak, which is assigned to $2e-2h$, was observed in the SLs with wider (GaMn)As layers. We have calculated the transition energies using the Kronig–Penny model assuming that (GaMn)As has the same band parameters as those of GaAs, marked in Fig. 10 by arrows. Good agreement is seen between the observed transition energies and the calculated ones. This is a clear demonstration of the quantum confinement effect and two-dimensional subband formation in III–V magnetic semiconductor systems.

TABLE I. Summarized magnetic properties of (GaMn)As SLs compared with a 1.4 μm thick $(Ga_{1-x}Mn_x)As$ film ($x=0.074$). The structure parameters [Mn content x (GaMn)As and AlAs widths], saturation magnetization M_s, coercive field H_c, and Curie temperature T_C are shown. At low temperatures, the SLs with wider (120 and 70 Å) (GaMn)As are ferromagnetic and conductive, whereas the SLs with narrower (50 and 45 Å) (GaMn)As are not ferromagnetic but are paramagnetic and highly resistive.

Mn content x	0.074	0.06	0.05	0.05	0.04
(GaMn)As width (Å)	1.4 μm	120	70	50	45
AlAs width (Å)	···	68	30	30	30
M_s (emu/cm^3)	35	32	28	···	···
H_c (Oe)	56	388	500	···	···
T_C (K)	60	63	45	···	···

V. CONCLUSIONS

A new class of III–V based ferromagnetic semiconductor $(Ga_{1-x}Mn_x)As$ was grown by low temperature MBE with the Mn content x up to 0.08, exceeding the equilibrium solubility of Mn in GaAs. The (GaMn)As films showed p-type conduction, and turned out to be ferromagnetic at transition temperatures T_c of 45–80 K. The Hall effect of the (GaMn)As films was dominated by the extraordinary Hall effect, from which the values of T_c were also estimated. A large magneto-optic effect was found in the MCD measurements of (GaMn)As, where the MCD signal was greatly enhanced near the critical point energies of GaAs. These MCD spectra indicate that the band structure of (GaMn)As is similar to that of zinc-blende type semiconductors. Furthermore, we have successfully grown superlattices consisting of magnetic (GaMnAs)/nonmagnetic (AlAs) III–V semiconductors. Structural characterizations showed that they have abrupt interfaces with good crystalline quality. The SLs with wider (GaMn)As quantum wells ($\geq$70 Å) are ferromagnetic at low temperature, while the SLs with narrower quantum wells ($\leq$65 Å) are paramagnetic even at low temperature (2 K).

The MCD spectra have shown clear evidence for the quantum confinement effect and the formation of subbands in these SLs. The ability to grow this new GaAs based magnetic semiconductor and its quantum heterostructures will open up unique possibilities for coupling spin-related phenomena with the well-established band engineering in III–V semiconductors.

ACKNOWLEDGMENTS

The author acknowledges the collaboration and help of T. Hayashi, H. Shimizu, K. Takahashi, and T. Nishinaga (Department of Electronic Engineering, University of Tokyo). He thanks H. Shimada, H. Tsuchiya, and Y. Ootuka (Cryogenic Center, University of Tokyo), K. Ando and R. Shioda (Electrotechnical Laboratory), and H. Hayashi and K. Niihara (Osaka University) for fruitful collaborations and discussions. This work was supported by the PRESTO (Sakigake-21) of Japan Science and Technology Corp. (JST), by the JSPS Research for the Future Program (JSPS-RFTF97P00202), and by a Grant-in-Aid for Scientific Research, "Spin Controlled Semiconductor Nanostructures," from the Ministry of Education, Science, Sports and Culture, Japan. Thanks are also due for the financial support from the Murata Foundation and from Asahi-Glass Foundation.

[1]G. A. Prinz, Science **250**, 1092 (1990).
[2]See, for example, M. Tanaka, Mater. Sci. Eng. B **31**, 117 (1995); M. Tanaka, Physica E (in press).
[3]R. R. Galazka, Posterpy Fiz. **28**, 601 (1977); Inst. Phys. Conf. Ser. **43**, 133 (1978).
[4]*Semiconductors and Semimetals*, edited by J. K. Furdyna and J. Kossut (Academic, Boston, 1988), Vol. 25; J. K. Furdyna, J. Appl. Phys. **64**, R29 (1988).
[5]K. Onodera, T. Matsumoto, and M. Kimura, Electron. Lett. **30**, 1954 (1994).
[6]H. Munekata, H. Ohno, S. von Molnar, A. Segmuller, L. L. Chang, and L. Esaki, Phys. Rev. Lett. **63**, 1849 (1989).
[7]H. Ohno, H. Munekata, T. Penny, S. v. Molnar, and L. L. Chang, Phys. Rev. Lett. **68**, 2664 (1992).
[8]S. Koshihara, A. Oiwa, S. Katsumoto, H. Hirasawa, Y. Iye, and H. Munekata, Phys. Rev. Lett. **78**, 4617 (1997).
[9]J. De Boeck, R. Oesterholt, A. Van Esch, H. Bender, C. Bruynseraede, C. Van Hoof, and G. Borghs, 2nd International Symposium on Metallic Multilayers, Cambridge, UK, September 1995; J. Magn. Magn. Mater. **156**, 148 (1996).
[10]J. De Boeck, R. Oesterholt, A. Van Esch, H. Bender, C. Bruynseraede, C. Van Hoof, and G. Borghs, Appl. Phys. Lett. **68**, 2744 (1996).
[11]T. Hayashi, M. Tanaka, T. Nishinaga, H. Shimada, H. Tsuchiya, and Y. Ootuka, 9th International Conference on Molecular Beam Epitaxy, Malibu, CA, August 1996; J. Cryst. Growth **175/176**, 1063 (1997).
[12]T. Hayashi, M. Tanaka, T. Nishinaga, and H. Shimada, 41st Conference on Magnetism and Magnetic Materials, Atlanta, November 1996; J. Appl. Phys. **81**, 4865 (1997).
[13]Other groups are also studying the properties of (GaMn)As; H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, and Y. Iye, Appl. Phys. Lett. **69**, 363 (1996); A. Shen, H. Ohno, F. Matsukura, Y. Sugawara, N. Akiba, T. Kuroiwa, A. Oiwa, A. Endo, S. Katsumoto, Y. Iye, J. Cryst. Growth **175/176**, 1069 (1997).
[14]M. Tanaka, H. Shimizu, and T. Hayashi (unpublished).
[15]R. Shioda, K. Ando, T. Hayashi, and M. Tanaka, Phys. Rev. B (in press).
[16]L. Berger and G. Bergmann, in *The Hall Effect and Its Applications*, edited by C. L. Chien and C. R. Westgate (Plenum, New York, 1980), p. 55.
[17]K. Ando, M. Tanaka, T. Hayashi, and A. Twardowski, 1997 APS March Meeting, Kansas City, March 1997, J. Appl. Phys. **83**, 6548 (1998).
[18]K. Ando, K. Takahashi, T. Okuda, and M. Umehara, Phys. Rev. B **46**, 12 289 (1992).
[19]J. Szczytko, W. Mac, A. Stachow, A. Twardowski, P. Becla, and J. Tworzydlo, Solid State Commun. **99**, 927 (1996).
[20]H. Munekata, A. Zaslavsky, P. Fumagalli, and R. J. Gambino, Appl. Phys. Lett. **63**, 2929 (1993); P. Fumagalli and H. Munekata, Phys. Rev. B **53**, 15 045 (1996).
[21]A. Shen, H. Ohno, F. Matsukura, Y. Sugawara, Y. Ohno, N. Akiba, and T. Kuroiwa, Jpn. J. Appl. Phys., Part 2 **36**, L73 (1997).
[22]T. Hayashi, M. Tanaka, K. Seto, T. Nishinaga, and K. Ando, Appl. Phys. Lett. **71**, 1825 (1997).

Ferromagnetic resonance imaging of Co films using magnetic resonance force microscopy

B. J. Suh, P. C. Hammel,[a] and Z. Zhang
Condensed Matter and Thermal Physics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

M. M. Midzor and M. L. Roukes
Condensed Matter Physics, California Institute of Technology, Pasadena, California 91125

J. R. Childress
University of Florida, Gainesville, Florida 32611

(Received 21 January 1998; accepted 21 May 1998)

Lateral one-dimensional imaging of cobalt (Co) films by means of microscopic ferromagnetic resonance (FMR) detected using the magnetic resonance force microscope (MRFM) is demonstrated. A novel approach involving scanning a localized magnetic probe is shown to enable FMR imaging in spite of the broad resonance linewidth. We introduce a spatially selective local field by means of a small, magnetically polarized spherical crystallite of yttrium iron garnet (YIG). Using MRFM-detected FMR signals from a sample consisting of two Co films, we can resolve the ~ 20 μm lateral separation between the films. The results can be qualitatively understood by consideration of the calculated spatial profiles of the magnetic field generated by the YIG sphere.
© *1998 American Vacuum Society.* [S0734-211X(98)08204-3]

I. INTRODUCTION

Magnetic field sensors comprised of layered magnetic materials are having a significant impact on magnetic recording technologies. The sensitivity of these layered materials to characteristics of the buried interfaces between layers highlights the need for a high resolution, spatial imaging probe of structural and magnetic properties of materials. The magnetic resonance force microscope (MRFM) can potentially fill this need. MRFM detection of both nuclear magnetic resonance (NMR)[1] and ferromagnetic resonance (FMR)[2] has been demonstrated. Each of these has advantages for microscopic imaging in magnetic materials. FMR benefits from very high signal sensitivity because it couples to fully polarized electronic moments, and conventional FMR has a demonstrated capability for determining crucial magnetic properties such as magnetic anisotropies of the thin films and the magnetic exchange coupling between nearby ferromagnetic layers.[3] However, microscopic FMR imaging cannot be performed using conventional techniques because conventional FMR is performed in a uniform magnetic field so there is no means to identify the spatial origin of a particular contribution to the FMR signal.

Magnetic resonance imaging employs a magnetic field gradient to identify the spatial origin of a resonance signal. Through the magnetic resonance condition ($\omega_0 = \gamma H_0$ for a noninteracting spin having gyromagnetic ratio γ, where H_0 is the applied field) the field gradient allows the spatial origin of the signal to be inferred from the resonance frequency. This assumes that the resonance frequency ω_0 is a *local* function of applied field H, that is, $\omega_0(\mathbf{r}) = f[H(\mathbf{r})]$. Because of strong dipole couplings to neighboring spins in a ferromagnet, the resonance frequency at a particular spatial location is nonlocal, i.e., it is determined by magnetization of neighboring regions in addition to the value of the field applied at that point. Thus, imaging by means of an applied field gradient is not as straightforward as in the case of noninteracting spins, such as occurs in NMR.

Here we address two aspects of this problem. First, we demonstrate an alternative approach to imaging using a spatially localized magnetic field source, and we present a scanning FMR image in a Co film obtained using a small yttrium iron garnet (YIG) grain as the magnetic probe. This approach is similar to that used in magnetic force microscopy (MFM),[4] where only the spin magnetization in the vicinity of the probe tip contributes to the signal. Spatial resolution in this approach is determined by the extent of the field produced by the magnetic probe, and this approach is not expected to yield resolution superior to that of the MFM.[4] However, FMR imaging has the advantage that it can provide microscopic determination of quantities not obtainable through MFM measurements such as the interlayer exchange coupling. Second, the field gradient due to the YIG particle is sufficiently large that we are able explore the affect of an applied field gradient on a ferromagnetic film with broad intrinsic linewidth and show that it can cause spatially separate regions of a contiguous film to resonate at distinct frequencies, thus indicating that imaging by means of an applied gradient is possible.

The MRFM mechanically detects the magnetic resonance signal by sensitively detecting the oscillatory response of a micromechanical resonator.[1,5–7] A small permanent magnet is used to produce a spatially inhomogeneous magnetic field which plays two crucial roles. First it establishes the coupling

$$\mathbf{F} = (\mathbf{m} \cdot \nabla)\mathbf{B} \tag{1}$$

[a]Author to whom correspondence should be addressed; electronic mail: pch@lanl.gov

between the spin magnetization **m** and the mechanical element. Second, it enables imaging as discussed earlier. The irradiation frequency ω_{rf} defines a surface of constant field called the "sensitive slice" in which the magnetic resonance condition is met, that is, in which $\omega_0=\omega_{rf}$; only those electron spins in this slice will couple to the rf field. Modulating the sample magnetization at the resonance frequency of the mechanical resonator drives it into oscillation; this is accomplished by modulating either the rf field intensity or the applied magnetic field, or both. The resultant time-varying force will be due only to the spins within the sensitive slice. The resonant oscillation of the cantilever is detected by means of an optical fiber interferometer. Images are obtained by scanning the sensitive slice throughout the sample.

In principle, the spatial resolution is given by the sensitive slice width Δz which is determined by the intrinsic resonance linewidth ΔH_{lw} and the applied field gradient $\partial H_0/\partial z$:

$$\Delta z \simeq \frac{\Delta H_{lw}}{\partial H_0/\partial z}. \tag{2}$$

An additional requirement is that signal detection sensitivity must be sufficient to observe the signal from the resolved volume. We have demonstrated earlier[2] that the sensitivity of MRFM detected FMR in YIG films is sufficient to enable studies of microscopic volumes. However, the FMR lines remained sharp in spite of the application field gradient sufficiently large that our sensitive slice width [given by Eq. (2)] should have been smaller than the sample size. That is, the gradient should have been sufficient to broaden the line. Clearly then, nonlocal effects due to dipole couplings dominate over the applied field gradient in this case. Similarly, Wago *et al.*[8] found that imaging in YIG by standard means involving an applied field gradient was not successful. This contrasts with successful demonstrations of microscopic MRFM imaging by means of both electron spin resonance (ESR) and NMR.[9,10]

We have recently focused our efforts on studies of Co films[11] whose FMR lines are even broader than YIG, making the task more difficult. Here we explore the requirements for spatial imaging of magnetic properties of Co films using magnetic resonance force microscopy.

II. EXPERIMENTAL DETAILS

A. Sample

A commercially available atomic force microscopy cantilever[12] serves as the resonant mechanical element used to detect the magnetic resonance signal. Two laterally separated Co films were placed on the cantilever by sputter deposition through a mask consisting of two 70 ± 5 μm wide slits separated by 20 ± 5 μm. Because one of the slits only partially overlapped the end of the cantilever the resulting sample geometry, starting at the free end of the cantilever, is approximately [20 μm (Co) | 20 μm (separation) | 70 μm (Co)]. Because the mask was not in direct contact with the cantilever surface, the film edges were not vertical. Each section of Co film is $\approx$600 Å thick and is protected by Ag layers above and below; the vertical profile is then [Si (cantilever) | Ag (35 Å) | Co (600 Å) | Ag (70 Å)].

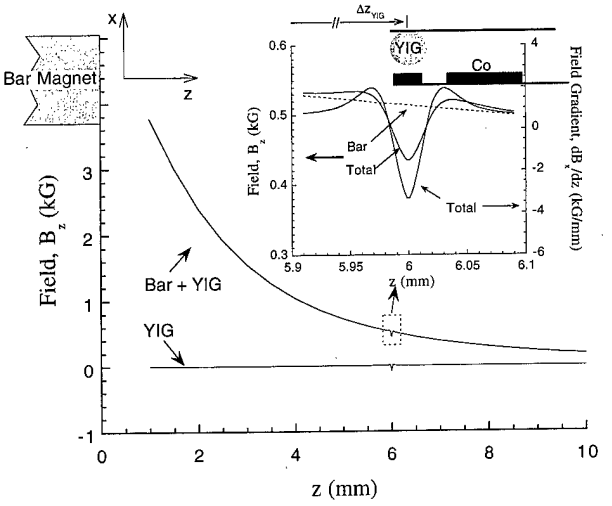


FIG. 1. Calculated field and field gradient profiles for the YIG sphere. A schematic diagram of the relative positions of the YIG sphere and the Co films is shown.

B. Measurements

The experiments were performed with the field applied in the Co film plane which allows small saturation and therefore, resonance fields, of order hundreds of Gauss.[11] A schematic illustration of the arrangement is shown in Fig. 1. The rf irradiation at a frequency, $\omega_{rf}/2\pi\simeq7.9$ GHz, was generated by microstrip resonator.[11,13,14] The cantilever resonance frequency was $f_c\simeq12$ kHz, and its Q value was $\sim10^4$ at 70 mTorr and room temperature.

The sample is positioned slightly off the axis of the bar magnet at $(x,z)\simeq$ (1 and 6 mm) with respect to a point at the center of the near face of the bar magnet (see Fig. 1). The bar magnet is 6.35 mm ($\frac{1}{4}$ in.) long and 6.35 mm ($\frac{1}{4}$ in.) in diameter. The field from an electromagnetic solenoid is scanned from -300 to 300 G.

An approximately spherical YIG grain ≈30 μm in diameter is mounted on a second cantilever and then scanned above the sample with a fixed vertical separation Δx ~30 μm as indicated in Fig. 1. The horizontal position (i.e., the z axis position) of the YIG grain with respect to the Co films is denoted by Δz_{YIG} with respect to an arbitrary reference as illustrated in Fig. 1.

Measurements were performed both by selecting a position for the YIG sphere (Δz_{YIG}), then sweeping the solenoid field, and by scanning the YIG sphere horizontally across the two films (i.e., varying Δz_{YIG}) at a fixed value of the solenoid field.

III. RESULTS AND DISCUSSION

A. YIG sphere magnetic field profile

Figure 1 shows the calculated spatial variation of the magnetic field and the field gradient of the YIG sphere. From Eq. (1) the appropriate expression for our geometry is

$$F_x = m_x \frac{\partial B_x}{\partial x} + m_z \frac{\partial B_x}{\partial z}, \tag{3}$$

where m_α refers to the magnetization of the Co film, and B_α to the various components of the applied field. Since the field due to the solenoid and bar magnet is nearly parallel to $\hat{z}$, that is, in the plane of the Co film, $m_z \gg m_x$, the first term in Eq. (3) can be neglected. The calculated results for B_z and $\partial B_x/\partial z$ are shown in Fig. 1.

YIG was chosen as a probe magnet because it has a small saturation magnetization value, $4\pi M_s = 1.6$ kG which allows it to be easily saturated in small applied fields. Therefore its magnetization does not change in response to variation of the external field due to field sweeping or displacement of the YIG grain with respect to the bar magnet.

We can expect that FMR signals arising from sample regions affected by the field of the YIG sphere will have three characteristics:

(1) As shown in the inset of Fig. 1, the YIG sphere locally reduces the field by ~100 G. Thus, a correspondingly higher applied field will be required to meet the resonance condition, so the resonance spectrum from the region affected by the local field generated by the YIG sphere will appear at higher solenoid field.
(2) The signal strength in a MRFM is enhanced by a larger field gradient.[15,16] The larger magnitude of the field gradient of the YIG sphere will then enhance the size of the signals originating from Co experiencing the field of the sphere.
(3) The sign of the gradient from the YIG sphere is reversed from that of the bar magnet; this will shift the phase of the signals originating from Co near the sphere by π relative to signals from other regions.

B. FMR spectra and scanning image

Figure 2 shows a series of in-phase FMR/MRFM spectra obtained by scanning the solenoid field at several values of Δz_{YIG}. A single FMR signal with a resonance linewidth ≈60 G is observed when the YIG sphere is located far from the sample region. The magnetic field gradient $\partial B_x/\partial z$ due to the bar magnet at the sample is ~0.2 G/μm; this corresponds to a field difference of less than 20 G across the sample. This is smaller than the observed resonance linewidth (≈60 G) so the field gradient is too small to resolve the two films laterally separated by ~20 μm.

When the YIG sphere approaches the sample region, an additional signal at higher field begins to appear. The maximum shift of the additional signal with respect to the original one is ≈170 G. This is larger than the calculated value (~100 G; see the inset of Fig. 1) of the additional field

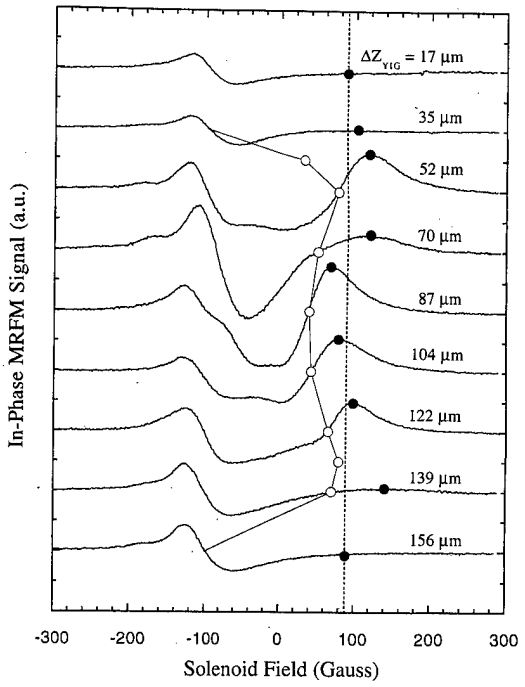


FIG. 2. Series of FMR spectra as a function of solenoid field are shown for several values of Δz_{YIG}: the positions of the maximum deviation (●) and the center of the additional signal (○) are shown. The dotted line indicates the value of the solenoid field at which the signal intensity as a function of Δz_{YIG} (shown as a dotted line in Fig. 3) is determined.

contributed by the YIG sphere. The error in the calculation is ascribed to uncertainties in the size and shape of the YIG grain and the vertical distance of the YIG from the sample (Δx). For Δz_{YIG} near 70 ± 10 μm, the additional signal is suppressed indicating the YIG sphere is between the two films. The intensity of the additional signal recovers when the YIG sphere is moved over the second Co film.

The additional signals generated by the YIG sphere shown in Fig. 2 have the expected features discussed in Sec. III A above, including the characteristic π phase shift. A more detailed analysis of the dependence of the signal intensity and of the spectral shift of the additional signal on Δz_{YIG} provides the basis for this approach to imaging the Co films.

The dependence of the MRFM signal strength $S(\Delta z_{\mathrm{YIG}})$ on the position Δz_{YIG} of the YIG sphere was extracted in two ways. First, $S(\Delta z_{\mathrm{YIG}})$ was determined at a fixed value of the solenoid field, $B = 90$ G (the average value of the peak position of the additional signals) shown as a dotted line in Fig. 2. $S(\Delta z_{\mathrm{YIG}})$ obtained in this way is shown as a dotted curve in Fig. 3. The amplitude of the additional signal is approximately determined by the area of the Co film affected by the selective local field generated by the YIG sphere. Therefore, $S(\Delta z_{\mathrm{YIG}})$ gives the lateral spatial profile of the Co sample. Two regions are clearly distinguished from the dotted curve in Fig. 3: one is ~20 μm wide and the other is ~60 μm wide separated by ~15 μm. This is in reasonable agreement with the actual sample profile, [20 μm |20 μm |70 μm]. The amplitude of the first region is observed to be small compared to the second one. We attribute this to a misreading of the sig-

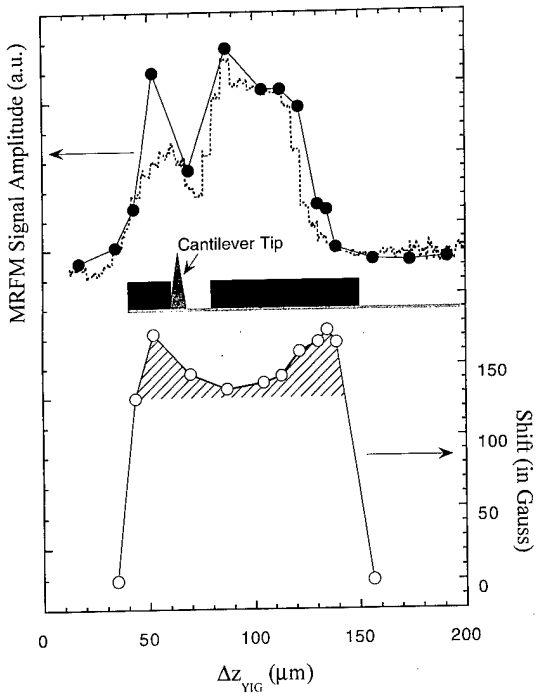


FIG. 3. (Upper) Dotted curve shows the variation of the signal amplitude as a function of the position of the YIG sphere, Δz_{YIG}, at a fixed value of solenoid field $B=90$ G. The curve with solid circles shows the variation of maximum deviation of the additional signal. (Lower) Curve with open circles indicates the magnetic field shift of the additional signal with respect to the position of the original signal. The estimated sample profile obtained from knowledge of the shadow mask dimensions and optical microscopy on the Co sample is illustrated in the middle of the figure for comparison.

nal amplitude resulting from shifts in the field at which the peak occurs. In fact, an alternative approach to determining $S(\Delta z_{YIG})$ in which it is extracted from the peak amplitude of the additional signal (marked as solid circles in Fig. 2) rather than at a constant value of applied field, gives the results shown by solid circles in Fig. 3. Using this approach, the amplitudes of the signals from the two films are more similar.

These shifts in peak position of the additional signal (open circles in Fig. 2) are expected as discussed in Sec. III A. The signal shape[16] from a sample whose dimensions are smaller than the sensitive slice width is a derivative of the magnetic resonance response. Therefore the center resonance spectrum will correspond to the point where this derivative vanishes.[16] Although the gradient is sufficiently large that this is no longer strictly appropriate, we determine the resonance field from the zero crossing of the signal. These results are shown as open circles in Figs. 2 and 3. We expect that any region directly adjacent to the YIG sphere will experience the maximum negative local field and therefore the maximum shift. As the sphere is laterally displaced away from the Co film, the Co sample will experience first the shoulder of the selective local field (see the inset of Fig. 1), giving rise to the smaller shift of the additional signal, then, when no additional field is experienced the shift will vanish. Therefore, the value of the shift also represents the sample profile. However, since the magnitude of the local field is sensitive to the vertical distance of the YIG sphere from the sample due to its small size, the uncertainty in the separation Δx of the YIG sphere from the sample as the sphere is scanned along the sample cantilever can introduce error. The sample profile deduced from the value of the shift as a function of Δz_{YIG} is shown by open circles in Fig. 3. The poorer agreement with the known sample profile is attributed to uncertainty in Δx.

A final point is that the additional signals associated with the YIG sphere in several of the spectra in Fig. 2 are clearly broadened by the application of the large applied field gradient of the YIG sphere. This demonstrates that imaging by means used in conventional magnetic resonance imaging (MRI) is possible in FMR. These results also provide an indication that, even with the modest gradients (<10 G/μm) used here, resolution on the scale of tens of microns can be achieved in FMR imaging with conventional MRI techniques.

IV. SUMMARY

We have presented FMR spectra which demonstrate scanning FMR imaging using the MRFM for the first time. By scanning a selective local field generated by a small YIG sphere, we are able to reconstruct the lateral sample profile along the $\hat{z}$ axis (20 μm | 15 μm | 60 μm), which should be compared with the profile expected on the basis of the mask used for depositing the sample: (20 μm | 20 μm | 70 μm).

Unlike conventional MRI where the resolution is determined by the strength of the applied field gradient [Eq. (2)], this approach has the disadvantage that the resolution is determined by the dimensions of the probe magnet. Although resolution superior to that of MFM is not expected, FMR imaging offers distinct advantages in that magnetic properties such as the interlayer exchange coupling can be microscopically mapped. The resolution using the present approach can be improved significantly by using a smaller magnetic probe and by improved control over the distance separating the probe and sample as the probe is scanned. Finally, we have for the first time, observed a broadening of the FMR linewidth arising from the field gradient of the small magnetic particle. This indicates that conventional magnetic resonance imaging techniques can be applied in FMR imaging. Studies to determine the limits of resolution using conventional MRI techniques for FMR imaging are underway.

ACKNOWLEDGMENTS

Work at Los Alamos National Laboratory was performed under the auspices of the U.S. Department of Energy. One of us (Z.Z.) acknowledges the support of the Center for Nonlinear Studies at Los Alamos National Laboratory.

[1]D. Rugar, O. Züger, S. T. Hoen, C. S. Yannoni, H.-M.Vieth, and R. D. Kendrick, Science **264**, 1560 (1994).
[2]Z. Zhang, P. C. Hammel, and P. E. Wigen, Appl. Phys. Lett. **68**, 2005 (1996).
[3]P. E. Wigen, in *Magnetic Multilayers*, edited by L. H. Bennett and R. E. Watson (World Scientific, Singapore, 1994), pp. 183–226.
[4]D. Rugar, H. J. Mamin, P. Guethner, S. E. Lambert, J. E. Stern, I. McFadyen, and T. Yogi, J. Appl. Phys. **68**, 1169 (1990).

[5]J. A. Sidles, Appl. Phys. Lett. **58**, 2854 (1991).
[6]J. A. Sidles, Phys. Rev. Lett. **68**, 1124 (1992).
[7]P. C. Hammel, Z. Zhang, G. J. Moore, and M. L. Roukes, J. Low Temp. Phys. **101**, 59 (1995).
[8]K. Wago, D. Botkin, C. S. Yannoni, and D. Rugar (preprint, 1998).
[9]O. Züger and D. Rugar, J. Appl. Phys. **75**, 6211 (1994).
[10]O. Züger, S. T. Hoen, C. S. Yannoni, and D. Rugar, J. Appl. Phys. **79**, 1881 (1996).
[11]Z. Zhang *et al.* (preprint, 1998).
[12]Digital Instruments, 520 E. Montecito St, Santa Barbara, CA. The experiments reported here use type ESP cantilevers which are 450 μm long, 30–40 μm wide and 1–3 μm thick.
[13]W. J. Wallace and R. H. Silsbee, Rev. Sci. Instrum. **62**, 1754 (1991).
[14]K. Wago, O. Züger, J. Wegener, R. Kendrick, C. S. Yannoni, and D. Rugar, Rev. Sci. Instrum. **68**, 1823 (1997).
[15]J. A. Sidles and D. Rugar, Phys. Rev. Lett. **70**, 3506 (1993).
[16]Z. Zhang, M. L. Roukes, and P. C. Hammel, J. Appl. Phys. **80**, 6931 (1996).

In situ controlled reactions and phase formation of thin films on GaAs

D. A. Caldwell, L. C. Chen, A. H. Bensaoula, J. K. Farrer, C. B. Carter, and C. J. Palmstrøm[a)]
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

(Received 20 January 1998; accepted 20 May 1998)

In situ sequential depositions and reactions are used to control a reaction sequence first to consume GaAs and then to release GaAs from the reaction product. The first reaction involves annealing *in situ* deposited Ni on molecular beam epitaxially grown GaAs (100) at 300 °C to form a Ni_3GaAs/GaAs structure. Exposure of this structure to As_4 results in a reaction which consumes the Ni_3GaAs by the layer-by-layer formation of NiAs at the surface and epitaxially regrown GaAs at the Ni_3GaAs/GaAs interface. The NiAs formation and GaAs regrowth are controlled by the As_4 flux. Ni diffusion dominates both the Ni_3GaAs formation and decomposition mechanisms. Reflection high-energy electron diffraction, Rutherford backscattering, x-ray diffraction, and transmission electron microscopy data are used to confirm the phase formation and reaction sequences. © *1998 American Vacuum Society.* [S0734-211X(98)08604-1]

I. INTRODUCTION

Reactions of metals with GaAs have been studied in order to investigate reaction kinetics and the stability of metallic contacts to GaAs.[1–4] Metal–Ga–As ternary phase diagrams[5–10] indicate the thermodynamic stability between the different phases in the system as a function of composition. Through the knowledge of reaction kinetics and thermodynamics, techniques for controlling reaction sequences and phase formation can be predicted. The approach adopted here is generic and can be applied to a large number of new materials systems. The example presented is for the Ni–Ga–As system. This system was chosen because the Ni–Ga–As phase diagram[8–10] and Ni/GaAs reactions[3,11–20] have been studied extensively. However, some controversy still exists over the details of the ternary phase diagram.[21,22] The controversy deals with the details of the isothermal section at 600 °C. The phase diagram of Guérin and Guivarc'h[8] appears to be more consistent with thin film reaction studies at lower temperatures.[21,22]

The Ni/GaAs reaction starts at annealing temperatures of less than 200 °C and is very rapid above 200 °C.[19] The reaction follows Eq. (1) and results in the formation of Ni_xGaAs at the Ni/GaAs interface, where x ranges from 2 to 4:[13,15,16,19]

$$x\mathrm{Ni}+\mathrm{GaAs}\rightarrow \mathrm{Ni}_x\mathrm{GaAs}. \tag{1}$$

Ni_xGaAs is a thermodynamically stable phase. However, Ni_xGaAs is not thermodynamically stable in contact with GaAs according to the ternary phase diagram. Furthermore, the ternary phase diagram allows the prediction that annealing of Ni_xGaAs in contact with GaAs should decompose the Ni_xGaAs. The phase diagram by Guérin and Guivarc'h[8] predicts the decomposition into NiGa and NiAs, which is observed for Ni_xGaAs on GaAs annealed at temperatures $\geqslant$400 °C.[3,11,16,23] The ternary phase diagram also indicates that Ni_2Ga_3, and possibly $NiAs_2$, are stable in contact with GaAs. The fact that $NiAs_2$ was not observed in the experimental verification of the ternary phase diagram suggests that $NiAs_2$ itself is not stable at the temperatures and pressures used in these studies. Using the phase diagram with the predicted lack of stability of $NiAs_2$ at the low pressures used in the present study, one would predict that if a Ni_xGaAs/GaAs structure were exposed to As, the resulting thermodynamically stable phases would be NiAs, GaAs, and As. This can occur through the following reaction:

$$x\mathrm{As}+\mathrm{Ni}_x\mathrm{GaAs}\rightarrow x\mathrm{NiAs}+\mathrm{GaAs}. \tag{2}$$

The GaAs formed through reaction (2) is expected to regrow epitaxially on the GaAs substrate by solid phase regrowth in a similar manner to that observed in reactions of bilayer structures deposited on GaAs such as Si/Ni/GaAs,[24] Si/Pd/GaAs,[25] and Al/Ni/GaAs.[26] This solid phase regrowth technique requires interplay between kinetics and thermodynamics in the bilayer structures.[24,27] This interplay is eliminated, however, by utilizing two sequential depositions and reactions, effectively separating kinetics and thermodynamics. With the process adopted here, each reaction is driven to completion enabling control of the penetration depth of the metallic compound into the GaAs and the amount of GaAs regrown. The penetration depth into the GaAs is controlled by the amount of Ni deposited through completion of reaction (1), as shown in Figs. 1(a) and 1(b). The amount of As supplied for reaction (2) controls the amount of regrown GaAs formed, as shown in Fig. 1(c). Arsenic will only stick to a surface at elevated temperatures through a compound formation reaction. Hence, once all of the Ni has reacted to form NiAs, excess As_4 will not stick to the sample. This makes reaction (2) self limiting in this particular case.

II. EXPERIMENTAL PROCEDURE

n^+-doped GaAs (100) substrates were In bonded to Mo sample blocks and loaded into a modified VG Semicon V80H molecular beam epitaxy (MBE) system with a base

[a)]Electronic mail: palms001@gold.tc.umn.edu

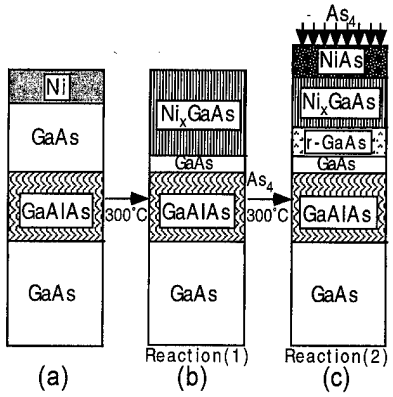


FIG. 1. Schematic of the sequential reaction procedure used for a Ni/GaAs/$Ga_{0.5}Al_{0.5}As$/GaAs structure: (a) as-deposited, (b) reaction (1) induced by 300 °C annealing, and (c) reaction (2) induced by As_4 exposure at 300 °C.

pressure of $\sim 5\times 10^{-11}$ mbar. For each substrate, after native oxide removal and annealing at 640 °C under an As_4 overpressure, a 0.65 μm thick buffer layer of GaAs was grown at 580 °C, followed by a 1000 Å thick $Ga_{0.5}Al_{0.5}As$ marker layer and an additional 1000 Å thick GaAs layer. The as-grown structure was annealed at 580 °C for 10 min in the growth chamber under an As_4 flux to form a smooth surface, then it was cooled until the sample temperature reached 350 °C, at which point the As_4 flux was shut off. This process results in the formation of an As-rich $c(4\times 4)$ surface reconstruction on the as-grown GaAs surface. The samples were transported in ultrahigh vacuum (UHV) to an integral chamber with electron beam evaporation sources for 500 Å Ni deposition at a rate of about 2 Å/s at room temperature. After Ni deposition, the samples were transported back to the MBE growth chamber under UHV for subsequent annealing and reactions at 300 °C. This enabled reflection high-energy diffraction (RHEED) studies to be performed during the annealing and reactions. A constant heater output power was used in order to ensure no temperature overshoot during the annealing. This resulted in a slow ramp rate. The annealing took place at a growth chamber pressure of $<2\times 10^{-10}$ mbar for 10 min at 300 °C prior to As_4 exposure. The As_4 flux used to induce reaction (2) was supplied from a valved cracking effusion cell. The use of a valved cracker allowed for controlled As_4 exposure times with negligible background As_4 pressure in the chamber. As-RHEED oscillations observed from As_4-induced GaAs growth from predeposited Ga on GaAs were used to determine the As_4 flux. The flux resulted in a GaAs growth rate of ~ 2 bilayers/s at 580 °C, which corresponds to an effective As incorporation flux of $\sim 1.2\times 10^{15}$ As atoms $cm^{-2}\,s^{-1}$.

Structural analysis *ex situ* included Rutherford backscattering spectrometry (RBS), x-ray diffraction (XRD), and cross-sectional transmission electron microscopy (TEM). Rotational random RBS spectra were obtained using a 2.3 MeV He^+ ion beam, a scattering angle of 160°, and a sample tilt angle of 5°. Simulations of the RBS spectra were performed using the Cornell (RUMP) RBS analysis program.[28] A Siemens D5005 wide-angle diffractometer was used for the XRD measurements. TEM analysis was performed on a Philips CM-30 microscope operating at 300 keV.

III. RESULTS AND DISCUSSION

A. Reaction (1): Ni/GaAs

After deposition a streaky RHEED pattern was observed from the Ni films with the electron beam incident along the ⟨010⟩ direction of GaAs, indicating a highly oriented or epitaxial Ni film. The streaky RHEED pattern started disappearing at 240 °C while heating the Ni/GaAs structure from room temperature to 300 °C. A ringed polycrystalline pattern appeared at 245 °C and finally another streaky pattern appeared between 250 and 260 °C. This pattern did not change upon further annealing at 300 °C for over 30 min.

Figure 2 shows the RBS spectra from as-deposited, annealed, and reacted samples. The RBS spectra in Figs. 2(a) and 2(b) correspond to the as-deposited and 300 °C annealed Ni/GaAs structures, respectively. There is a dramatic difference between the two spectra. The inset in each figure corresponds to the structure used for the simulation of the RBS spectrum. The simulated spectra have been deconvoluted to show the contributions from each element. The dip in the simulated Ga signal at about channel 310 corresponds to the Ga signal from the $Ga_{0.5}Al_{0.5}As$ layer and the peak in the simulated Ga signal at about channel 325 corresponds to the Ga signal from the GaAs layer on top of $Ga_{0.5}Al_{0.5}As$. An excellent fit to the as-deposited sample spectrum [Fig. 2(a)] was achieved with a thin layer of Ni_3GaAs between the Ni and GaAs. This reacted layer may have formed during the Ni deposition or more likely during the heating ($\geq$160 °C) required to melt the In during sample removal from the Mo sample holder. The spectrum in Fig. 2(b) shows that the Ni has completely reacted with the GaAs to form Ni_3GaAs during the *in situ* 300 °C 10 min anneal. The GaAs layer beneath this reacted layer is clearly much thinner as is evident from the reduced simulated Ga signal in the peak at about channel 325.

Results from XRD studies confirm the interpretation of the RBS data. Figure 3 shows the XRD patterns obtained from the as-deposited, annealed, and reacted samples. For the as-deposited sample (200) Ni and the (200) and (400) GaAs substrate peaks are observed [Fig. 3(a)]. This indicates that the Ni grows with a preferred orientation of (100) Ni‖(100) GaAs, which is consistent with the RHEED data. The diffraction pattern from the 300 °C annealed sample [Fig. 3(b)] shows peaks corresponding to $(10\bar{1}1)$ and $(20\bar{2}2)$ Ni_xGaAs in addition to the substrate peaks. No elemental Ni peaks are observed, indicating complete Ni_xGaAs formation. A cross-sectional TEM micrograph of the annealed sample is shown in Fig. 4(a). It shows a fine vertical-grained Ni_3GaAs layer $\sim$1000 Å thick above a thin ($\sim$300 Å) unreacted GaAs layer over the $\sim$1000 Å thick $Ga_{0.5}Al_{0.5}As$ marker layer. The selected area diffraction pattern from the Ni_3GaAs corresponds to diffraction patterns obtained along the $\langle 2\bar{1}\bar{1}0\rangle$ zone

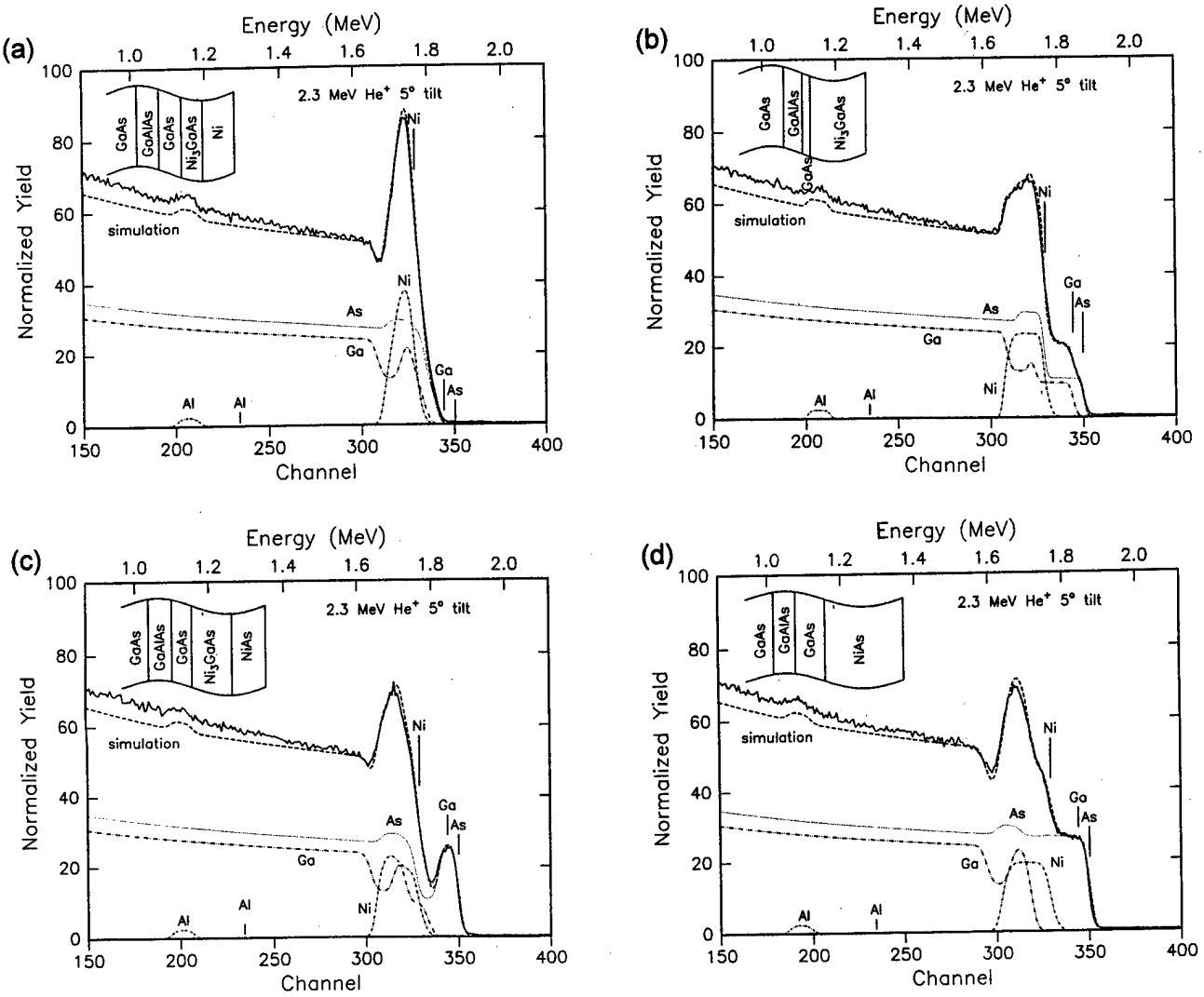


FIG. 2. RBS spectra of Ni/GaAs/$Ga_{0.5}Al_{0.5}As$/GaAs structures: (a) as-deposited, and annealed at 300 °C for 10 min followed by As_4 exposure times of (b) 0, (c) 3, and (d) 10 min. The simulated spectra (dashed lines) and deconvolutions of the simulated spectra indicating the contribution from each element are also shown.

axis of two variants. The observed orientation relationship $(10\bar{1}1)\langle 2\bar{1}\bar{1}0\rangle$ $Ni_3GaAs\|(100)\langle 011\rangle$ GaAs is consistent with that reported by Sands *et al.*[15]

The composition of Ni_xGaAs has been reported to range from $x=2$ to 4.[13,15,16,19] Simulations of the RBS spectra in Fig. 2 indicate a value of 3 for x. The Ni_xGaAs lattice parameters as a function of x have been reported by Chen *et al.*[19] using electron diffraction. Their data would correspond to $d_{(10\bar{1}1)}=2.75$ Å when $x=2.5$ and $d_{(10\bar{1}1)}=2.81$ Å when $x=4$. Assuming a linear interpolation, the value of $d_{(10\bar{1}1)}=2.77$ Å determined from the XRD data in Fig. 3(b) would also be consistent with an x value of 3. The lattice parameters of $a=3.835$ Å and $c=5.032$ Å ($c/a=1.31$) were determined from TEM selective area diffraction patterns and are consistent with the XRD data. These data are in excellent agreement with the values reported by Lahav *et al.*[16] ($a=3.83$ Å, $c=5.04$ Å, $c/a=1.316$) for a 300 °C annealed sample, which was reported to have a composition of Ni_2GaAs. A detailed study of the many different Ni_xGaAs phases has been made by Guérin and Guivarc'h.[8] The lattice parameters reported here are intermediate between the Guérin and Guivarc'h[8] reported values for the B and A Ni_xGaAs phases, which both have x values close to 3. Sands *et al.*[15] have measured the lattice parameters of Ni_xGaAs formed by thin film reactions as a function of annealing conditions. The composition of Ni_xGaAs can also be determined from the micrograph in Fig. 4(a). 500 Å of Ni (4.6×10^{17} atoms/cm^2) would consume 690 Å (3.05×10^{17} atoms/cm^2) GaAs to form Ni_3GaAs. As the initial GaAs layer was 1000 Å thick, this would leave 310 Å unreacted GaAs, as is observed in the micrograph.

Based on the changes in the RHEED patterns during annealing, it appeared that the Ni/GaAs reaction [reaction (1)] proceeded rapidly at temperatures $\leq$250 °C and was completed before the sample reached 300 °C. This is consistent with the reported Ni/GaAs reaction kinetics.[19] Ni has been reported to be the dominant diffusing species.[16,19,20] Palmstrøm *et al.*[29] pointed out in the Co/GaAs system that

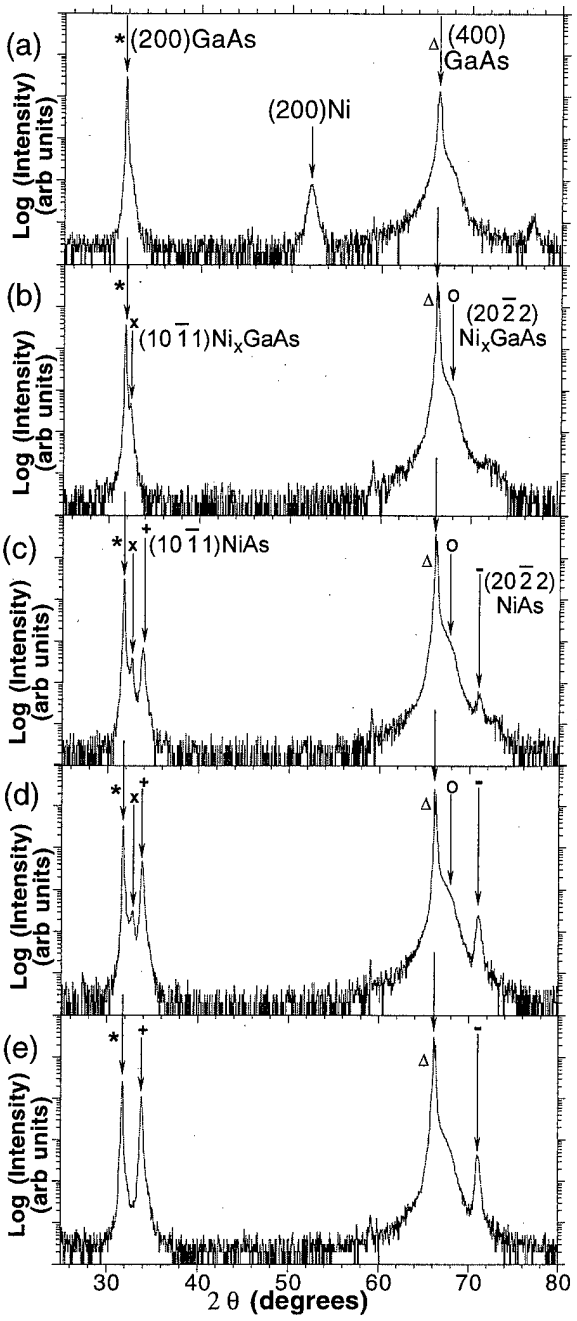


FIG. 3. X-ray diffraction spectra of Ni/GaAs/$Ga_{0.5}Al_{0.5}As$/GaAs structures: (a) as-deposited and annealed at 300 °C for 10 min followed by As_4 exposure times of (b) 0, (c) 1, (d) 3, and (e) 10 min. (a) shows a (200)Ni peak in addition to (200)GaAs (*) and (400)GaAs (△) peaks. Annealing results in Ni_xGaAs [×=(10$\bar{1}$1) and o=(20$\bar{2}$2)] formation and with subsequent As_4 exposure the Ni_xGaAs peak intensities decrease and the NiAs [+=(10$\bar{1}$1) and −=(20$\bar{2}$2)] increase.

Co_2GaAs formation results from Co indiffusion into the GaAs and that CoAs and CoGa can only form when there is diffusion of Ga and As. The formation of Ni_3GaAs without NiGa or NiAs formation results from Ni indiffusion into the GaAs with Ga and As being essentially immobile. This explains why Ni_3GaAs is formed at these low temperatures rather than NiGa and NiAs, which the Ni–Ga–As ternary phase diagram predicts.

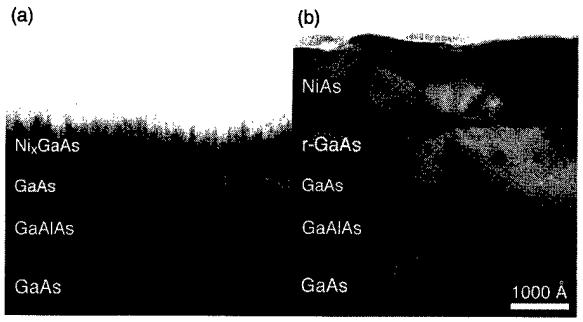


FIG. 4. Cross-sectional TEM micrographs of Ni/GaAs/$Ga_{0.5}Al_{0.5}As$/GaAs structures annealed to 300 °C for 10 min (a) without subsequent As_4 exposure and (b) with 10 min of subsequent As_4 exposure at 300 °C. (The rough GaAs/$Ga_{0.5}Al_{0.5}As$ interface is a result of a rough $Ga_{0.5}Al_{0.5}As$ surface during MBE growth of this layer.)

B. Reaction (2): As_4 reactions with Ni_3GaAs/GaAs

Reaction (2) was induced by exposing the Ni_3GaAs/GaAs structure formed by reaction (1) to an As_4 flux while the sample was held at 300 °C. With the electron beam incident along the GaAs ⟨010⟩, the streaky RHEED pattern from the Ni_3GaAs changed abruptly to a more spotty pattern, indicating diffraction from a rough surface. The diffraction pattern kept changing during the As_4 exposure. Within 1 min of exposure, the RHEED pattern became more streaky and after 2 min additional faint streaks, possibly due to a 3× surface reconstruction, were observed. After 6.5 min, the RHEED pattern consisted of streaks with chevrons, indicating a faceted surface. Further exposure to As_4 did not result in a change in the RHEED pattern. This indicates that reaction (2) was completed after 6.5 min of As_4 exposure.

The Ni_xGaAs XRD peaks decrease and the NiAs peaks increase with As_4 exposure (Fig. 3). Figure 3(e) has no diffraction peaks for Ni_3GaAs, showing that reaction (2) is complete after 10 min of As_4 exposure. These x-ray diffraction data also show that NiAs grows with the same perpendicular orientation as Ni_3GaAs, i.e., (10$\bar{1}$1) NiAs‖(100) GaAs. Further, the lattice parameters of the Ni_3GaAs change slightly with As_4 exposure, $d_{(10\bar{1}1)}$=2.750 and 2.737 Å after 1 and 3 min exposure, respectively, indicating a possible change in composition. However, as GaAs regrowth is observed, compositional variation in the Ni_3GaAs is unlikely to involve significant changes in the Ga to As ratio, but may involve a change in the Ni content of the Ni_3GaAs. Figures 2(c) and 2(d) show the RBS spectra obtained for Ni_3GaAs/GaAs structures after exposure to a constant As_4 flux for 3 and 10 min, respectively. Figure 2(c) clearly shows that NiAs forms on the Ni_3GaAs surface and that GaAs regrows at the Ni_3GaAs/GaAs interface. This is evident from the deconvoluted simulated spectrum. The large peak at about channel 345 is due to the As in the NiAs surface layer and the peak in the Ga simulated signal at channel 320 is due to the GaAs layer on top of the $Ga_{0.5}Al_{0.5}As$. The fact that this peak is larger than that observed for the Ni_3GaAs/GaAs structure prior to As_4 exposure demonstrates that regrowth of GaAs has occurred. Comparing the size and position of this

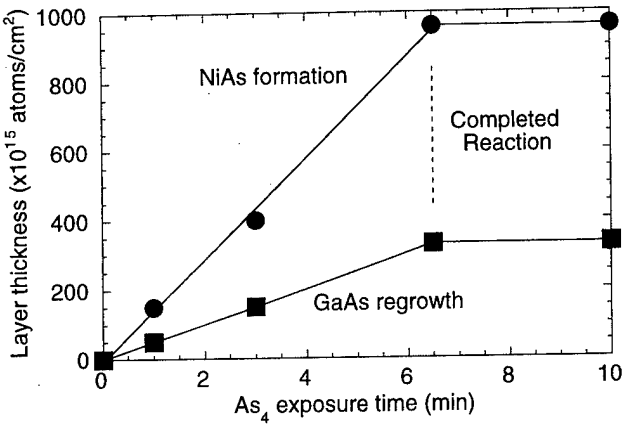


FIG. 5. Plot of thickness of reaction products as a function of As_4 exposure time at 300 °C for $Ni_xGaAs/GaAs$ structures. A linear relationship is found between thickness and exposure time, implying that reaction (2) is limited by arrival of As_4 at the sample surface.

Ga peak in Figs. 2(b), 2(c), and 2(d) indicates that it grows in size and moves to lower energy (channel number) as a result of As_4 exposure. The Al peak can also be seen to move to lower energy as a result of the As_4 exposure. The increase in size of the Ga peak corresponds to an increase in the thickness of the GaAs layer on top of the $Ga_{0.5}Al_{0.5}As$ (i.e., GaAs regrowth). The shift in energy of the Ga and Al peaks results from the increase in thickness of material above the GaAs layer due to the incorporation of As from the As_4 flux into the NiAs layer at the surface. The simulated spectrum for Fig. 2(d) corresponds to the completion of reaction (2) with the formation of a $NiAs/GaAs/Ga_{0.5}Al_{0.5}As/GaAs$ structure, which is confirmed by the cross-sectional TEM of the same sample shown in Fig. 4(b). The layered structure of NiAs/regrown $GaAs/GaAs/Ga_{0.5}Al_{0.5}As/GaAs$ can easily be seen. The grain structures of Ni_3GaAs and NiAs are distinctly different with grains of NiAs being much larger and less elongated (Fig. 4). The regrown GaAs layer is epitaxial. However, it has some inclusions and stacking faults. The inclusions are probably NiAs. NiAs precipitates have been reported to form in GaAs regrown using Si/Ni/GaAs structures.[24] The stacking faults are likely to have nucleated on inclusion. The rough $GaAs/Ga_{0.5}Al_{0.5}As$ interface is a result of a rough $Ga_{0.5}Al_{0.5}As$ surface during the MBE growth of this layer.

The reaction rate for the NiAs formation and the GaAs regrowth can be determined from RBS data. Figure 5 shows a plot of the NiAs and regrown GaAs thicknesses (in atoms/cm^2) as a function of the As_4 exposure time. A linear increase in both the NiAs and GaAs is observed up to 6.5 min exposure, after which no further increase is observed. No further increase results from reaction (2) being completed as no further Ni is available for NiAs formation. The linear dependence on time indicates reaction limited growth. A fit to the NiAs linear growth rate corresponds to an As incorporation rate of 1.2×10^{15} atoms cm^{-2} s^{-1}, which is in excellent agreement with the As incorporation flux measured by As-induced RHEED oscillations during GaAs growth. This demonstrates that the arrival rate of As is the limiting factor in the reaction rate. The detailed mechanism involved in the reaction is not known. Either the Ni is diffusing out from the Ni_3GaAs, resulting in GaAs regrowth or As is diffusing in and displacing the GaAs. Chen *et al.*[19] estimated the Ni interdiffusion coefficient in Ni_xGaAs from lateral Ni–GaAs diffusion couples to be 1.47×10^{-11} cm^2/s at 300 °C, which would result in a characteristic diffusion length ($\sqrt{Dt}$) of 7500 Å for a 6.5 min anneal. This characteristic diffusion length is much larger than the Ni_3GaAs layer thickness (~1000 Å). This suggests that the rate of supply of Ni to the surface as a result of diffusion out of the Ni_3GaAs is not the limiting factor for the NiAs formation rate. Whether the NiAs growth occurs at the surface by Ni outdiffusion or at the $NiAs/Ni_3GaAs$ interface by As diffusion through the NiAs has yet to be determined.

Possible models for the reaction mechanisms in reactions (1) and (2) can be suggested. In reaction (1) the Ni_3GaAs forms as a result of Ni indiffusion with Ga and As remaining essentially immobile at 300 °C. When the surface of Ni_3GaAs is exposed to As_4, the chemical potential for Ni is lower in NiAs than in Ni_3GaAs [the driving force for reaction (2)] and Ni diffuses out of the Ni_3GaAs to form NiAs on the surface. As the Ga and As atoms have essentially remained immobile during the process, GaAs regrows as the Ni leaves the Ni_3GaAs. One may think of the GaAs acting as a sponge for the Ni.

The model can also explain the precipitate formation in the regrown GaAs. By comparing the micrographs in Figs. 4(a) and 4(b), the spacing of the small vertical lines of inclusion in the regrown GaAs [Fig. 4(b)] can be seen to correspond to the grain boundary spacing of the Ni_3GaAs prior to the GaAs regrowth reaction [Fig. 4(a)]. Hence, a reasonable argument can be made that these precipitates formed in the grain boundaries of the Ni_3GaAs prior to the regrowth reaction (2). Thermodynamics dictates that there is a driving force for Ni_3GaAs decomposition into NiGa and NiAs.[8] Guivarc'h *et al.*[23] observed that the initial decomposition involves the formation of two ternary $Ni_xGa_{1-y}As_y$ phases, one richer in Ga and the other richer in As, which they label C and A, respectively. This decomposition can only occur through Ga or As diffusion, making precipitate formation easier at grain boundaries than in the bulk. The long annealing time (10 min) at 300 °C prior to As_4 exposure would enhance Ni_3GaAs decomposition. If NiAs precipitates have formed, there is no driving force for the Ni to diffuse out of the precipitate to the surface in order to react with the As_4 flux as it already is in the form of NiAs. As the activation energies for As and Ga interdiffusion in Ni_xGaAs are larger than for Ni,[20] lower temperature would decrease the relative diffusion of Ga and As compared to Ni. Hence, shorter annealing times and lower annealing temperatures should result in a dramatic decrease in the precipitate density. Using the Ni interdiffusion coefficient reported by Chen *et al.*,[19] a 20 s anneal should be sufficient to complete reaction (1) at 300 °C and 180 s at 200 °C.

Thermodynamics suggests that Ga deposition on the

Ni_3GaAs/GaAs structure should also induce a GaAs regrowth reaction. Preliminary results indicate that Ga can be used instead of As_4 for the regrowth reaction (2). In the Ga case the Ni_3GaAs reacts with Ga to form Ni_2Ga_3 on the surface and regrown GaAs at the Ni_3GaAs/GaAs interface.

IV. CONCLUSIONS

This study shows that through the knowledge of thermodynamics and reaction kinetics, techniques for controlling reaction sequences and phase formation can be predicted. This was demonstrated by the use of two sequential reaction sequences. The first reaction (1) involved Ni/GaAs and resulted in a reaction product of Ni_3GaAs. The second reaction (2), induced by exposure of the reaction product of the first reaction to As_4, resulted in two reaction products, one of which was one of the reactants in the first reaction (GaAs). These reaction sequences resulted in the Ni being driven into GaAs and then pulled back out again. The consumption of GaAs in reaction (1) is controlled by the thickness of Ni deposited and the amount of regrown GaAs in reaction (2) is controlled by the As_4 exposure. At 300 °C, the rate of reaction (2) was found to be limited by the As_4 arrival rate. During this reaction, NiAs forms at the surface and GaAs regrows epitaxially on the GaAs at the Ni_3GaAs/GaAs interface. The proposed reaction mechanism for both reactions involves Ni being the dominant diffusing species with minimal Ga or As diffusion. Improvements in the quality of regrown GaAs are expected utilizing shorter annealing times and lower temperatures to minimize Ga and As diffusion.

ACKNOWLEDGMENT

This research was supported by the U.S. Air Force Office of Scientific Research, AFOSR Contract No. AF/F49620-95-1-0360.

[1]A. Piotrowska, A. Guivarc'h, and G. Pelous, Solid-State Electron. **26**, 179 (1983).
[2]C. J. Palmstrøm and D. V. Morgan, in *Gallium Arsenide Materials, Devices, and Circuits*, edited by M. J. Howes and D. V. Morgan (Wiley, Chichester, 1985), p. 195.
[3]T. Sands, V. G. Keramidas, A. J. Yu, K.-M. Yu, R. Gronsky, and J. Washburn, J. Mater. Res. **2**, 262 (1987).
[4]T. Sands, Mater. Sci. Eng., B **1**, 289 (1989).
[5]T. Sands, J. Met. **38**, 31 (1986).
[6]R. Beyers, K. B. Kim, and R. Sinclair, J. Appl. Phys. **61**, 2195 (1987).
[7]R. Schmid-Fetzer, J. Electron. Mater. **17**, 193 (1988).
[8]R. Guérin and A. Guivarc'h, J. Appl. Phys. **66**, 2122 (1989).
[9]X. Y. Zheng, J. C. Lin, D. Swenson, K. C. Hsieh, and Y. A. Chang, Mater. Sci. Eng., B **5**, 63 (1989).
[10]D. B. Ingerly, D. Swenson, C.-H. Jan, and Y. A. Chang, J. Appl. Phys. **80**, 543 (1996).
[11]M. Ogawa, Thin Solid Films **70**, 181 (1980).
[12]T. G. Finstad and J. S. Johannessen, Proceedings of the 10th Nordic Semiconductor Meeting, Elsinore, Denmark, 9–11 June 1982, I4:1.
[13]S. H. Chen, C. B. Carter, C. J. Palmstrøm, and T. Ohashi, Mater. Res. Soc. Symp. Proc. **54**, 361 (1986).
[14]S. H. Chen, C. B. Carter, C. J. Palmstrøm, and T. Ohashi, Appl. Phys. Lett. **48**, 803 (1986).
[15]T. Sands, V. G. Keramidas, J. Washburn, and R. Gronsky, Appl. Phys. Lett. **48**, 402 (1986).
[16]A. Lahav, M. Eizenberg, and Y. Komem, J. Appl. Phys. **60**, 991 (1986).
[17]D. J. Chadi, J. Vac. Sci. Technol. A **5**, 834 (1987).
[18]T. Sands, Appl. Phys. Lett. **52**, 197 (1988).
[19]S. H. Chen, C. B. Carter, and C. J. Palmstrøm, J. Mater. Res. **3**, 1385 (1988).
[20]C. H. Jan, D. Swenson, X. Y. Zheng, J. C. Lin, and Y. A. Chang, Acta Metall. Mater. **39**, 303 (1991).
[21]R. Guerin and A. Guivarc'h, J. Appl. Phys. **82**, 493 (1997).
[22]D. B. Ingerly, D. Swenson, C. H. Jan, and Y. A. Chang, J. Appl. Phys. **82**, 496 (1997).
[23]A. Guivarc'h, R. Guérin, J. Caulet, A. Poudoulec, and J. Fontenille, J. Appl. Phys. **66**, 2129 (1989).
[24]T. Sands, E. D. Marshall, and L. C. Wang, J. Mater. Res. **3**, 914 (1988).
[25]L. C. Wang, B. Zhang, F. Fang, E. D. Marshall, S. S. Lau, T. Sands, and T. Kuech, J. Mater. Res. **3**, 922 (1988).
[26]T. Sands, J. P. Harbison, W. K. Chan, S. A. Schwarz, C. C. Chang, C. J. Palmstrøm, and V. G. Keramidas, Appl. Phys. Lett. **52**, 1216 (1988).
[27]L. C. Wang, Mater. Res. Soc. Symp. Proc. **319**, 93 (1994).
[28]L. R. Doolittle, Nucl. Instrum. Methods Phys. Res. B **15**, 227 (1986).
[29]C. J. Palmstrøm, C. C. Chang, A. Yu, G. J. Galvin, and J. W. Mayer, J. Appl. Phys. **62**, 3755 (1987).

Metal/GaN Schottky barriers characterized by ballistic-electron-emission microscopy and spectroscopy

L. D. Bell and R. P. Smith
Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109

B. T. McDermott, E. R. Gertner, R. Pittman, R. L. Pierson, and G. J. Sullivan
Rockwell Science Center, 1049 Camino Dos Rios, Thousand Oaks, California 91360

(Received 20 January 1998; accepted 20 May 1998)

Ballistic-electron-emission microscopy (BEEM) and spectroscopy have been used to characterize the Pd/GaN and Au/GaN interfaces. BEEM spectra yield a Schottky barrier height for Au/GaN of ~1.05 eV that agrees well with the highest values measured by conventional methods. For both Pd and Au, a second threshold is observed in the spectra at about 0.2–0.3 V above the first threshold. Imaging of these metal/GaN interfaces reveals transmission in nearly all areas, although the magnitude is small and spatially varies. Attempts to perform BEEM measurements on other GaN material have resulted in no detectable transmission in any areas, even at voltages as high as 3.5 V.
© *1998 American Vacuum Society.* [S0734-211X(98)08904-5]

I. INTRODUCTION

Recent progress in the growth of GaN and related nitrides has stimulated interest in the use of these materials for short-wavelength light-emitting diodes (LEDs), high-electron-mobility transistors (HEMTs), and metal–semiconductor field-effect transistors (MESFETs). Moreover, GaN exhibits stability at high temperatures and a general unreactivity in caustic environments. However, the lack of a suitable lattice-matched substrate for epitaxial growth has hindered the achievement of low defect densities in GaN. Shrinking device dimensions also emphasize as never before the role of defects in device uniformity. As one example, optimization and reproducibility of device characteristics requires stable, high-quality Schottky contacts to these materials. Unfortunately, electronic properties of interfaces to GaN and its alloys are still not well understood, and macroscopic electrical characterization is hindered by the presence of defects.

Ballistic-electron-emission microscopy[1] (BEEM) was developed from scanning tunneling microscopy[2] (STM) as a high-resolution, nondestructive probe of buried interfaces and hot-carrier transport. The high spatial resolution of BEEM derives from the high spatial resolution inherent in STM and allows the characterization of interfaces on a nanometer scale. An energy spectroscopy of interface electronic properties may be performed, and these features may be imaged directly by scanning the STM tip. BEEM provides a local measurement of Schottky barrier height and is also sensitive to other details of interface band structure.[1] This article describes the application of BEEM spectroscopy and imaging to a characterization of metal/semiconductor Schottky contacts to GaN.

II. EXPERIMENTAL PROCEDURES

The GaN layer used in this work was grown by metalorganic chemical vapor deposition (MOCVD) on a (0001) oriented sapphire substrate using trimethyl gallium and ammonia as source materials and disilane to supply the donor dopant. A low temperature nucleation layer was first grown, followed by a higher temperature GaN layer. The GaN growth temperature was 1040 °C as measured by optical pyrometer. The 1 h growth at a pressure of 150 T produced a 2.2 μm thick GaN layer. Room temperature carrier concentration was 1.2×10^{17} cm^{-3} as determined by Hall measurement, and measured electron mobility was 534 cm^2/V s. The measured full width at half maximum of the (004) x-ray peak was 284 arcsec. Ti/Al ohmic contacts were processed by rapid thermal annealing in a nitrogen environment at 900 °C for 90 s.

Samples were divided into 4 mm squares. Metallization was preceded by transfer into a nitrogen-purged glove-box for chemical cleaning, where the sample was spin-etched[3] using 1:10 HCl:ethanol. For the case of Au contacts, the sample was then directly transferred into the load-lock attached to the evaporation chamber. The Au layer was deposited at a pressure of 2×10^{-9} T. For Pd contacts, the sample was transferred into a nitrogen-filled container to another glove-box attached to a second evaporation chamber. Pd was deposited at a pressure of approximately 1×10^{-7} T. After metal deposition the sample was moved to the STM, also located in a nitrogen glove-box. BEEM measurements were performed at room temperature unless otherwise stated. Additional sample fabrication and measurement details have been presented elsewhere.[4]

III. RESULTS

The first attempts to perform BEEM measurements on GaN were unsuccessful, initially due to difficulty in forming low-leakage Schottky barriers. A decrease in contact area and an increase in material quality resulted in low-leakage contacts; however, most of the GaN layers did not produce measurable BEEM current (down to a detection level of several fA). Many BEEM samples were made using different surface treatments, including hot aqua regia and HCl, and the

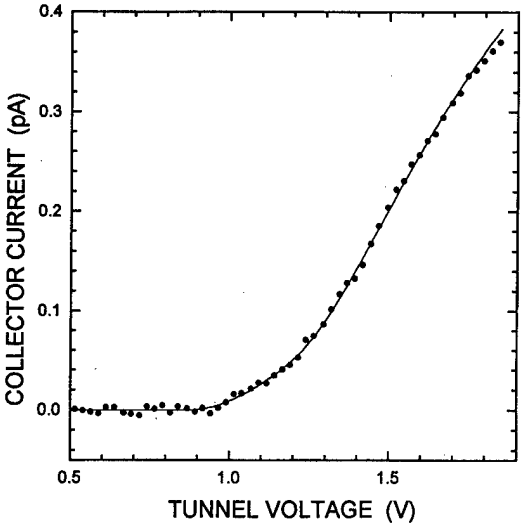


FIG. 1. BEEM I_c-V spectrum (circles) for Pd(2 nm)/GaN, taken at a tunnel current of 2 nA. Also shown is a fit (solid line) to the data using a two-threshold model, which yields thresholds of 0.91 and 1.20 V. Leakage in this sample was larger than average, decreasing the signal-to-noise for I_c measurement.

possibility of a defective surface GaN layer was investigated by etching a fraction of the GaN away using hot KOH. Au, Pd, and Pt were used for metallization. None of these attempts produced measurable collector current, even at tunnel voltages as high as 3.5 V.

All results discussed below were from contacts fabricated on material grown at Rockwell Science Center. Samples made with this material produced a measurable BEEM signal, although the magnitude was extremely small, even using thin (2 nm) metal base layers. Average observed collector current was on the order of 0.5 pA (for a 2 nA tunnel current and 1.6 V tunnel voltage), which is more than two orders of magnitude smaller than theory would predict. A tunnel current of 2 nA was used for BEEM spectroscopy and imaging in order to increase collector current and the signal-to-noise ratio. Due to the extremely low level of BEEM current, many spectra ($\gtrsim$10) were averaged together to increase signal-to-noise.

Figure 1 illustrates such an average for a Pd/GaN sample. Also shown is a fit to the data, using the simple phase-space model.[1] Interestingly, it was necessary to allow two different thresholds in the fit to obtain good agreement with the data. The two-threshold fit to this average yields threshold energies of 0.91 and 1.20 eV with a total estimated uncertainty of ~0.1 eV. The extremely small signal produced Pd/GaN spectra with noticeable noise, even for averages of large numbers of spectra. The average spectrum in Fig. 1 is typical of all averages obtained for two different Pd/GaN samples. The average Schottky barrier height for all data was approximately 0.95 eV as measured by BEEM, although the accuracy of the threshold determination was limited by the noise level.

Imaging was also performed on these Pd/GaN samples. Figure 2 presents a STM image/BEEM image pair obtained on one Pd/GaN sample. Many image sets have been recorded

FIG. 2. STM/BEEM image pair for a Pd(2 nm)/GaN sample. The topograph of the Au surface was obtained at $V=0.5$ V, $I_t=1$ nA, and the BEEM image was recorded at $V=1.8$ V, $I_t=2$ nA. I_c ranges from about 0 to 0.5 pA. Small areas of contamination provide a calibration in the BEEM image for zero current. Except for these areas, current is observed in nearly all areas. Imaged area is 196×150 nm.

on each sample fabricated, and all show similar behavior. Collector current is measured in nearly all areas, although the magnitude varies from 0 to about 0.5 pA (for a 2 nA tunnel current). Small areas of contamination produced areas in the BEEM image where no collector current is measured, providing a visual calibration for zero current, but nearly all other areas show measurable transmission.

In order to test whether samples using other metals exhibited the same small transmission, Au/GaN structures were also investigated. Preliminary characterization by conventional electrical measurement was first performed. Figure 3 shows two current–voltage ($I-V$) curves taken on two different Au/GaN samples, both fabricated on the Rockwell material. Despite comparable ideality factors for the two curves, a different barrier height is extracted from each. This irreproducibility likely results from defects, and illustrates the difficulty of using macroscopic electrical probes to characterize heterogeneous materials.

Although these $I-V$ measurements indicated the presence of defects, BEEM spectra on both of these samples (and on all measured Au/GaN samples) yielded a reproducible Schottky barrier height between 1.02 and 1.07 eV. Figure 4(a) shows an average of many spectra on one Au/GaN sample. Leakage current was generally somewhat smaller for Au than for Pd, resulting in lower noise in the BEEM spec-

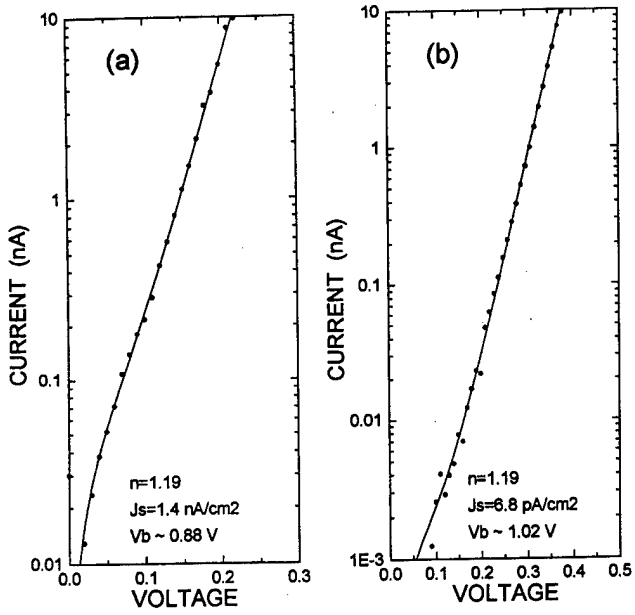


FIG. 3. Conventional macroscopic $I-V$ curves taken on two different Au(5 nm)/GaN samples. GaN material and Schottky contact preparation were identical on the two samples. Extracted barrier height values assumed an effective mass of 0.2 and an effective Richardson constant of $4\pi em^*k^2/h^3$.

tra. The higher data quality allows a more precise assignment of threshold position. For these Au/GaN samples the BEEM spectra also exhibit two thresholds. This is more apparent in Fig. 4(b), where the attempted fit to a model with a single threshold resulted in poor agreement over the entire spectral range.

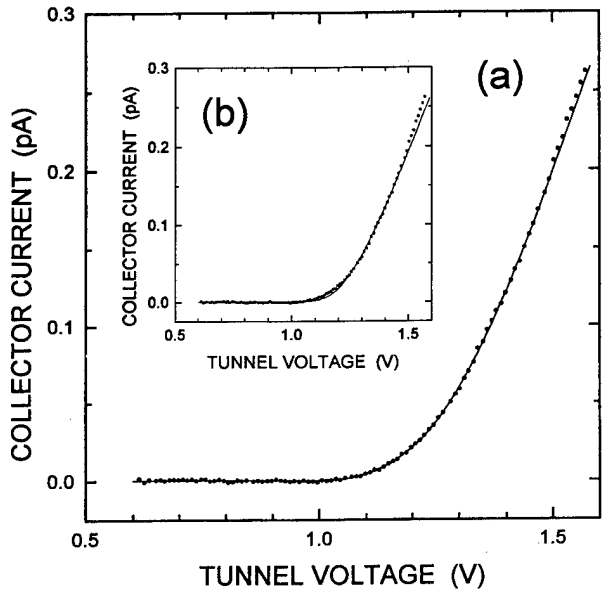


FIG. 4. BEEM I_c-V spectrum (circles) for Au(6 nm)/GaN, taken at a tunnel current of 2 nA. (a) Fit (solid line) to the data using a two-threshold model, which yields thresholds of 1.06 and 1.22 V. (b) Fit to the data assuming only one threshold, yielding a value of 1.12 V. The same data are shown in both (a) and (b).

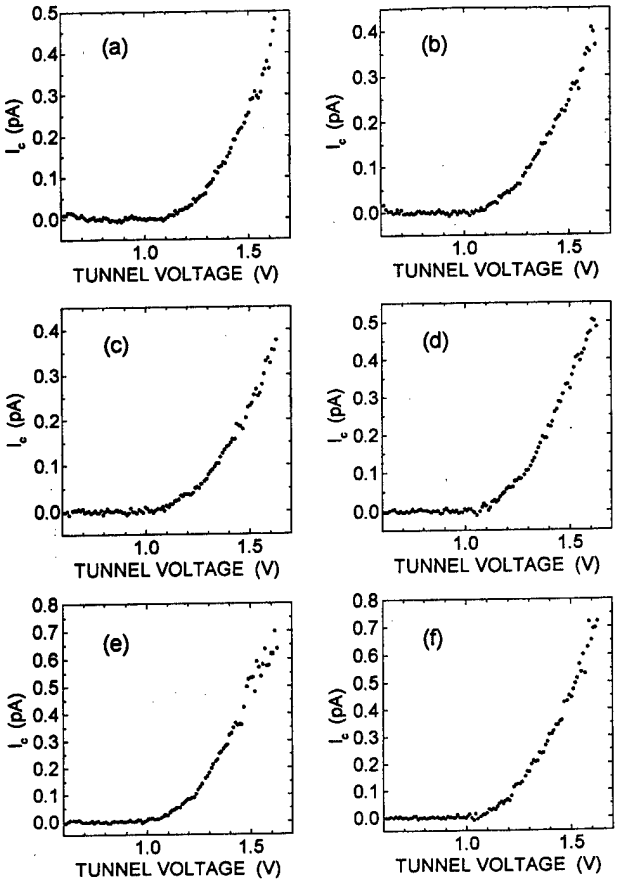


FIG. 5. Six individual I_c-V spectra taken on two different Au/GaN samples at a tunnel current of 2 nA. Although the noise level in these unaveraged spectra is high, two thresholds can usually be observed.

One problem created by the extremely small collector current is the necessity to average many spectra to obtain acceptable signal-to-noise. The capability for resolving the two thresholds is in turn reduced by this averaging process. Although signal-to-noise is improved by averaging, any variation in the energies of the thresholds at different points on the surface will produce a broadening which decreases energy resolution. Spectral averaging also allows the possibility that a two-threshold spectrum is actually a linear combination of two single-threshold spectra, obtained at different locations, and with different threshold energies. In order to rule out this possibility, individual spectra were examined. Six examples are shown in Fig. 5. Although the noise level is higher for individual spectra, it is apparent that the first threshold always occurs around 1 V, and in most cases a second threshold is apparent at approximately 1.2–1.3 V. However, the thresholds do appear to vary somewhat in position, possibly due to variations in strain[5] or to the presence of defects,[6] implying that the signal averaging process is broadening the thresholds to some degree.

Figure 6(a) presents a STM image/BEEM image pair obtained on one Au/GaN sample. Many image sets have been recorded on each sample fabricated, and all show similar behavior. Collector current is measured in nearly all areas, although in this image the magnitude varies from about 0.02

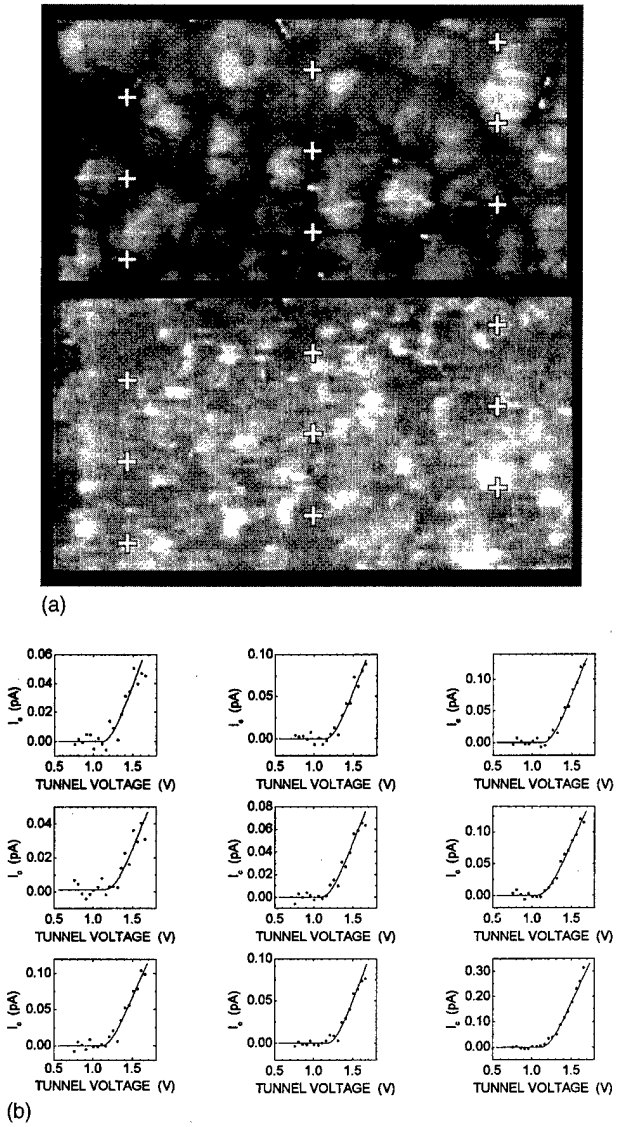


FIG. 6. STM/BEEM image pair for a Au(6 nm)/GaN sample. The topograph of the Au surface was obtained at $V=0.5$ V, $I_t=1$ nA, and the BEEM image was recorded at $V=1.6$ V, $I_t=2$ nA. I_c ranges from about 0.02 to 0.3 pA, and transmission is observed in nearly all areas. Imaged area is 196 $\times$100 nm. Also shown are individual spectra obtained at the positions indicated by the + marks in the images. Here also $I_t=2$ nA. The plotted fits include only a single threshold, and are shown only to guide the eye.

to 0.3 pA (for a 2 nA tunnel current). Individual BEEM spectra were also taken at the locations indicated by + symbols, and these are shown in Fig. 6(b). Collector current in this image is even smaller than for most Au/GaN samples, thus signal-to-noise does not allow a meaningful fit to two thresholds. However, it is apparent that current is observed in all locations, even though the magnitude varies substantially.

IV. DISCUSSION

The extremely small BEEM transmission observed for metal/GaN structures is unusual. GaN is a direct semiconductor, thus the small currents cannot be explained by small coupling to conduction-band minima at large parallel momentum. An insulating or heavily scattering interfacial layer might also attenuate transmission, as might strong scattering in the GaN later itself, if the effect were to backscatter most electrons into the metal layer.[7] Preliminary measurements at 77 K indicate that there is not a strong change in transmission at low temperature, arguing against temperature-dependent mechanisms such as phonon scattering.

Earlier BEEM spectra on Au/GaN by Brazel *et al.*[8] presented somewhat different results. These spectra also displayed two thresholds; however, in that case the measured first threshold was much lower ($\sim$0.7 V), and the second threshold was measured to be at approximately 1.04 V. It is interesting that this second threshold in energy agrees quite well with the first threshold in the present spectra for Au/GaN, and that the magnitude of the BEEM current was measured to be much larger than in the present work.

Interface transport in a defected area might produce a threshold lower in energy than that of the normal Schottky barrier height. Since the highest barrier heights measured by conventional $I-V$[9–11] are in excess of 1 eV, it is expected that BEEM measurements on high-quality GaN would usually produce values equal to or greater than this value. The BEEM spectra obtained here never displayed a first threshold lower than 1 V. If BEEM spectra with a threshold at 0.7 eV represent defected areas, they would have to be infrequent in order that they not dominate the $I-V$ measurements. However, the second threshold in the prior work might then represent the Schottky barrier height energy, in agreement with the value measured in the present experiments.

It is also possible, however, that high strain might perturb the band structure by decreasing both thresholds in energy. In this case the two thresholds here and in the work of Brazel *et al.* would represent the same band minima. Large strains at the GaN/sapphire interface exist[12] due to the large lattice mismatch; the range of surface strains, however, could depend on growth conditions and methods, and might be substantially different for epilayers from different sources.

The existence of a second threshold in the BEEM spectra is unexpected, since a second conduction-band minimum is not predicted at these energies by band structure calculations. Recent calculations[13,14] only produce a second band about 2 eV above the lowest minimum Γ_1, which is much higher than the second threshold observed here. Strain effects due to the large lattice mismatch between GaN and sapphire, as described above, could strongly distort the conduction bands[15] and perturb the splitting observed in BEEM spectra, and some degree of variation in strain and BEEM threshold would be expected.

The variation in threshold splitting will be examined more completely when samples exhibiting larger transmission are achieved. It is also important to verify that the two thresholds can be attributed to the nitride conduction-band structure. Further work on AlGaN epilayers would determine whether a systematic change of threshold separation with Al fraction occurs, providing an indication of the mechanism for the second threshold. The observation in BEEM measurements of zero transmission on some GaN layers and small trans-

mission on others will be investigated further in future experiments. This lack of transmission, and the extremely small BEEM currents observed in the transmitting samples, are possibly indicative of the same attenuation mechanism. This suggests that optimization of the GaN layers in some respect (most probably defect-related) might produce a further increase in transmission. Experiments involving the effects of annealing will also be performed.

In conclusion, BEEM spectroscopy and imaging have been performed on Pd/GaN and Au/GaN Schottky barrier structures. These experiments yield a value of Schottky barrier height which is reproducible and occurs at all observed positions at the metal/GaN interface. A second threshold is also reproducibly observed in the BEEM spectra of both structures, raising the possibility of a secondary conduction-band minimum about 0.2–0.3 eV above the primary minimum. BEEM imaging reveals transmission in most areas of the interface, although the magnitude is unusually small and varies strongly, a feature that also appears with both metals. This observation, together with the lack of measurable interface transmission when using other GaN material, suggests a persistent attenuation mechanism which may be defect-related. Further work on other GaN material, as well as BEEM measurements on AlGaN layers, are planned to clarify these issues.

ACKNOWLEDGMENTS

The research described in this article was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, and was jointly sponsored by the Office of Naval Research, the Ballistic Missile Defense Organization/Innovative Science and Technology Office, and the Ballistic Missile Defense Organization/GaN Microwave Power Amplifier Program through an agreement with the National Aeronautics and Space Administration (NASA).

[1]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988); L. D. Bell and W. J. Kaiser, *ibid.* **61**, 2368 (1988).
[2]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).
[3]P. J. Grunthaner, F. J. Grunthaner, R. W. Fathauer, T. L. Lin, M. H. Hecht, L. D. Bell, W. J. Kaiser, F. D. Schowengerdt, and J. H. Mazur, Thin Solid Films **183**, 197 (1989).
[4]L. D. Bell, W. J. Kaiser, M. H. Hecht, and L. C. Davis, in *Scanning Tunneling Microscopy*, edited by J. A. Stroscio and W. J. Kaiser (Academic, San Diego, CA, 1993), pp. 307–48.
[5]L. D. Bell, W. J. Kaiser, S. J. Manion, A. M. Milliken, R. W. Fathauer, and W. T. Pike, Phys. Rev. B **52**, 12081 (1995).
[6]H. Sirringhaus, T. Meyer, E. Y. Lee, and H. von Känel, Phys. Rev. B **53**, 15944 (1996).
[7]E. Y. Lee and L. J. Schowalter, J. Appl. Phys. **70**, 2156 (1991).
[8]E. G. Brazel, M. A. Chin, V. Narayanamurti, D. Kapolnek, E. J. Tarsa, and S. P. DenBaars, Appl. Phys. Lett. **70**, 330 (1997).
[9]M. Khan, T. Detchprohm, P. Hacke, K. Hiramatsu, and N. Sawaki, J. Phys. D **28**, 1169 (1995).
[10]E. V. Kalina, N. I. Kuznetsov, V. A. Dmitriev, K. G. Irvine, and C. H. Carter, Jr., J. Electron. Mater. **25**, 831 (1996).
[11]A. C. Schmitz, A. T. Ping, M. A. Khan, Q. Chen, J. W. Yang, and I. Adesida, Semicond. Sci. Technol. **11**, 1464 (1996).
[12]H. Siegle, A. Hoffman, L. Eckey, C. Thomsen, J. Christen, F. Bertram, D. Schmidt, D. Rudloff, and K. Hiramatsu, Appl. Phys. Lett. **71**, 2490 (1997).
[13]W. R. L. Lambrecht, B. Segall, J. Rife, W. R. Hunter, and D. K. Wickenden, Phys. Rev. B **51**, 13516 (1995).
[14]M. van Schilfgararde, A. Sher, and A.-B. Chen, J. Cryst. Growth **178**, 8 (1997).
[15]Z. Yang and Z. Xu, Phys. Rev. B **54**, 17577 (1996).

Ballistic electron emission microscopy spectroscopy study of AlSb and InAs/AlSb superlattice barriers*

X.-C. Cheng and T. C. McGill
Department of Applied Physics, California Institute of Technology, Pasadena, California 91125

(Received 21 January 1998; accepted 28 May 1998)

Due to its large band gap, AlSb is often used as a barrier in antimonide heterostructure devices. However, its transport characteristics are not totally clear. We have employed ballistic electron emission microscopy (BEEM) to directly probe AlSb barriers as well as more complicated structures such as selectively doped *n*-type InAs/AlSb superlattices. The aforementioned structures were grown by molecular beam epitaxy on GaSb substrates. A 100 Å InAs or 50 Å GaSb capping layer was used to prevent surface oxidation from *ex situ* processing. Different substrate and capping layer combinations were explored to suppress background current and maximize transport of BEEM current. The samples were finished with a sputter deposited 100 Å metal layer so that the final BEEM structure was of the form of a metal/capping layer/semiconductor. Of note is that we have found that hole current contributed significantly to BEEM noise due to type II band alignment in the antimonide system. BEEM data revealed that the electron barrier height of Al/AlSb centered around 1.17 eV, which was attributed to transport through the conduction band minimum near the AlSb *X* point. Variation in the BEEM threshold indicated unevenness at the Al/AlSb interface. The metal on semiconductor barrier height was too low for the superlattice to allow consistent probing by BEEM spectroscopy. However, the superlattice BEEM signal was elevated above the background noise after repeated stressing of the metal surface. A BEEM threshold of 0.8 eV was observed for the Au/24 Å period superlattice system after the stress treatment. © *1998 American Vacuum Society.* [S0734-211X(98)12004-8]

I. INTRODUCTION

There has been much interest in the GaSb, AlSb, and InAs lattice matched material system. Due to the unique type II band alignment and narrow band gaps of these materials, much research has focused on developing antimonides for infrared lasers[1,2] and detectors,[3] as well as for high speed integrated circuits.[4] Common to all these devices is the need for a barrier like constituent such as AlSb. However, AlSb barriers often appear leaky, especially in Schottky gate type applications.[5,6] Moreover, AlSb is *p* type when left unintentionally doped, and *n*-type doping of AlSb is not always convenient. Recently, it has been reported that InAs/AlSb superlattices can be an attractive alternative for *n*-type barrier applications such as making cladding layers in laser structures.[7] The superlattice approximately lattice matches with GaSb when the InAs and AlSb constituent layers are given the same thickness. The band gap and effective barrier height of the superlattice are tunable by varying the superlattice period, and *N*-type doping can be achieved by incorporating Si in the InAs well.

In this article, ballistic electron emission microscopy (BEEM) was used to characterize these antimonide barrier structures. BEEM is a technique based on scanning tunneling microscopy (STM) and is especially suited to study of local transport properties.[8,9] In a BEEM experiment, the semiconductor of interest is coated with a thin layer of metal, which supports STM tunnel current while a collector terminal at the back of the sample conducts away electrons that leak across the metal–semiconductor interface. In BEEM spectroscopy, the collector current is monitored as a function of tunneling voltage while the tunneling current is kept constant by varying the STM tip to sample distance. As the STM tip potential rises above the bottom of the semiconductor conduction band, electrons can travel ballistically across the thin metal region and enter the semiconductor unimpeded, causing a noticeable increase in collector current. Thus the BEEM turn on threshold depends on the underlying semiconductor band structure and can be used to evaluate Schottky barrier heights and semiconductor band edges in the antimonide material system. The scanning probe nature of BEEM allows local variations of these properties to be mapped and was utilized to examine the integrity of AlSb barriers.

Since its inception, BEEM spectroscopy has been used extensively to study metal on semiconductor structures. It has been adapted by several workers to study the Au/Si (Ref. 9) and the $Au/Al_xGa_{1-x}As$ system.[10,11] Recently, the technique has been successfully applied to GaAs/AlAs (Ref. 12) and InAs/AlSb (Ref. 13) double barrier heterostructures, which indicates that the technique is promising for probing less conventional systems such as the aforementioned InAs/AlSb superlattice.

II. EXPERIMENT

The samples used in this experiment were grown in a Perkin–Elmer 430 molecular beam epitaxy (MBE) system equipped with a valved As cracker. Figure 1 shows the structure of typical BEEM samples. Te doped ($n=5\times10^{17}$ cm^{-3}) GaSb wafers were used for most of the growth. This was to ensure that the substrate would be conductive enough in subsequent BEEM experiments. Follow-

*No proof corrections received from author prior to publication.

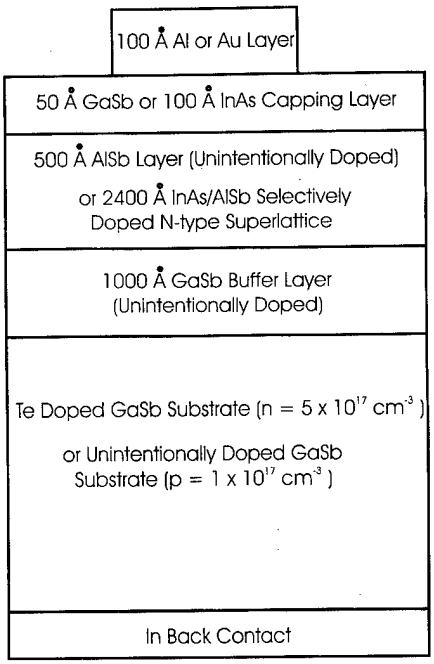


FIG. 1. Structure of the BEEM samples.

ing oxide desorption under Sb overpressure, an unintentionally doped GaSb buffer layer was grown. Since GaSb is *p* type from background doping, the buffer layer was kept as thin as possible without compromising the structural quality of subsequent growth. At low growth rate, a 1000 Å thick buffer layer was found to be adequate. At the end of the buffer growth, samples were soaked in Sb, yielding the (1 ×3) reflection high energy electron diffraction (RHEED) pattern characteristic of a reconstructed GaSb surface.

For AlSb studies, a 500 Å layer of unintentionally doped AlSb was grown over the smoothed GaSb surface. The thickness was selected so that the bulk properties could be examined while at the same time the layer was thin enough to support transport of BEEM current. Because the AlSb layer was relatively thin, the substrate temperature was kept at 520 °C, the same as for GaSb growth. RHEED for the AlSb layer was less streaky but still showed the characteristic 1 ×3 pattern. To prevent AlSb oxidation, the samples were capped off at the end of the growth by either a 50 Å GaSb layer or a 100 Å InAs layer. The substrate temperature was lowered to 470 °C for growing the InAs capping layer.

For growth of the selectively doped InAs/AlSb superlattice, the substrate temperature had to be lowered to prevent excessive As incorporation in the antimonide layers. The structural quality of the superlattice was significantly improved when the growth temperature was lowered to 420 °C, at which point the GaSb surface turns Sb rich and the RHEED pattern changes from 1×3 to 1×5. During growth of the InAs constituent layer, the Si dopant cell shutter was opened, and As flux was minimized by using the valved cracker while maintaining an As stabilized growth front. A 10 s Sb soak was applied between each InAs and AlSb interface to ensure a InSb like interface, which is known to produce material of superior quality.[14] The RHEED pattern remained streaky throughout the growth and exhibited sharp 2×4 and 1×3 reconstructions for the InAs and AlSb layers, respectively. The samples were grown with superlattice periods of 17, 24, and 48 Å. The period thickness was split between the InAs and AlSb layers, which were under alternative compressive and tensile strain. The total thickness of the superlattice was kept constant for all the samples at 2400 Å. To prevent oxidation, the superlattice was capped with 50 Å of GaSb following completion of the last AlSb layer.

A sputter-etch deposition system was used for postgrowth metallization. Aluminum and gold were sputtered off solid targets by Ar plasma and deposited onto the sample at rates up to 0.4 Å/s. The samples were placed behind a mask and patterned with arrays of metal dots 1 mm^2 in area. Metal layer thicknesses up to 100 Å were experimented. Atomic force microscopy (AFM) studies showed that the typical metal layer had a root mean square (rms) roughness on the order of 5 Å. For most samples, the surface morphology was smooth and appeared suitable for BEEM studies. Prior to metallization, samples were taken out of the ultrahigh vacuum (UHV) growth environment and exposed to the ambient. Hence a 20–30 Å thick native oxide was present between the metal and semiconductor layer. Talin *et al.* have shown that the oxide layer does not affect BEEM results for Au/GaAs structures.[15] In our study, it was found that the antimonide samples with native oxides were stable for up to several weeks. To minimize contamination from handling, a degreasing procedure was followed before the sample was introduced to the metallization chamber which may partially remove the native oxide layer. It consisted of sequential ultrasonic rinses in acetone, isopropanol, and de-ionized water, with each rinse lasting 2 min. The procedure helped generate more consistent BEEM results, especially for the samples that have been stored in air for some time.

Our BEEM setup was configured for experiments in air at room temperature. It was based on a Digital Instruments scanning tunneling microscope unit (Nanoscope III). A fine Au wire was spring mounted against the top of the sample as the STM base contact, while indium left from the growth, on the back side of the sample, served as the BEEM back contact. The setup was tested and calibrated with the Au/Si(100) system. The equipment has also been successfully used to study the $Au/Al_xGa_{1-x}As$ system.[11]

III. RESULTS AND DISCUSSION

Prior to BEEM experiments, samples were characterized by $I-V$ measurements. Figure 2 shows the results from two types of AlSb samples. Sample A was from an early growth on an unintentionally doped *p*-type GaSb substrate and capped by an InAs layer. Sample B was grown on an *n*-type GaSb substrate and capped by GaSb. The sample B substrate/capping layer combination was standard for most of the AlSb samples and for all superlattice samples. As can be seen from the $I-V$ behavior of sample A, AlSb acted as a poor barrier for the InAs/GaSb junction, which allowed holes to tunnel out of the underlying *p*-type substrate, raising the

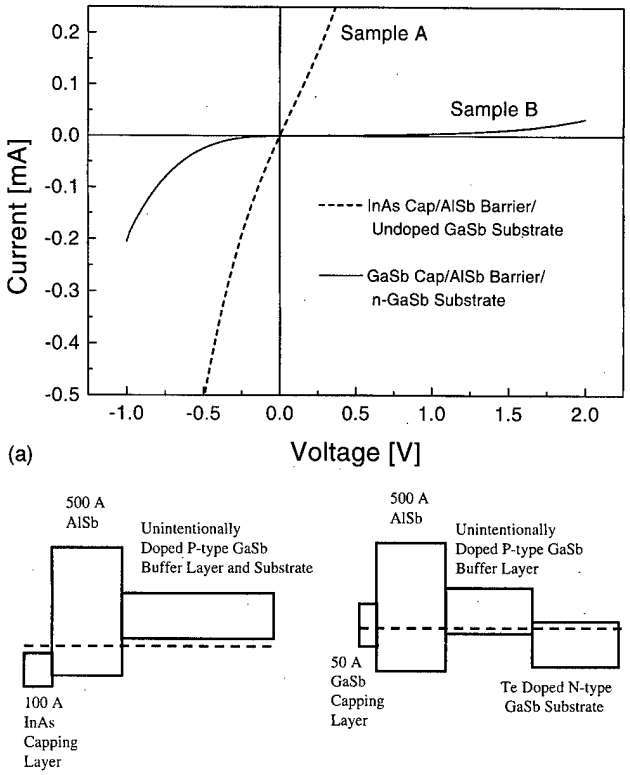


FIG. 2. $I-V$ curves (a) and flat band diagrams (b) for two types of AlSb BEEM samples. Both samples were metallized with 100 Å of Al. The Al mesa had an area of 1 mm^2. Sample B yielded BEEM curves but sample A was too noisy for BEEM measurements.

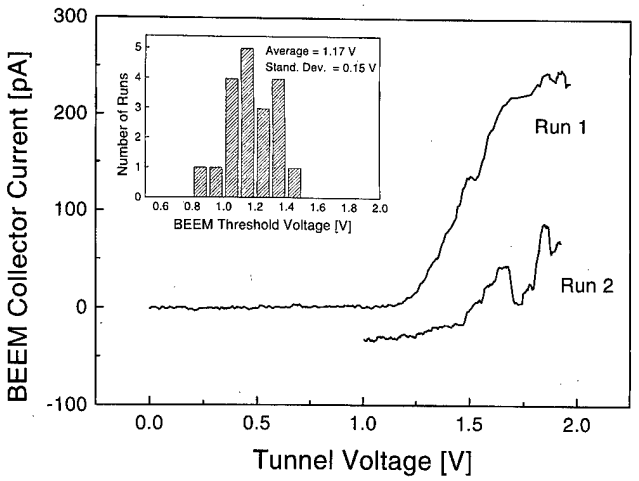


FIG. 3. BEEM $I-V$ curves for AlSb barriers grown on n-type GaSb substrates. The samples were capped with 50 Å of GaSb and metallized with 100 Å of Al. The tunneling current was held constant at 10 nA. Run 1 was over the whole spectral range, and run 2 was for the high voltage region, which is offset for clarity. The spectrum was not averaged. The inset shows the spread in BEEM threshold from a number of runs.

background BEEM current to hundreds of picoamps and rendering the device unsuitable for BEEM spectroscopy. By contrast, sample B was much more resistive because the n-type substrate blocked much of the tunneling current.

Figure 3 shows two BEEM $I-V$ curves from a type B AlSb sample. Each curve was taken from a different place on the sample surface and took approximately 10 s to generate. Since the experiment was carried out in air at room temperature, there was some tip drift even after the system had been given hours to equilibrate. Typical drift rates were about a few nm/min. Hence BEEM scans were not averaged in order to preserve spatial resolution in the experiment. The BEEM $I-V$ curves were analyzed by using the well known Bell–Kaiser model,[9] which assumes that the BEEM threshold behavior takes on the form

$$I_c=\sum_{i=1}^{n}(V-V_i)^2,$$

where I_c is the BEEM collector current, V the tunnel voltage, and V_i the threshold voltage. By examining a large number of runs, it was found that the turn on voltage centered around 1.17 eV with a standard deviation of 0.15 eV. This was in fair agreement with the result obtained by Walachova *et al.*[13] in their study of InAs/AlSb double barrier heterostructures. The BEEM turn on threshold was attributed to the conduction band minimum near the AlSb X point and verified the importance of transport through indirect band minima in AlSb. The L and Γ points of AlSb lie higher and could not be delineated from the BEEM data. As shown in Fig. 3, there was significant variation among the individual BEEM $I-V$ curves. The large variation in individual BEEM thresholds indicated uneveness at the metal–semiconductor interface, consistent with the fact that AlSb barriers are often leaky. This is in contrast with the BEEM study of AlAs, where the BEEM turn on voltage exhibited minimal variation across the wafer.[11]

It should be noted that the BEEM current background noise in the AlSb sample was on the order of 5 pA, which was higher than similarly prepared AlAs samples even through the barrier height in both systems was about 1.2 eV. We attribute this discrepancy to the fact that the background doping was p type for AlSb and n type for AlAs thus the Fermi level and tunneling barriers were likely different even for similar bias voltage. The increased background BEEM current was accounted for by additional hole current in the AlSb system due to the smaller hole barrier. The dominance of hole current was evident in the $I-V$ response of the sample to ambient light.

Background noise was also a significant problem in BEEM spectroscopy of the InAs/AlSb superlattices. This is because the effective superlattice band gap is substantially smaller than that of AlSb, even for samples with a very short period. We have attempted to grow these superlattice structures with period thicknesses of 17, 24, and 48 Å. According to a calculation performed through an eight-band $k\cdot p$ model that included the effects of strain, the band gaps for these structures should be 1.2, 1.15, and 0.88 eV, respectively.[7] The band gaps could be made larger by growing structures with shorter superlattice periods, but the structural quality of the material deteriorated rapidly as the superlattice period

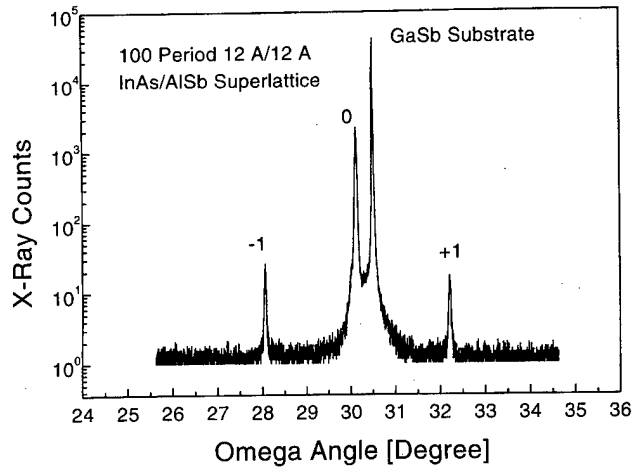


FIG. 4. High resolution x-ray diffraction scan from a 100 period, 12 Å/12 Å/InAs/AlSb superlattice grown on a GaSb substrate.

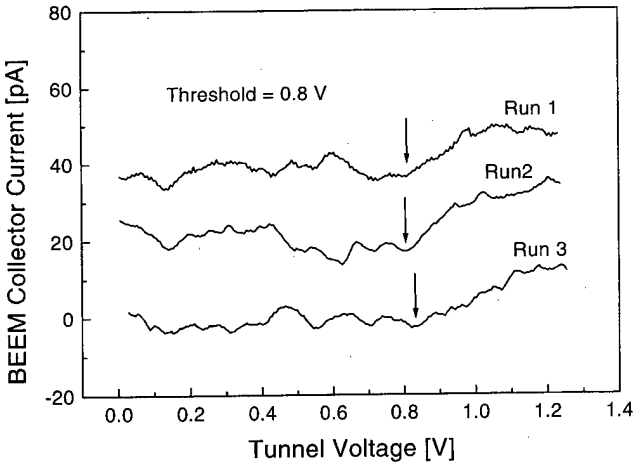


FIG. 6. BEEM $I-V$ curves from the 24 Å period InAs/AlSb superlattice. The superlattice was capped with 50 Å of GaSb and metallized with 100 Å of Au. The samples had previously been stressed by ramping the tip voltage to −3 V while holding the tip current at 50 nA. The tunneling current was held constant at 10 nA during the BEEM scan. The curves were averaged over 255 runs to reduce background noise and are offset for clarity.

was decreased. In fact, x-ray rocking curves for samples with a 17 Å period showed multiple splits at the superlattice peak, indicating that the layer had relaxed from too much strain. The inferior quality of these samples rendered them unsuitable for BEEM studies. The 48 Å longer period sample exhibited good structural integrity but its band gap was too small to keep background BEEM current at a reasonable level. Thus only the 24 Å period samples seemed suitable for BEEM experiments.

Figure 4 shows a high resolution x-ray diffraction scan of the 24 Å period superlattice. The sharp x-ray diffraction satellites were indicative of good structural quality. The $I-V$ curve of the metallized device is shown in Fig. 5 and indicated that the underlying superlattice was n type. The curve deviated significantly from ideal Schottky diode behavior at high voltages. But the low voltage portion of the curve could be used to extract a Schottky barrier height of 0.6 eV.[16]

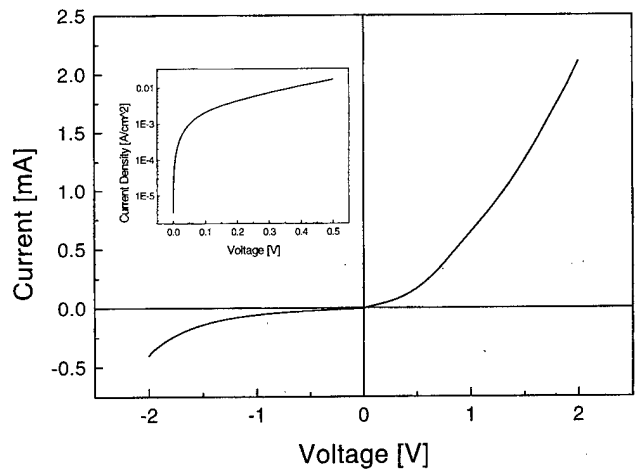


FIG. 5. $I-V$ curve of a BEEM device based on the 24 Å period InAs/AlSb superlattice. Samples were capped with 50 Å of GaSb and metallized with 100 Å of Au. The Au mesa had an area of 1 mm^2. The inset shows a log plot of the current density when the sample was under forward bias, from which a Schottky barrier height of 0.6 eV was extracted.

When these samples were inserted into the BEEM setup, the background BEEM current noise was on the order of 100 pA, which overwhelmed any conventional BEEM signal that would be present. However, it was found that, after the surface was stressed by running a high voltage and current (−3 V and 50 nA) through the STM tip, the metal layer could become deformed resulting in regions where the metal layer was tenuous. When the STM tip was placed over these regions, possible BEEM thresholds could be observed in the BEEM spectroscopy curve. Figure 6 shows some typical BEEM scans after the stress treatment. The estimated possible threshold occurred at around 0.8 eV for the Au/superlattice system and could be reproduced by retracting the STM tip and using it to stress a new region.

IV. SUMMARY AND CONCLUSION

We have demonstrated that BEEM spectroscopy can be applied to the InAs/GaSb/AlSb material system. The Al/AlSb system yielded a BEEM threshold of 1.17 eV, which was attributed to transport through the conduction band minimum near the AlSb X point. The BEEM threshold varied by up to 0.2 eV across the wafer, indicating degradation of the AlSb barrier due to unevenness at the metal–semiconductor interface. It was found that the surface capping layer and the underlying substrate played important roles in suppressing background BEEM current. The junctions formed by InAs/thin AlSb barrier/p-GaSb leaked hole current, which was detrimental to BEEM measurement. In the case of a selectively doped n-type superlattice, BEEM spectroscopy was hampered by considerable background BEEM current. BEEM scans yielded a threshold of 0.8 eV for the Au/24 Å period superlattice system only after considerable stressing of the metal layer. More work would be needed to produce a short period, large band gap superlattice

with superior material quality, the large band gap being needed to suppress background BEEM current.

ACKNOWLEDGMENTS

The authors would like to thank D. H. Chow of Hughes Research Lab for helpful discussions of antimonide growth. This study was supported in part by the Office of Naval Research under Grant No. N00019-89-J-00014 and by the Air Force Office of Scientific Research under Grant No. F49620-93-J-0258.

[1]H. Lee, P. K. York, R. J. Menna, R. U. Martinelli, D. Z. Garbuzov, S. Y. Narayan, and J. C. Connolly, Appl. Phys. Lett. **66**, 1942 (1995).
[2]R. H. Miles, D. H. Chow, Y.-H. Zhang, P. D. Brewer, and R. G. Wilson, Appl. Phys. Lett. **60**, 1921 (1995).
[3]D. H. Chow, R. H. Miles, C. W. Nieh, and T. C. McGill, J. Cryst. Growth **111**, 683 (1991).
[4]D. H. Chow, H. L. Dunlap, W. Williamson, III, S. Enquist, B. K. Gilbert, S. Subramaniam, P.-M. Lei, and G. H. Berstein, IEEE Electron Device Lett. **17**, 69 (1996).
[5]M. Drndic, M. P. Grimshaw, L. J. Cooper, and D. A. Ritchie, Appl. Phys. Lett. **70**, 481 (1997).
[6]M. J. Yang, F.-C. Wang, C. H. Yang, B. R. Bennett, and T. Q. Do, Appl. Phys. Lett. **69**, 85 (1996).
[7]D. H. Chow, Y. H. Zhang, R. H. Miles, and H. L. Dunlap, J. Cryst. Growth **150**, 879 (1995).
[8]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[9]L. D. Bell and W. J. Kaiser, Phys. Rev. Lett. **61**, 2368 (1988).
[10]J. J. O'Shea, T. Sajoto, S. Bhargava, D. Leonard, M. A. Chin, and V. Narayanamurti, J. Vac. Sci. Technol. B **12**, 2625 (1994).
[11]X.-C. Cheng, D. A. Collins, and T. C. McGill, J. Vac. Sci. Technol. A **15**, 2063 (1997).
[12]T. Sajoto, J. J. O'Shea, S. Bhargava, D. Leonard, M. A. Chin, and V. Narayanamurti, Phys. Rev. Lett. **74**, 3427 (1995).
[13]J. Walachova, J. Zelinka, J. Vanis, D. H. H. Chow, J. N. Schulman, S. Karamazov, M. Cukr, P. Zich, J. Kral, and T. C. McGill, Appl. Phys. Lett. **70**, 3588 (1997).
[14]G. Tuttle, H. Kromer, and J. H. English, J. Appl. Phys. **67**, 3032 (1990).
[15]A. A. Talin, D. A. Ohlberg, R. S. Williams, P. Sullivan, I. Koutselas, B. Williams, and K. L. Kavanagh, Appl. Phys. Lett. **62**, 2965 (1993).
[16]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981), p. 279.

Current oscillations in thin metal–oxide–semiconductor structures observed by ballistic electron emission microscopy

H. J. Wen[a] and R. Ludeke[b]
IBM T. J. Watson Research Center, Yorktown Heights, New York 10598

Andreas Schenk
Integrated Systems Laboratory, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland

(Received 21 January 1998; accepted 11 May 1998)

Quantum interference oscillations of electrons in a thin SiO_2 layer were observed by ballistic electron emission microscopy (BEEM). With BEEM, electrons are injected across the gate of a metal–oxide–semiconductor (MOS) structure and directly into the conduction band of the SiO_2. The MOS capacitor consisted of a 5 nm thick Pd film deposited on a 2.8 ± 0.2 nm oxide thermally grown on Si(100). Oscillations with up to four peaks in an energy range of 0–3 eV above the injection threshold were noted. Their magnitude is of the order of 30% of the underlying BEEM current. The oscillations were most salient and their energy location repeatable at points of the sample that were previously not exposed to the electron beam. Even modest exposures caused a buildup of positive charge. This charge resulted in energy shifts, as well as a weakening of the oscillations, both of which are a consequence of the added scattering and local field inhomogeneities associated with the random distribution of the positive charge. Solutions of the Schrödinger equation that included a built-in oxide potential of 0.20 V and image force effects at both interfaces gave excellent fits to the experimental data for an effective electron mass in the oxide $m_{ox}=0.63\pm0.09m_o$. The uncertainty in m_{ox} arises from an uncertainty of ±0.2 nm in the determination of the oxide thickness by ellipsometric methods. Nevertheless, the obtained value is well above the generally accepted value of $0.5m_o$. © *1998 American Vacuum Society.* [S0734-211X(98)05804-1]

I. INTRODUCTION

The notion that electron wave interference should occur in Fowler–Nordheim (FN) tunneling of electrons into thin (<6 nm) oxide layers comprising a metal–oxide–semiconductor (MOS) structure was proposed by Gundlach over three decades ago.[1] Oscillations in the (oxide)-bias-dependent current arise from the interference of electron waves reflected at the oxide–semiconductor (OS) interface and at the point of emergence of the tunneling electrons at the bottom of the tilted SiO_2 conduction band. The latter point is a "hard" turning point, whereas the OS interface is relatively transparent, with reflections occurring mainly from the wave-function mismatch across the interface. Experimental verification of a weak oscillatory structure in the FN current was reported by Maserjian and Petersson in 1974,[2,3] and by others in subsequent years.[4–8] Invariably, these data were analyzed in terms of Gundlach's theory based on a trapezoidal barrier (i.e., neglecting image force effect),[1] from which an estimation of the conduction-band effective mass m_{ox} of SiO_2 can be made. Values for m_{ox} ranged from $0.32m_o$ (Ref. 1) to as high as $0.85m_o$,[2,6] where m_o is the free-electron mass. The conduction-band effective mass was estimated as well from fits of the theoretical FN current to experimental data covering many orders of magnitude of the current. Again, image force effects were neglected and values reported ranged from a low of $0.3m_o$ (Ref. 9) to $0.5m_o$.[10] When image forces were included, the values were somewhat higher.[11,12] Although a defacto value of $0.5m_o$ is almost exclusively used in transport simulations,[13] the reported discrepancies in m_{ox} and its dependence on fitting assumptions questions the acceptance of this standard, and prompts one to look for alternative experimental methods to extract a value for m_{ox}. Intrinsically, interference phenomena represent the most straight forward method to arrive at a value of m_{ox}, provided the inherent difficulties in the experiment and simulation of FN tunneling can be overcome. These include a position-dependent energy of the electron, uncertain tunneling (injection) description, a weak signal superimposed on a strongly rising current and averaging effects due to a large and often inhomogeneous device area. These constraints, as we shall see, can be circumvented by using the local injection scheme of ballistic electron emission microscopy (BEEM). This scanning tunneling microscope (STM) based method allows the injection of variable energy and nearly monochromatic electrons into the thin gate of a MOS structure and then directly into the conduction band of the SiO_2.[14] Both the high lateral confinement (<2 nm) of the injected electrons in thin oxides (<4 nm),[15] which enhances the attainment of a homogeneous local oxide potential, and the relatively weak power dependency on energy for the transmitted current are contributory to the realization of pronounced interference oscillations. Moreover, the modeling of transport in SiO_2 for "over the barrier" injection depends on fewer unknowns than FN and direct tunneling, which assures a higher degree of confidence in the calculated parameters obtained through fits to the data. We present here both oscillatory BEEM current data for a 2.8 nm SiO_2 layer and fits

[a]Present address: IBM Austin Research Laboratory, Austin, TX 78758.
[b]Corresponding author; electronic mail: ludeke@watson.ibm.com

using solutions of the Schrödinger equation that include image force effects, with the relevant adjustable parameter being the effective mass m_{ox} and oxide thickness d. With the assumption that d is known within ± 0.2 nm, a "best-fit" value of $m_{ox}=0.63\pm 0.09 m_o$ was obtained. Uncertainties in other parameters needed for the fits, such as the built-in oxide potential V_{ox} and the effective dielectric constant ϵ_{ox} proved to be of minor consequence to the error in m_{ox}.

II. EXPERIMENTAL DETAILS

A. Ballistic electron emission microscopy (BEEM)

BEEM is an adaptation of the conventional STM and is characterized by a special sample configuration that consists of a thin conducting layer, usually a metal, deposited on top of the semiconductor structure to be measured. Although historically a Schottky barrier, here the sample is a MOS structure. The metal provides a ground contact relative to which both the STM tip bias V_T and the applied oxide bias V_b are referenced. The tip bias thus defines the energy eV_T of the electrons injected into the metal. For metal film thicknesses comparable or less than the electron mean-free path, most of the electrons will traverse the metal and reach the far interface without scattering (ballistically). If the electrons encounter a potential barrier at that interface, they will backscatter unless their energy exceeds that of the metal–oxide barrier. In this case, a fraction of the electrons will be injected into the conduction band of the SiO_2. Here, they may undergo electron–phonon scattering, which may cause some of them to return to the metal.[14] The remainder, if not trapped,[16,17] proceed towards the Si substrate to emerge as a collector current I_c. The STM is operated under constant I_T conditions. For the experiments reported here $I_T=2$ nA. In the spectroscopy mode of BEEM, the STM image acquisition is interrupted at a predetermined point on the surface and the collector current I_c is measured as V_T is ramped over a voltage range that includes the barrier potential. Initially, I_c is zero until V_T exceeds a threshold value V_o that represents the maximum in the barrier potential of the MOS structure.

An energy-band diagram for a BEEM experiment on a MOS structure is schematically shown in Fig. 1 for $V_b=0$. The curved leading edge of the oxide barrier results from the inclusion of image force lowering,[18] a corresponding, but weaker effect at the SiO_2–Si interface was omitted for clarity. Conditions for injection into the conduction band of the SiO_2 are shown ($V_T>V_o$). The threshold for injection V_o is about 4 V in the absence of a negative trapped charge.[17] The application of an external oxide bias V_b further modifies the electron energies as they move across the oxide. In the present experiments $V_b=0$, however, $V_{ox}\approx 0.2$ V due to work-function differences between the n-type Si and the Pd layer.[18] It is worth pointing out that the energy distribution of electrons injected by the STM tip fall off in near-exponential fashion from its maximum value, with an energy spread that decreases with increasing energy eV_T. A theoretical full width at half maximum of ~0.150 eV was estimated for $eV_T=6$ eV.[19] This energy spread is sufficiently monochromatic for characterizing most hot electron phenomena in oxides.

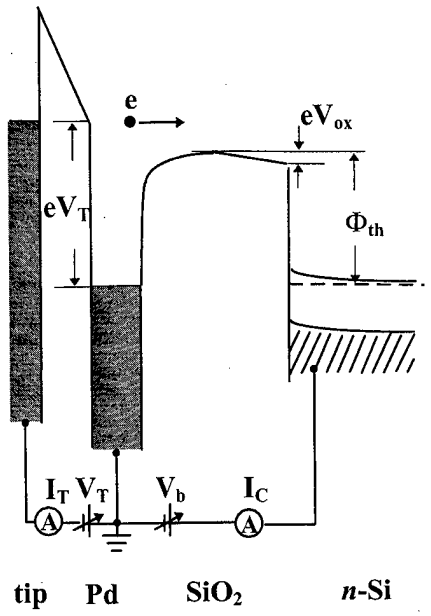


FIG. 1. Energy-band diagram of a BEEM experiment applied to a MOS structure. V_T is the STM tip bias that determines the energy eV_T of the electrons injected by the tip into the Pd metal gate. An optional oxide bias V_b may also be applied. Here, $V_b=0$. The oxide potential V_{ox} shown is attributable to work-function differences between Pd and n-Si, $V_{ox}\approx 0.2$ V in the absence of oxide charge.

B. Sample preparation

The device-grade SiO_2 layers were thermally grown near 800 °C in dry oxygen on 125 mm Si(100) wafers doped in the low 10^{17} cm^{-3} range. The wafer was subsequently annealed in forming gas at 500 °C. Working samples of ~8 $\times$15 mm^2 were cleaved from the wafers and introduced into the ultrahigh vacuum (UHV) preparation chamber, where they were outgassed over night at ~200 °C to desorb water and other surface contaminants. The sample was then transferred under UHV to the metal deposition chamber, where Pd dots 0.2 mm in diameter were thermally evaporated onto the SiO_2 through a shadow mask. The substrate was held near 30 K during deposition in order to smooth the surface morphology of the thin (~5 nm) Pd films. This process produced films with a nodular structure, typically, 8 nm in diameter that protruded <2 nm above the valleys.[17] A smooth surface morphology is desirable to reduce BEEM image contrast arising from the surface topography of the metal.[14] The finished sample was allowed to warm up to room temperature and was then transferred under UHV into the STM chamber. The grounding contact was carefully positioned onto a selected Pd dot by means of three orthogonally mounted Inchworms™. STM images and sets of BEEM spectra were then taken. Typically, 9–25 BEEM spectra were measured in a grid pattern covering 25×25 to 50×50 nm^2 areas. As will become apparent in the next section, it is desirable to widely separate the acquisition points for each spectrum to avoid charging effects arising from electrons in-

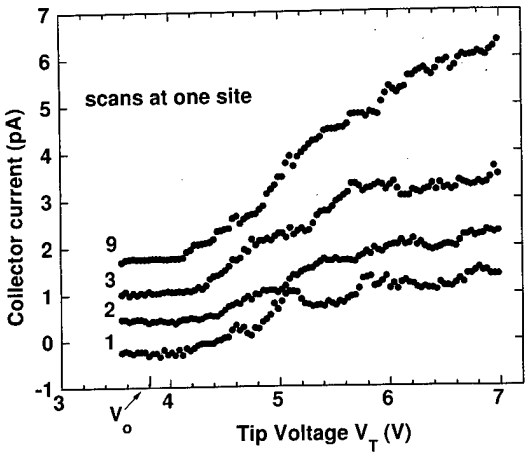


FIG. 2. Single scan BEEM spectra on a 5 nm Pd/2.8 nm SiO_2/n-Si(100) MOS structure measured at the same point on the sample. The numerals indicate the number of the scan. The first spectrum was taken on a previously unexposed area of the sample. V_0 marks the current threshold bias, obtained by computer-aided fits. The spectra are vertically displaced for clarity. $I_T=2$ nA, $V_b=0$ V.

jected at a prior location and then scattered into the area to be probed. In order to study charging effects, sequential spectra were acquired at the same location of the sample. Although spectral acquisition times are of the order of 10 s, tip drifts were of the order 0.1 nm/min, which assured that injections occurred at the same point of the sample.

The oxide thickness was determined by spectroscopic ellipsometry on two different instruments operating at a 632.8 nm wavelength. The thickness measured by both instruments agreed within 0.3 nm. We found that outgassing the sample in UHV at ~200 °C prior to the measurement in the same stage and conditions as the BEEM samples prior to metallization, reduced the measured thickness by ~0.3 nm. The difference was attributed to adsorbed contaminants. The value for $d=2.8\pm0.2$ nm was the best estimate for the oxide thickness after the anneal and was used in the simulations. In unpublished studies at IBM, the oxide thickness determined by $C-V$ methods on oxides of comparable thickness, after correction for polydepletion and quantization effects, was, in general, ~0.2–0.3 nm thinner than that obtain with the ellipsometer.[20] Consequently, the value quoted above, except for instrument error, most likely represents a slight overestimate in the thickness.

III. COLLECTOR CURRENT OSCILLATIONS

A sequence of single spectral scans taken at the same point of the 5 nm Pd/2.8 SiO_2/Si(100) sample is shown in Fig. 2. The digits next to the curves indicate the number of the spectral scan. The first scan shows a strong modulation of the collector current that progressively broadens and weakens as the number of scans increases. The oscillatory structure is attributed to quantum interference effects in the thin SiO_2 layer, which arise from the constructive/destructive interference of electron waves reflected at the metal–SiO_2 and SiO_2–Si boundaries of the SiO_2 "cavity." In the simplest realization, transmission maxima follow the quantization condition: $E=(n\pi\hbar/d)^2/2m^*$ $(n=1,2,3,...)$, where E is the electron energy, d the cavity width, and m^* the effective electron mass.[1] No oscillatory structure due to the metallic film was ever observed by us, a failure that we attribute to the uneven nodular character of the Pd film.[17,21] The oscillatory structure in the first scan of Fig. 2 is repeatable provided the scan is made on a new, previously unexposed point of the sample. However, only about 1/3 of the virgin points yielded a spectral structure with similar periodicities; the remainder lacked the oscillatory structure entirely or exhibited a weak and smeared out structure of varying periodicities. After a number of scans, shown here at the ninth, the structure is altered and strongly suppressed. A second point to be noted is the progressive increase in the collector current with each scan. Such increases were previously observed,[21] and were attributed to the buildup of stress-induced positive charge near the anode (OS interface). Although for thicker oxides ($d>4$ nm) electrons are trapped in the oxide, resulting in an increase in V_o and a decrease I_c, for thinner oxides the electrons leak out, thereby revealing the presence of the positive charge.[17] Its presence at the anode has a small effect (<0.1 V) on V_o, primarily due to image force lowering, that is consistent with our observations (~0.05 V). Yet, the positive charge creates an accelerating field that, again with the inclusion of image force effects, results in an enhancement of transmission probabilities and an increase in I_c.[22] Since the positive charge is randomly distributed, its buildup during the scans progressively distorts the local potential the electrons see as they traverse the oxide. Such potential fluctuations affect the ability of the electron waves to interfere coherently.

The oscillatory structure commonly observed on a virgin portion of the surface, such as the bottom curve in Fig. 2, exhibits peaks for $V_T\approx4.6$, 5.1, 5.8, and 6.8 V. We have limited ourselves to $V_T<7$ V to minimize hot electron damage. We also convinced ourselves that an additional weak peak appears near 4.1 V, but is not readily discernible in an average spectrum because of the background noise. It should be realized that less than 1 in 1000 electrons injected by the STM tip are collected in the Si substrate. The oscillatory structure is also shown in Fig. 3, where we have averaged spectra from five data sets. The observation of pronounced oscillation due to quantum interference effects is at first somewhat surprising, since the "cavity" of the SiO_2 film is quite leaky, with allowed states in both cladding regions (metal and Si) expected to reduce the reflectivity at the interfaces. Moreover, the extensive electron–phonon scattering in SiO_2, with a mean-free path of 1–2 nm that is, typically, less than the SiO_2 film thickness,[14,23] is expected to further suppress the interference oscillations. Of course, interference oscillations have been observed, albeit weak, in MOS capacitors using the FN injection method.[2–7] Support for the correctness of this interpretation also comes from the theoretical modeling that will be described next. Compared to the early work by Gundlach[1] for over the barrier interference, the present work incorporates image force effects, which

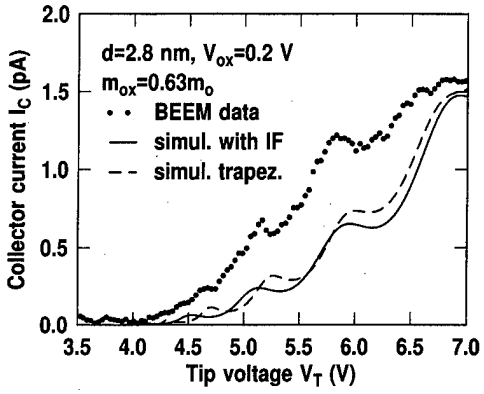


FIG. 3. Composite BEEM spectrum of five sets of data, each taken on a previously unexposed location of a 5 nm Pd/2.8 nm SiO_2/n-Si(100) MOS capacitor. The solid line represents the best fit of the peak location to the data using theoretical transmission coefficients calculated with the indicated parameters and with image force effects included. The dashed curve was obtained by omitting the image force effects.

represent the screening of the electrons in SiO_2 by nearby conduction electrons in the metal and Si. We will fit the theoretical curves to our data and extract from it the relevant parameter of the effective conduction-band mass m_{ox} of SiO_2.

IV. THEORETICAL TRANSMISSION COEFFICIENTS

The transmission coefficient (TC) $T(E)$ was calculated by a numerical solution of the one-dimensional Schrödinger equation assuming an idealized potential barrier with (i) the classical image potential, (ii) its divergencies removed as discussed below, (iii) neglecting oxide charges, and (iv) using the optical dielectric constant of the present thin SiO_2 layer. For dispersion in the latter, we assumed a parabolic $E(k)$ relation with an effective mass m_{ox} as parameter. The barrier was discretized by N partial subbarriers of rectangular shape which covered the whole oxide layer of thickness d. From the continuity of wave-function and quantum-mechanical current density at each boundary, the TC is then found by (see, e.g., Ref. 24)

$$T(E)=\frac{m_0}{m_{N+1}}\frac{k_{N+1}}{k_0}\frac{|\det M|^2}{|M_{22}|^2}, \tag{1}$$

where M is a (2×2) product matrix $M=\Pi_{l=0}^{N}M_l$ with transfer matrices M_l given by

$$M_l=\frac{1}{2}\begin{vmatrix}(1+S_l)\exp[-i(k_{l+1}-k_l)x_l] & (1-S_l)\exp[-i(k_{l+1}+k_l)x_l]\\ (1-S_l)\exp[+i(k_{l+1}+k_l)x_l] & (1+S_l)\exp[+i(k_{l+1}-k_1)x_l]\end{vmatrix}. \tag{2}$$

In Eq. (2) $S_l=m_{l+1}k_l/(m_lk_{l+1})$, and the effective masses and momenta are discretized as $m_l=m^*[(x_{l-1}+x_l)/2]$ and $k_l=k[(x_{l-1}+x_l)/2]$, respectively, x_l being the position of the lth boundary. If the metal–oxide interface is at x_0 and the oxide–silicon interface at x_N, then $m_o=m_M$ denotes an effective mass in the metal electrode and $m_N=m_{Si}$ an "effective" mass in silicon. For all other l we have $m_l=m_{ox}$. Because of the assumed parabolic dispersion within the oxide, the momentum takes the form

$$k(x)=\sqrt{2m_{ox}/\hbar^2}\sqrt{E-[\Phi_B+eF_{ox}x+E_{im}(x)]}, \tag{3}$$

there, with the image potential[25]

$$E_{im}(x)=-\frac{e^2}{16\pi\epsilon_{ox}}\sum_{n=0}^{\infty}(-\kappa)^n \times\left[\frac{1}{nd+x}-\frac{\kappa}{d(n+1)-x}+\frac{2\kappa}{d(n+1)}\right], \tag{4}$$

which includes the effect of all images in the two electrodes. In Eq. (4) κ is given by $\kappa=(\epsilon_{ox}-\epsilon_{Si})/(\epsilon_{ox}+\epsilon_{Si})$. The remaining quantities are the metal–SiO_2 barrier height for electrons Φ_B, the built-in potential drop over the oxide layer $eF_{ox}x$, and the dielectric constants ϵ_{ox} and ϵ_{Si} in oxide and silicon, respectively. Neglecting the image force, $T(E)$ can be written analytically in terms of Airy functions as was first done by Gundlach.[1] In our simulations we used the following parameters: d=2.8 nm, N=30, Φ_B=4.1 eV, $F_{ox}=V_{ox}/d=-0.071\,43$ V/nm (potential peak at the oxide–metal boundary), $m_{Si}=0.19m_o$, $m_M=m_o$, $\epsilon_{Si}=11.7$, and $\epsilon_{ox}=2.13$.

The significance of the "classical" image force in tunneling experiments was supported by Binnig *et al.*[26] They showed that it is indispensable in order to describe correctly the barrier-width dependence and absolute value of the vacuum tunnel current. The existence of image force effects for over the barrier transport in MOS structures was also shown recently by Wen *et al.*[18] On the other hand, the "classical" form can only be used asymptotically, i.e., a few Bohrs off the image plane. In the vicinity of that plane, the classical singularity has to be replaced by a self-consistent potential shape.[27] This shape is smooth throughout the interface and can be modeled by a smooth variation of the dielectric constant.[28] For simplicity, we have removed the singularity of the classical image potential in a more simplistic way by a straight continuation of both, the band edge in the semiconductor and the gate Fermi level.

The simulations involve further simplifications. Any possible band-structure mismatch at the Si–SiO_2 interface was disregarded. In above-barrier transitions electrons tunnel into highly excited states in the silicon near the Si–SiO_2 boundary. Here, the "effective" mass m_{Si} is merely a fitting parameter. Fortunately, m_{Si} only enters the preexponential factor of the TC and has no significant influence on the

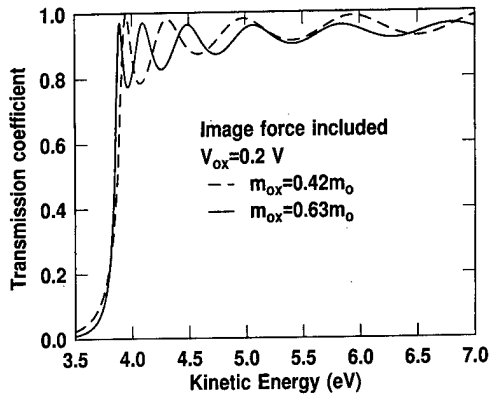


FIG. 4. Calculated transmission coefficients, image force effects included, for $m_{ox}=0.63m_0$ (solid line) and $m_{ox}=0.42m_0$ (dashed line), with the remaining parameters determined from experiment. The solid line was used to determine the best fit in Fig. 3.

interpretation of the measurements. The missing knowledge about the band structure of the ultrathin SiO_2 layer is covered by the "tunneling" mass m_{ox}. This parameter can be adjusted when the oscillations of the simulated TC are brought in coincidence with the periodicity of the measured current.

Results of the calculation for the selected parameters are shown in Fig. 4 (solid line). The transmission coefficient $T(E)$ exhibits initially a strong oscillatory structure whose period increases and whose amplitude decreases with energy. Although a barrier of $\Phi_B=4.1$ eV was used, the transmission threshold is closer to 3.9 eV. The decrease is consistent with experimental observations and predictions of image force theory, based on thicker oxides,[18] for the image force lowering expected for a 2.8 nm oxide and $V_{ox}=0.2$ V. *T(E)* is actually finite at energies <3.5 eV, which is a consequence of electrons tunneling through the rounded barrier near its top (see Fig. 1). The effect on the oscillatory structure of a change in the electron mass is illustrated by the dashed curve in Fig. 4, for which $m_{ox}=0.42m_o$, all other parameters remaining the same. A lower mass increases the effective wavelength of the electron for a given energy, which thereby requires a larger value to "fit" the wavelength into the cavity. Hence, the periodicity increases, as it also would by decreasing the cavity width (oxide thickness). The lighter electron mass also enhances the probability for the electron to tunnel through the top of the barrier, as indicated by the increased intensity of the tail in *T(E)* near 3.5 eV.

V. DISCUSSION AND CONCLUSIONS

To fit the theoretical transmission function to the experimental data either a corresponding function should be derived from experiment or the experimental collector current should be simulated. Neither is, at present, a realizable option due to the complexity of the transmission process of hot electrons through a MOS structure. Certain inelastic aspects of transmission across the oxide have been simulated by Monte Carlo calculations.[14,15] However, the role of transport across the metal, scattering at the interfaces, including details of the band structures, have generally been ignored. And this for good reason, as much of this information is neither available, nor calculable within the framework of present knowledge of the structurally incoherent system that the MOS structure represents.

As BEEM spectra in general exhibit little structure and obey, at least near threshold, a power-law dependence of the collector current I_c on V_T,[29] we can attempt to represent the BEEM spectra by multiplying the transmission coefficient with a simple quadratic power law, i.e., $I_c \propto (V_T - V_o)^2$. The result is shown for $V_o=3.9$ V and $T(E)$ calculated with $m_{ox}=0.63m_o$ and with image forces included in Fig. 3 as a solid line. As can be ascertained, an energy-independent $m_{ox}\approx 0.63m_o$ is sufficient for a reasonable fit over the entire voltage range. No attempt has been made to make the curves overlap, but merely to show the position of the structure. The sensitivity of the structure on m_{ox} (aside from that shown in Fig. 4) can be estimated from differentiating the quantization condition $E=(n\pi\hbar/d)^2/2m^*$. Thus, $\delta m_{ox}=-(\delta E/E)m_{ox}$, with a change in the peak location, let us say near 6 eV, of 0.1 eV resulting in a $\delta m_{ox}=0.01m_o$. Thus, a conservative estimate of the error gives a best-fit value of $m_{ox}=0.63\pm 0.02m_o$. In contrast, the error in d of ± 0.2 nm yields a substantially larger uncertainty of $\pm 0.09m_o$. Other uncertainties due to a lack of knowledge in $V_{ox}(<\pm 0.1$ V) or choice of ϵ_{ox} give uncertainties of $<0.01m_o$, so that our present best estimate for m_{ox} is $m_{ox}=0.63\pm 0.09m_o$.

Also shown in Fig. 3 is the dashed curve generated with a *T(E)* calculated with a trapezoidal barrier, i.e., with the omission of image force effects, but otherwise with identical parameters as the solid curve. The latter clearly gives a better fit to the data. Ignoring image force effects and optimizing the fit to the experimental data yield a $m_{ox}\approx 0.65m_o$. However, we find no physical reason to ignore image force effects, but merely show its role, having been motivated to show it by reason of a consistent historical neglect of image force effects by most practitioners of the art of electrical characterization of MOS structures.

The consequence of omission of image force effects is rather small on the magnitude of m_{ox}, however, its inclusion has a dramatic effect on the transmission coefficient in the presence of an oxide field. This is illustrated for relatively moderate fields in Fig. 5. The lowering of the threshold for increasing V_{ox} is clearly observable, with an accompanying expansion of the period for low n values (i.e., initial periodicity for $V_{ox}=0$ is less than that for $V_{ox}>0$). It is interesting to note that for $V_{ox}=1$ V the weak first peak corresponds to electrons partially tunneling through the top of the barrier before interference occurs (somewhat akin to the FN case). The rapid shifting of the structure for even moderate fields clearly suggests that in the presence of field inhomogeneities on the local sampling scale the observation of the interference effect would be quickly suppressed. In our experiment, positive charge near the anode is (randomly) generated, a conclusion also reached from other stressing experiments $(I-V)$ on MOS capacitor structures.[30] Even a single charge would generate a field of order 1 V/nm at the injecting point of the SiO_2 layer. Even at a distance of several nanometers

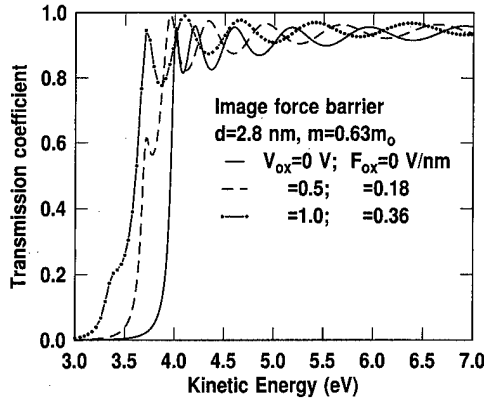


FIG. 5. Oxide field dependence of transmission coefficients calculated with the best-fit parameters, excepting V_{ox}. Image force effects are included.

away the resulting fields are comparable or larger than those in the simulations of Fig. 5. Because of scattering at the interfaces and in the SiO_2, electrons will experience somewhat different paths. Consequently, electrons injected locally near an area of charge would experience both a local- and time-dependent variation of the fields, which leads to a suppression of the interference effects, as was observed in our time-dependent experiments (Fig. 2). Thus, a condition for the observation of quantum interference is a region essentially free of oxide charge. The lateral extent of this region can only be guestimated from our observation that we needed to move at least 10 nm away from any point of prior exposure (and hence, positive charge buildup) before we could observe a new oscillatory structure in the BEEM spectrum. The fact that a virgin area did not necessarily lead to the observation of quantum interference suggests that the presence of local charge was quite pervasive. Since device quality oxides have generally quite low densities of defects, traps, etc. ($<10^{11}/cm^2$), it appears somewhat surprising that we do not observe quantum interference most of the time. Two reasons for this failure may be sited: (a) Pd metallization induces a large ($\sim 10^{13}/cm^2$) density of electron traps,[17] and (b) multiple scattering in the metal and SiO_2 layer can effectively broaden the sampling area. Although no net electron charge has been observed in the thin layers used here, electrons can nevertheless momentarily be trapped before leaking out by tunneling to the electrodes.[21] A trapped electron acts as a scattering center and distorts the field sensed by the other electrons passing in its vicinity.

We have shown here that direct electron injection into a laterally confined area of a MOS structure can lead to the observation of strong quantum interference oscillations. A theoretical analysis yields an effective conduction-band mass of $0.63m_o$, and further reveals an extreme sensitivity of the oscillatory structure to the oxide thickness and the oxide field. The technique has, thus, the potential of an extremely sensitive local probe to address issues of local structural and electric homogeneity, issues of great importance in the area of future ultrasmall devices. It is to be expected that the use of polygates with their drastically lower density of trap states would considerably enhance the observation of quantum interference in thin oxides, and thereby facilitate the realization of this technique as a potentially powerful local probe to assess dielectric quality.

[1]K. H. Gundlach, Solid-State Electron. **9**, 949 (1966).
[2]J. Maserjian and G. P. Peterson, Appl. Phys. Lett. **25**, 50 (1974).
[3]J. Maserjian, J. Vac. Sci. Technol. **11**, 996 (1974).
[4]M. V. Fischetti, D. J. DiMaria, L. Dori, J. Batey, E. Tierney, and J. Stasiak, Phys. Rev. B **35**, 4044 (1987).
[5]J. C. Poler, K. K. McKay, and E. A. Irene, J. Vac. Sci. Technol. B **12**, 88 (1994).
[6]S. Zafar, K. A. Conrad, Q. Liu, E. A. Irene, G. Hames, R. Kuehn, and J. J. Wortman, Appl. Phys. Lett. **67**, 1031 (1995).
[7]S. Zafar, Q. Liu, and E. A. Irene, J. Vac. Sci. Technol. A **13**, 47 (1995).
[8]H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguru, S. Nakamura, M. Saito, and H. Iwai, IEEE Trans. Electron Devices **45**, 1233 (1996).
[9]T. Yoshida, D. Imafuku, J. L. Alay, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys., Part 2 **34**, L903 (1995).
[10]Z. A. Weinberg, J. Appl. Phys. **53**, 5052 (1982).
[11]Reference 10 lists earlier results with image force effects included.
[12]A. Schenk and G. Heiser, J. Appl. Phys. **81**, 7900 (1997).
[13]M. V. Fischetti, S. E. Laux, and E. Crabbé, J. Appl. Phys. **78**, 1058 (1995), and references therein.
[14]R. Ludeke, A. Bauer, and E. Cartier, Appl. Phys. Lett. **66**, 730 (1995); J. Vac. Sci. Technol. B **13**, 1830 (1995).
[15]R. Ludeke, H. J. Wen, and E. Cartier, J. Vac. Sci. Technol. B **14**, 2855 (1996).
[16]B. Kaczer and J. P. Pelz, J. Vac. Sci. Technol. B **14**, 2864 (1996).
[17]H. J. Wen and R. Ludeke, J. Vac. Sci. Technol. B **15**, 1080 (1997).
[18]H. J. Wen, R. Ludeke, D. M. Newns, and S. H. Low, J. Vac. Sci. Technol. A **15**, 784 (1997).
[19]N. D. Lang, A. Yacoby, and Y. Imry, Phys. Rev. Lett. **63**, 1499 (1989); N. D. Lang (private communication).
[20]C. P. D'Emic (private communications).
[21]H. J. Wen and R. Ludeke, J. Vac. Sci. Technol. A **16**, 1735 (1998).
[22]R. Ludeke, E. Cartier, and H. J. Wen (unpublished results). The transmission probability here is the one associated with inelastic scattering events due to electron–phonon interactions, see Ref. 14 for further details.
[23]E. Cartier, D. Arnold, D. J. DiMaria, M. V. Fischetti, P. Braunlich, S. C. Jones, X. A. Shen, R. T. Casper, and P. J. Kelly, Rev. Solid State Sci. **5**, 531 (1991).
[24]Y. Ando and T. Itoh, J. Appl. Phys. **61**, 1497 (1987).
[25]M. Kleefstra and G. C. Herman, J. Appl. Phys. **51**, 4923 (1980).
[26]G. Binnig, N. Garcia, H. Rohrer, J. M. Soler, and F. Flores, Phys. Rev. B **30**, 4816 (1984).
[27]P. A. Serena, J. M. Soler, and N. Garcia, Phys. Rev. B **34**, 6767 (1986).
[28]F. Stern, Phys. Rev. B **17**, 5009 (1978).
[29]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988); L. D. Bell and W. J. Kaiser, *ibid.* **61**, 2368 (1988).
[30]D. A. Buchanan, D. J. DiMaria, C.-A. Chang, and Y. Taur, Appl. Phys. Lett. **65**, 1820 (1994).

Investigation of ultrathin SiO_2 film thickness variations by ballistic electron emission microscopy

B. Kaczer,[a] H.-J. Im, and J. P. Pelz
Department of Physics, The Ohio State University, Columbus, Ohio 43210

(Received 21 January 1998; accepted 21 May 1998)

We investigate the feasibility of using ballistic electron emission microscopy (BEEM) to study possible thickness variations of ultrathin SiO_2, which might exist at substrate defects, such as steps. We find that simple BEEM imaging of the oxide film sandwiched into a metal–oxide–semiconductor (MOS) structure does not reveal any features that could be related to the oxide film. We further present initial results suggesting that hot-electron resonance in the oxide conduction band could be observed by BEEM and could be sensitive to local film thickness. We also confirm the ability of oxide film to sustain injection of very high densities of hot electrons without any observable damage. In some cases we observe local damage of the MOS structure induced by BEEM measurements, but we conclude that it is most likely related to failure of the metal overlayer and probably not related to hot-electron breakdown of the oxide. © *1998 American Vacuum Society.* [S0734-211X(98)07804-4]

I. INTRODUCTION

Although alternative materials are being considered for metal–oxide–semiconductor (MOS) field-effect-transistor gate insulation, silicon dioxide (SiO_2) so far remains the insulator of choice, and is likely to remain so for the next several MOS technology generations below the 0.25 μm design rule. Devices scheduled into production over the next several years will require oxide film thickness around 3 nm and below.[1] Direct tunneling through the oxide at these thicknesses will become considerable. Although direct tunneling may be "built in" into the device characteristic,[2] hot-electron flux could influence the device reliability and lifetime. Moreover, any nonuniformities of ultrathin oxide films, due to, for example, underlying Si substrate steps, could constitute a relatively substantial thickness change. Since the tunneling probability through such an oxide film can change by an order of magnitude if its thickness is changed just by a faction of a nm,[3] variation in the thickness of these very thin films could further adversely influence the device performance.

The initial idea behind studying the ultrathin, 2.7 nm oxides was to test whether ballistic electron emission microscopy (BEEM) can be used to study possible thickness variations in the oxide. In this article we report on initial BEEM measurements on these ultrathin films. We find that the magnitude of the BEEM current, modulated by local properties of the overlaying metal layer, does not show any obvious variations related to the oxide film. On the other hand, we observe faint oscillations in some BEEM spectra, which, as we try to demonstrate, could be related to interference of hot BEEM electrons with the oxide film. These oscillations could be potentially used in the future to precisely measure the local thickness of ultrathin films.

In addition to this, we occasionally observe dramatic local damage of the MOS structure caused by the BEEM probe. Since hot electrons are injected into the structure prior and during the actual failure, a possibility exists that hot-electron-related breakdown of the oxide is taking place. We present evidence suggesting that this damage is most likely related to the failure of the metal overlayer. In contrast to that, we are able to locally inject very large densities of very energetic electrons into the oxide without observing any physical damage to the MOS structure.[4]

II. EXPERIMENT

A. Ballistic electron emission microscopy

BEEM, an extension of scanning tunneling microscopy (STM),[5] was first used by Kaiser and Bell to investigate the microscopic properties of metal–semiconductor Schottky barriers.[6] The application of this technique to MOS structures has been described previously.[7–9] The schematic configuration of the BEEM experiment is shown in Fig. 1(a).[7–9] Electrons, injected from the STM tip into the metal overlayer, are peaked in energy near the Fermi energy of the tip.[10] The STM tip thus serves as a highly spatially localized tunable source of hot electrons of a very high current density.

Due to the small thickness of the metal overlayer, a nonnegligible fraction of these hot electrons is able to traverse the metal layer and reach the metal–oxide interface *ballistically*, i.e., without losing energy,[6] and enter the oxide provided they have enough energy (controlled by V_T) to surmount the barrier at the metal–oxide interface. Some of these electrons may conduct across the oxide into the Si substrate and be measured as the external collector current I_c. By measuring a I_c–V_T curve we can, therefore, determine the *local* maximum barrier height, with the barrier height simply corresponding to the turn-on in the curve at the threshold voltage V_{th}.[6–9] When the tip is scanned across the sample at above-threshold voltages, BEEM current I_c can be recorded as a function of position. Such a BEEM image then could carry information about the lateral homogeneity of the MOS structure.[7–9]

[a]Electronic mail: kaczer@mps.ohio-state.edu

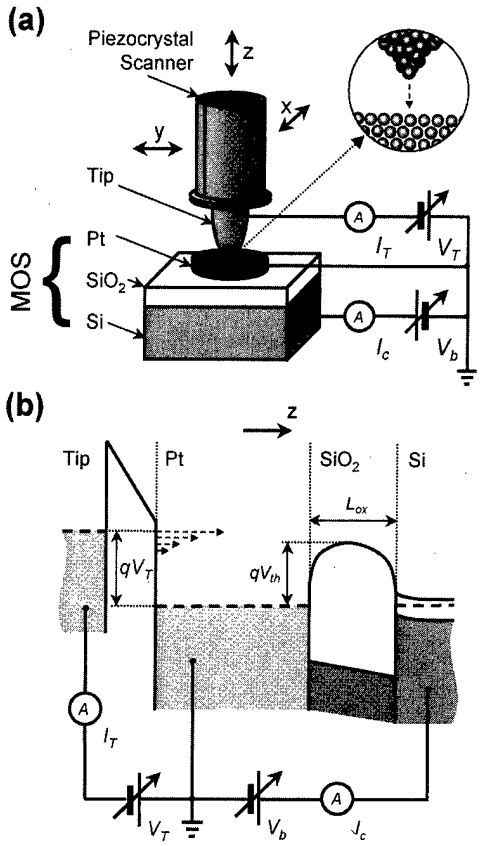


FIG. 1. (a) Schematic configuration of BEEM experiment on MOS structures and (b) the corresponding energy-level diagram, with the apparent oxide barrier height qV_{th}. Hot electrons of adjustable energy (corresponding to tip voltage V_T) and high flux density can be injected into the structure from the STM tip that can be positioned over the structure with nm-scale precision. The electrons traversing the oxide film are registered as collector current I_c.

B. Tip locking

It is well known in the scanning probe microscopy community that the STM tip can slowly drift with respect to the sample. This motion could be due to subtle temperature changes in the apparatus and/or to the tip piezocrystal scanner creep. This tip/sample drift can be, typically, units of nm per minute in our setup, and typically, diminishes slowly over a period of hours. In some of the experiments described below, however, it was necessary to perform the BEEM measurements over one well-defined location of the sample. We, therefore, devised a technique that allowed us to "lock" the tip onto a local extreme in the sample topography and follow that exact position throughout the measurement (up to about 20 h—the duration of the longest injection experiment). A similar technique has been used, for example, for "atom tracking" by others.[11]

We have chosen and tested a simple algorithm for the tip locking: after the z position [see Fig. 1(a)] of the tip in the constant current mode is evaluated, the tip is moved by about 0.3 nm in the x direction [parallel with the sample surface, see Fig. 1(a)], where the z position is evaluated again. Depending on the sign of the height (z position) difference, the tip proceeds in the same or the reverse direction in the next step. The same procedure is alternately performed in the y direction. This way, the STM tip seeks the nearest maximum or minimum (depending on the height difference condition) in the sample topography and locks on this extreme, even if the sample is drifting underneath the tip.

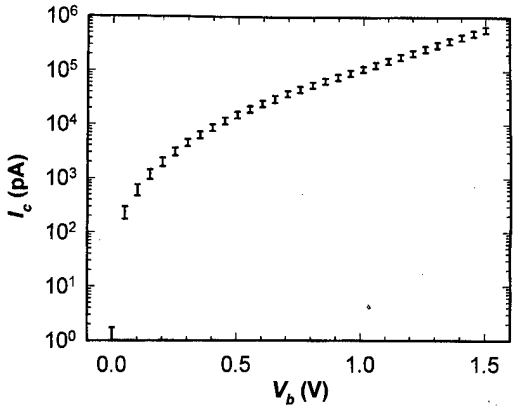


FIG. 2. Average of I_c-V_b curves measured on 19 different Pt dots documenting the quality of the MOS structures.

C. Sample preparation

Samples used in our experiments were prepared from n-Si(100) wafers with nominally 2.7 nm of high-quality oxide (courtesy of Texas Instruments, Inc.). Two samples, sized $\sim 5 \times 20$ mm^2, were cut from the wafer and dusted off in a stream of nitrogen. One sample was then used in standard atomic force microscopy (Digital Instruments Nanoscope AFM) imaging of the top oxide surface, the other sample without any further cleaning was introduced through an air-lock into the UHV preparation chamber, where it was outgassed by direct current heating at 350–400 °C for 10 min. Then, a number of ~4-nm-thick, 0.5-mm-diam Pt dots were deposited onto the sample through a shadow mask from a 4 kV e-gun evaporator. The complete sample was then transported under UHV into the STM/BEEM chamber for measurements. A dot on the sample was then selected by grounding it gently with a 0.1-mm-diam Au wire positioned with a mechanical manipulator inside the STM chamber. Macroscopic I_c-V_b curves measured on all dots were reproducible from dot to dot, and in agreement with the measurements in Ref. 3 (after proper area scaling, see Fig. 2). No faulty or leaky dots were observed. Finally, the STM tip was approached to one grounded dot to perform BEEM measurements.

III. RESULTS AND DISCUSSION

A. Basic measurements

The ATM topograph of the top oxide surface in Fig. 3 reveals a steplike structure of the surface, with the steps separated roughly by 130–140 nm. In addition to that, the AFM scan shows pits of varying depth (typically, 1–2-nm-

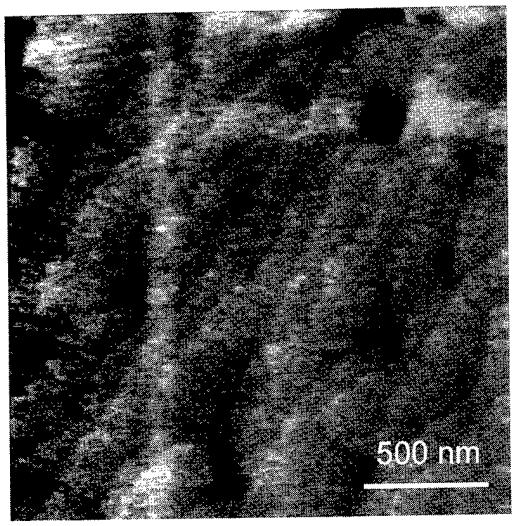


FIG. 3. AFM topograph of the top surface of the oxide film. Steps are visible, separated on average by 130–140 nm. The large pits are coated with the oxide film.

deep, but a pit ~10-nm-deep has been observed). These pits were created in the substrate accidentally during substrate cleaning prior to oxidation.[12]

A typical BEEM spectrum, shown in Fig. 4(a), represents an average of 25 BEEM I_c-V_T curves taken sequentially at a single location of the MOS sample (stabilized by locking the STM tip to a local topographical minimum). The threshold voltage V_{th}, corresponding to the barrier height in the oxide, is about 3.9 V. When the tip voltage V_T is held constant at 10 V and the I_c-t curve is recorded [Fig. 4(b)], a strong *enhancement* in the current I_c of hot electrons across the oxide is observed in most cases. This is in contrast to the *suppression* of I_c observed on thicker oxide films (compare with Fig. 3 of Ref. 8), which we have previously shown to be linked to a buildup of negative charge in the oxide.[7] One possible mechanism for the enhancement, suggested by Wen and Ludeke, is related to a buildup of positive charge at the SiO_2/Si interface. This should be accompanied by a decrease of the barrier in the oxide due to image force lowering effects [see Fig. 1(b)]. In order to quantify this effect, we evaluate the threshold voltages of BEEM spectra taken sequentially at a single location, but all our attempts to determine this effect were inconclusive, due to relatively large noise in the collector current I_c (a lowering of $\sim 0.1 \pm 0.2$ V is measured following injection of ~10 nC of collected charge). For comparison, the decrease of 0.1 V would correspond to a positive surface charge density of $\sim 10^{14}$ cm^{-2} for this sample, suggesting that BEEM is fairly insensitive to trapped charge at the SiO_2/Si interface.

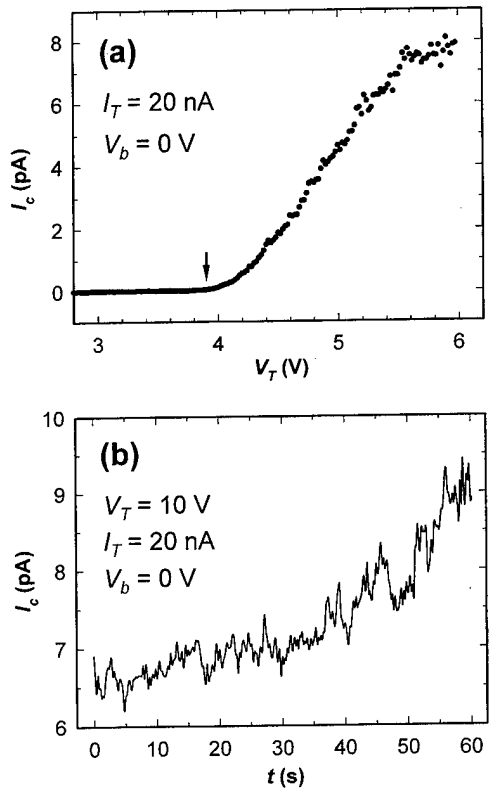


FIG. 4. (a) BEEM spectrum, consisting of an average of 25 I_c-V_T curves, taken sequentially at one location of the sample at $I_T=20$ nA and $V_b=0$ V. The threshold voltage V_{th} is ~3.9 V. (b) I_c-t curve, taken at $V_T=10$ V, $I_T=20$ nA and $V_b=0$ V illustrates the *enhancement* of the BEEM current I_c with hot-electron injection.

B. BEEM imaging

A typical large-area BEEM scan at $V_T=5.5$ V and $I_T=10$ nA is shown in Figs. 5(a) and 5(b). Some faint, large-scale contrast in the topography is visible, possibly related to the steps on the underlying oxide film. The nodular contrast in the BEEM image reflects the varying thickness of the overlaying Pt film (Pt nodules), but no contrast is apparent that could be associated with possible oxide thickness variation. This is even more pronounced when a scan is taken over an edge of a pit [Figs. 5(c) and 5(d)]. The pit itself is clearly visible in the topography, but no contrast related to the pit edges is apparent in the BEEM image.

It is, however, uncertain how much can be concluded about the oxide film uniformity from these measurements. It is possible that the transmission probability through the SiO_2 conduction band is only a weak function of the film thickness. Since the BEEM electrons *conduct* across the oxide film (instead of tunneling through a barrier), small variations in local film thickness only affect this transport indirectly, by varying the local electric field in the oxide, or slightly changing the total distance the electrons must cross. In principle, the tip voltage could be lowered to allow BEEM electrons to tunnel *through* the barrier, where the electron transmission probability through the oxide depends strongly on the oxide thickness. The flux of ballistic electrons, however, is too small to produce a measurable signal in this regime, and therefore, this most natural mode of measuring variations in oxide thickness could not be used.

C. Local damage of MOS structure

Interestingly, these large-area scans were intermittently interrupted by random crashes of the STM tip, usually ac-

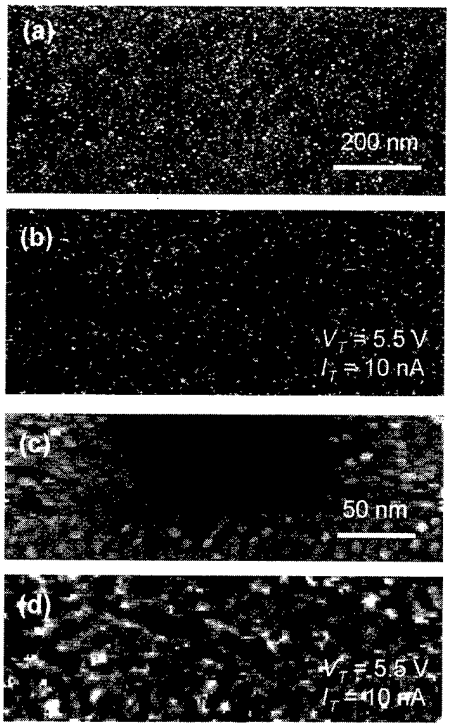


FIG. 5. (a) and (b), a section of a large-area scan. Topography of the top metal surface (a) reveals its nodular character, which is reflected in the corresponding BEEM image. (b) No contrast related to the oxide film is discernible. (c) and (d), a section of a scan over a pit edge. The pit itself is clearly visible in the topography (c), but no contrast related to the pit edges is apparent in the BEEM image (d).

companied by a deep crater (~10 nm or more, see Fig. 6), with high BEEM transmittance through the structure (equal to unity at $V_T = 5.5$ V) and a low threshold voltage (~1 V) inside the crater. The MOS structure, however, was not shorted by this effect. It was not immediately obvious whether these crashes are related to failure of the STM tip or the metal overlayer, or are due to a more fundamental failure within the oxide film due to hot-electron injection. Hot-electron injection into oxide films has been known to produce pronounced physical damage to MOS structures, although these effects are of different character and usually seen in thicker oxides only.[13]

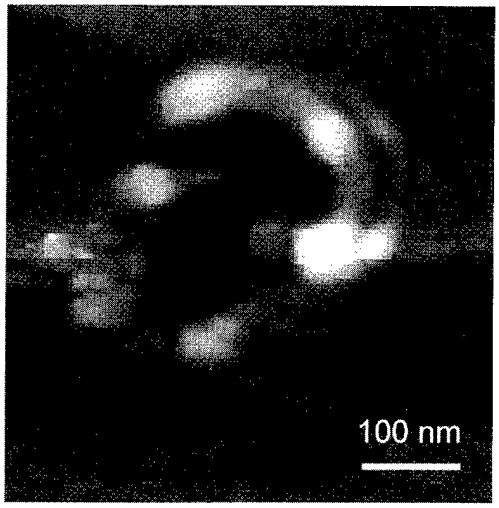


FIG. 6. STM scan of the region damaged by an apparent STM tip crash. The depth of the crater is about 11 nm with respect to the flat top metal surface. The sloped left side of the crater wall is an artifact related to insufficient speed of the STM feedback loop.

We, however, find the tip crash frequency to be related inversely to the STM scanning speed: when the tip scanning speed was kept below a given value throughout the whole scan, we were able to take many large-area scans at $V_T = 10$ V without a single crash. During a typical, $0.5 \times 0.5\ \mu m^2$ scan, ~80% of the scanned area received 150–200 C cm^{-2} of ~6 eV ($V_T = 10$ V) electrons. No physical damage to the MOS structure was *ever* observed, even if large bias V_b (up to 3 V, i.e., ~7–9 MV cm^{-1}) was applied across the structure. The energy of injected electrons of ~6 eV for this oxide thickness corresponds to the applied field of about 22 MV cm^{-1} in an equivalent Fowler–Nordheim experiment.[4]

In addition to that, very long (~20 h) injections of ~6 eV electrons were performed at one location of the sample. Tip locking was used to maintain the tip at the same location. During such super-long injections, several hundreds of nC of total charge were seen, typically, going through the oxide, presumably over an area ~10 nm across. The area was deduced from the enhancement in a BEEM image taken after an injection, such as the one shown in Fig. 4(b). Even after the injection of ~10^5 C cm^{-2} of these very high-energy electrons, no physical damage to the oxide was observed.

From this, we conclude that failure of the metal film is the most likely mechanism for the creation of these craters. If the failure of the oxide were the case, the probability of which should be proportional to the injected charge density, we would expect the crash rate to go *up* with lower scanning speed. In a possible scenario, the part of the metal overlayer directly underneath the tip becomes electrically discontinued from the rest of the film due to modifications to the metal induced by the tip getting too close to the sample. The STM feedback loop then pushes the tip into the structure, until sufficient tunneling current is restored by injecting electrons into the Si substrate. This mechanism involving the metal overlayer failure is further supported by the fact that at $V_T = 5.5$ V, when the tip is closer to the sample, occasional crashes were still observed, although their occurrence diminished when the scanning speed was lowered. In addition to that, we observed that by driving the tip against the sample by applying a short pulse to the z piezo, which we routinely use to "sharpen" the STM tip, frequently resulted in similar damage to the MOS structure. This occurred at V_T as low as ~3.5 V, i.e., when the hot-electron energy is insufficient for entering the oxide conduction band. Furthermore, from studying the topography and the BEEM signal traces, we did not see any obvious increase of the BEEM current just prior to the crash, as we would expect in the case of the oxide film failure.

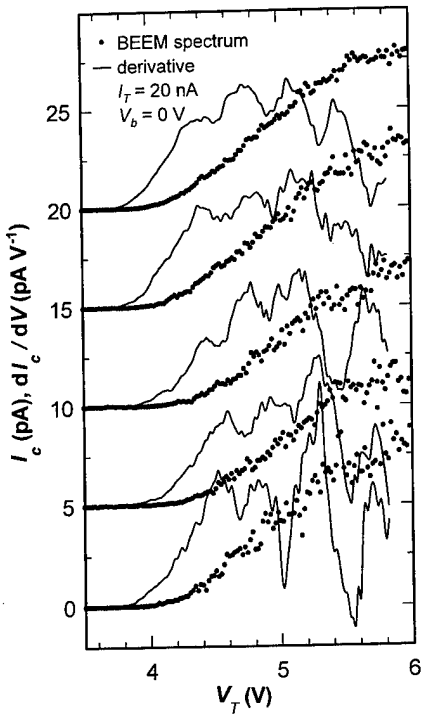


FIG. 7. Five BEEM spectra taken at different locations of the sample. Each spectrum constitutes an average of 7–25 I_c-V_T curves taken successively at the same location at $I_T=20$ nA and $V_b=0$ V. First derivatives of the BEEM spectra reveal weak oscillations in the spectra, suggesting possible BEEM electron resonance in the SiO_2 film. Derivative peaks seem to be shifted from spectrum to spectrum, possibly due to thickness variations of the oxide film.

D. Local hot-electron resonance

As stated above, the BEEM hot-electron flux is insufficient to produce a collector current signal by tunneling through the oxide film, and therefore, cannot be used to locally measure the oxide thickness. We, therefore, investigated the theoretical possibility for estimating local thickness variations of the ultrathin oxide films based on measuring possible hot-electron resonance in the *conduction band* of the film. Depending on the mean-free path of hot electrons in the oxide conduction band, it is possible that the small thickness of the ultrathin oxide allows some hot electrons to traverse the oxide without scattering, i.e., stay coherent throughout the thickness of the oxide. Such electrons could then, depending on their energy, constructively or destructively interfere with the oxide film. Electron transmittance through the oxide should then become an oscillatory function of energy, which in turn should be observable in the BEEM spectra. Such behavior was observed by Maserjian and others.[14,15]

In Fig. 7 we show several BEEM spectra taken at different locations of the sample. Each spectrum constitutes an average of 7–25 I_c-V_T curves taken successively at the same location maintained by actively locking the tip to a local minimum. Pre- and post measurement scans were taken to check that the STM tip position was held throughout the I_c-V_T curve taking. When we calculate the first derivative of the BEEM spectra in Fig. 7, weak oscillations in the spectra are revealed. This oscillatory behavior is found at about half of the locations were I_c-V_T curves are taken. We also notice that the first several I_c-V_T curves, out of the sequence of 25 curves taken at each location, show the most pronounced oscillations. With the increasing number of I_c-V_T curves taken at one location, this oscillatory behavior tends to go away. We explain this by possible changes to the oxide layer caused by hot-electron injection, the possibility of which is documented in Fig. 4(b).

The oscillatory behavior of the BEEM spectra is very suggestive of the hot BEEM electron interference with the ultrathin oxide film. To investigate this possibility, we compare the measurements with our own theoretical calculations. We approximate the MOS structure with the oxide conduction band in the middle by a "square" potential barrier with the transmission probability over this barrier given by a well-known formula.[16] The positions of the maxima of this probability are then given by

$$[2m^*_{ox}q(V_{T,n}-V'_0)]^{1/2}L_{ox}=\pi\hbar n, \tag{1}$$

where m^*_{ox} is the effective electron mass in the oxide, $V_{T,n}$ is the tip voltage corresponding to the nth maximum, and qV'_0 is the apparent height of the barrier in the oxide, as determined in Fig. 4 (i.e., equal to ~3.9 eV). L_{ox} is the oxide film thickness (2.7 nm) and q is the elementary charge.

We compose the theoretical BEEM spectrum from the oscillatory part and a smoothly varying background. Here, we assume that most hot electrons in the oxide will undergo some scattering and will not interfere with the film, contributing to the background only, and only a smaller part of BEEM electrons will remain coherent throughout the thickness of the oxide. The oscillatory part is calculated using a standard BEEM calculation,[17] with the above-mentioned transmission probability added to it.

A result of this calculation is shown in Fig. 8(a). This spectrum is calculated assuming constant $m^*_{ox}=0.8\,m_e$, an average value suggested by Maserjian and others.[14,15] The shape of the first derivative is reminiscent of the BEEM spectra derivatives shown in Fig. 7, although the theoretical peak positions and their spacing do not precisely match the experiment. The mismatch is most obvious for the peaks at lower V_T. From inspection of Eq. (1) we see that if the effective electron mass m^*_{ox} were energy dependent, it would be possible to match the experimental observation better. We note that a varying effective oxide electron mass was, in fact, suggested, for example, by Maserjian.[14] We, therefore, choose a varying m^*_{ox} according to the following prescription: $m^*_{ox}=0.42\,m_e$ at the bottom of the conduction band, in line with Maserjian's suggestion, and then slowly increasing to m_e over the range of 3 eV. The resulting BEEM spectrum is shown in Fig. 8(b). The derivative peak positions are in better agreement with the positions of the experimental spectra in Fig. 7.

We also notice that all peak positions of all BEEM spectra in Fig. 7 do not actually line up (peaks are shifted toward higher voltages for some spectra). One possible explanation of this could be that the oxide thickness is not strictly uni-

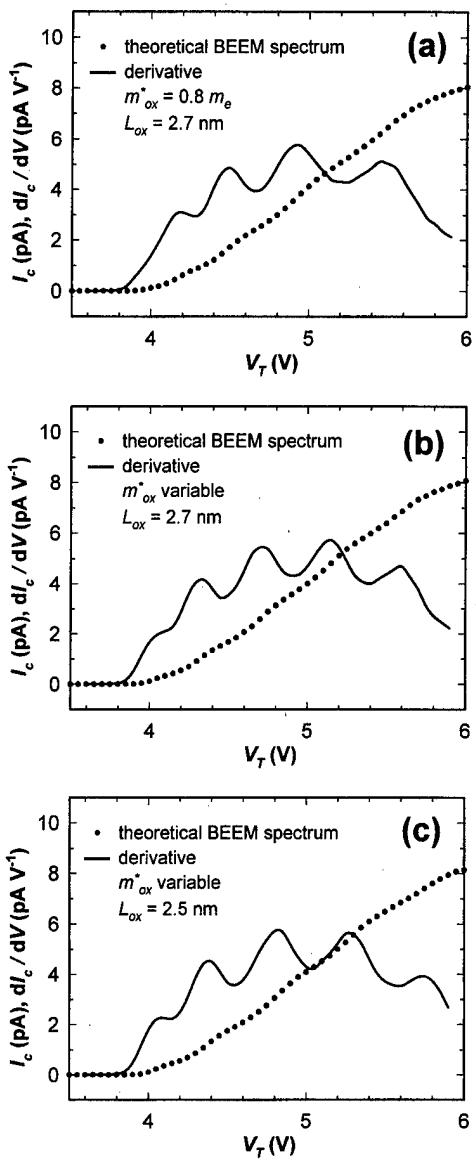


FIG. 8. Theoretical BEEM spectra assuming quantum-mechanical resonance of BEEM electrons in the oxide film. (a) Calculation assuming a constant effective electron mass m^*_{ox} in the oxide. The shape of the first derivative is reminiscent of the BEEM spectra derivatives shown in Fig. 7, although the theoretical peak positions and their spacing do not precisely match. (b) Calculation assuming a varying effective electron mass m^*_{ox}. Derivative peak positions are in better agreement with the experiment in Fig. 7(c). Calculation assuming a varying effective electron mass m^*_{ox} and smaller oxide thickness L_{ox}. The derivative peak positions are shifted upward in voltage, demonstrating the effect of oxide film thickness variation. Similar shifts are visible from different experimental BEEM spectra in Fig. 7.

form (this, of course, assumes that m^*_{ox} does not change with position). For comparison, we show a theoretical spectrum calculated with for L_{ox}=2.5 nm, which shows that peak positions shifted by ~0.25 V (at higher V_T), is approximately consistent with observed shifts between different BEEM spectra. This suggests that this type of measurement could, in principle, be sensitive to local oxide thickness variations of the order of 0.2 nm or less.

IV. CONCLUSIONS

We performed initial BEEM measurements on ultrathin oxide film to investigate the feasibility of studying possible thickness variations in the film. We conclude that while BEEM imaging does not reveal any contrast that can be linked to the oxide film, a certain potential exists for probing the ultrathin oxide thickness variations by evaluating hot-electron resonance peaks in the BEEM spectra. In this article we argue that oscillations caused by hot-electron resonance in the conduction band of the SiO_2 film may occur. We demonstrate that with some experimental care the oscillations can actually be observed in the BEEM spectra, which are in reasonable agreement with the theoretical expectations for hot-electron resonance. We further show that varying m^*_{ox}, in line with some predictions, improves the agreement between theory and the experiment. Obviously, more work is required to test these observations and to see if they can be used to map out local oxide film properties.

As part of our initial work on ultrathin oxide films we also confirm the possibility of injection of very high charge densities of very hot electrons into the oxide film without damaging the MOS structure on one hand.[4] On the other hand, we trace the origins of the BEEM-induced damage of the MOS structure and are able to reduce the problems related to it.

ACKNOWLEDGMENTS

The authors wish to thank Dr. R. M. Wallace of Texas Instruments, Inc., for his support and for helpful discussions. This work was supported by ONR and Texas Instruments, Inc.

[1] *The National Technology Roadmap for Semiconductors* [Semiconductor Industry Association (SIA), 1997], found at http://notes.sematech.org/roadmap5.pdf
[2] H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, IEEE Trans. Electron Devices **43**, 1233 (1996).
[3] B. Brar, G. D. Wilk, and A. C. Seabaugh, Appl. Phys. Lett. **69**, 2728 (1996).
[4] R. Ludeke, H. J. Wen, and E. Cartier, J. Vac. Sci. Technol. B **14**, 2855 (1996).
[5] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).
[6] W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988); **61**, 2368 (1988).
[7] B. Kaczer, Z. Meng, and J. P. Pelz, Phys. Rev. Lett. **77**, 91 (1996).
[8] B. Kaczer and J. P. Pelz, J. Vac. Sci. Technol. B **14**, 2864 (1996).
[9] R. Ludeke, A. Bauer, and E. Cartier, J. Vac. Sci. Technol. B **13**, 1830 (1995).
[10] E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford University Press, New York, 1989).
[11] M. Krueger, B. Borovsky, and E. Ganz, Surf. Sci. **385**, 146 (1997).
[12] R. M. Wallace and G. D. Wilk (private communication).
[13] M. Mankos, R. M. Tromp, M. C. Reuter, and E. Cartier, Phys. Rev. Lett. **76**, 3200 (1996).
[14] J. Maserjian, J. Vac. Sci. Technol. **11**, 996 (1974).
[15] S. Zafar, K. A. Conrad, Q. Liu, E. A. Irene, G. Hames, R. Kuehn, and J. J. Wortman, Appl. Phys. Lett. **67**, 1031 (1995).
[16] For example, C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics* (Wiley, New York, 1997).
[17] L. D. Bell, Phys. Rev. Lett. **77**, 3893 (1996).

Compliant substrate technology: Status and prospects

April S. Brown[a)]
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

(Received 21 January 1998; accepted 11 May 1998)

Compliant substrates offer a new approach to strain management in lattice-mismatched structures. The role of the compliant substrate is to reduce the strain in a mismatched overlayer by sharing the strain via deformation of the substrate, or by nucleating and confining defects in the substrate. This can be accomplished by using a thin, "free-standing" substrate. Current efforts are primarily focused on the specification, design, and fabrication of an "ideal" compliant substrate. Key processing issues include the etching of the substrate and the bonding of the substrate to a mechanical handle wafer. Dramatically reduced dislocation densities, as well as reduced relaxation rates, have been observed for the growth of mismatched overlayers on compliant substrates.
© *1998 American Vacuum Society.* [S0734-211X(98)05004-5]

I. INTRODUCTION

Materials scientists and engineers have long addressed the need to efficiently integrate devices fabricated from dissimilar materials. This goal has been driven in part by economic factors, in particular the relatively low cost of Si, as compared to GaAs and InP. In addition, the integration of compound semiconductor components with Si circuits enables systems with a higher degree of functionality.

The lattice and thermal mismatches between arbitrary materials are the primary challenges to integration. For the direct growth of materials on an arbitrary substrate, these mismatches lead to the production of deleterious defects, primarily high densities of threading dislocations. Schemes have been developed, such as the growth of graded lattice constant buffer layers, to minimize the dislocation density in the vicinity of the active device. However, these schemes offer minimal and, in most cases, inadequate improvements with respect to required device performance and reliability.

Another approach to the integration of dissimilar materials is the physical merging, or bonding, of device structures, either prior to, during, or after fabrication. Wafer bonding technology continues to be developed to meet this goal, and ultimately this approach may have great utility. However, the use of wafer bonding for material and/or device integration requires significant modification of processes. In contrast, an integration solution that can be merged with existing fabrication approaches is ideal. The compliant substrate approach, described in detail below, potentially offers such a solution. This approach looks to the modification of the substrate prior to growth in such a way that defects produced necessarily by lattice and thermal mismatches are confined in the substrate. This technology also should enable the production of materials that do not possess a lattice-matched substrate, such as GaN, with a higher quality than on standard mismatched substrates.

[a)]Electronic mail: april.brown@ee.gatech.edu

II. DEVELOPMENT OF COMPLIANT SUBSTRATE CONCEPT AND TECHNOLOGY

Lo first presented the advantages offered by a *compliant* or thin (on the same order thickness as an epitaxial film) substrate[1] in 1991.[2] Specifically, Lo proposed that a thin, "free-standing" substrate offers a reduction in the concentration of defects in a mismatched overlayer due to (1) a larger effective critical thickness, and (2) a smaller dislocation image force. The critical thickness is increased due to the partitioning of the strain between the overlayer and the substrate. Using an equilibrium argument, Lo relates the effective critical thickness (t_{eff}) to the conventional critical thickness (t_c) and the substrate thickness (t_s) by

$$1/t_{\text{eff}} = 1/t_c - 1/t_s . \tag{1}$$

If the substrate is thinner than the conventional critical thickness, then the effective critical thickness is infinite. Likewise, if the substrate is thinner than the epitaxial film, then the image force tends to pull dislocations down into the substrate, introducing a "dislocation gettering" mechanism. The adequacy of this model in describing observed behavior has yet to be determined. A primary impediment to the simple understanding of experimental results has to do with the difficulty in achieving an ideal compliant substrate. A summary is given below of different compliant substrate implementations and results obtained from mismatched growths on these substrates.

III. COMPLIANT SUBSTRATE IMPLEMENTATIONS

Figure 1 shows a schematic of a compliant substrate system, consisting of the thin substrate, the bond layer, and the mechanical handle wafer. A generalized process flow for a GaAs compliant substrate is also shown. Before summarizing the various compliant substrates that have been proposed or produced, it is useful to consider the perceived criteria for achieving an ideal complaint substrate.[3]

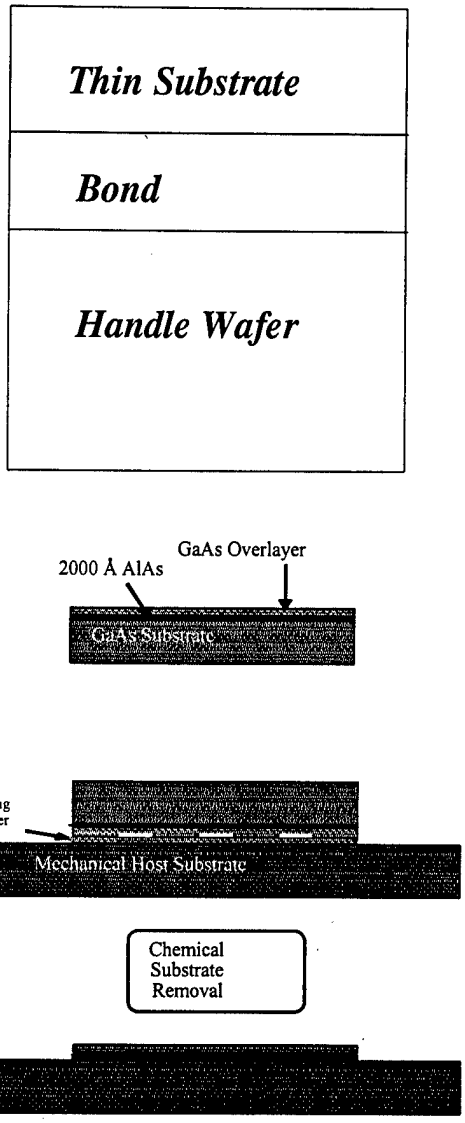


FIG. 1. Generic diagram and process flow for a compliant substrate.

A compliant substrate should

(1) be as thin as possible (for most large misfit systems, this means less than 10 nm);
(2) possess a bond to a mechanical handle wafer that has low viscosity at low temperatures (to allow substantial plastic flow);
(3) be comprised of materials that are bondable and have good strength;
(4) be comprised of materials that are stable to heat and water.

The first implementation of a GaAs compliant substrate was a membrane supported at each corner by a pedestal. The membrane was square with a typical size of tens of microns per side.[4] A 200 nm thick InGaAs layer (with 1% strain) was grown on a 80 nm thick membrane. The conventional critical thickness for the InGaAs is 120 nm. While dislocations were observed as a cross-hatched morphology in the InGaAs film on the conventional substrate, a smooth morphology was observed for the film on the membrane. In addition, x-ray diffraction measurements indicated that the InGaAs was not relaxed compared to the film on the control substrate.

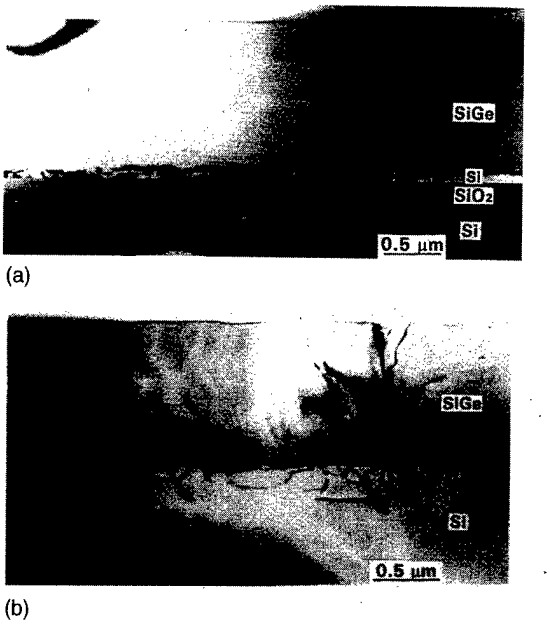


FIG. 2. SiGe on (a) SOI and (b) Si substrates (after Ref. 7).

In 1994, Powell *et al.* published the results of an experimental study of the growth of SiGe on silicon on insulator (SOI) substrates. The Si surface was etched back to a thickness of 50 nm. They interpreted the data with respect to a compliant substrate model.[5,6] They examined the relaxation behavior of low-misfit, $Si_{0.85}Ge_{0.15}$–Si, structures as a function of annealing. The annealing temperature was varied from 700 to 1050 °C for a duration of 1 h. X-ray diffraction measurements showed that the samples were partially relaxed with the anneal. A transmission electron microscope (TEM) analysis was used to observe the dislocation structure produced with the relaxation. The misfit dislocations present at the Si–SiGe interface had associated threading segments that penetrated *down into the Si* substrate, rather than into the SiGe overlayer. The authors speculate that this resulted from the introduction of tensile strain in the thin Si, which can only occur from slippage at the Si/SiO_2 interface. The growth and anneal temperatures were too low to allow flow of the oxide.

Recent work by Wang and co-workers shows a significant improvement in SiGe on SOI can be achieved *during the growth process*. Figure 2 shows a comparison of $Si_{0.6}Ge_{0.4}$ (1.0 μm) grown on SOI (with 20 nm of Si) and Si. The threading dislocation density is significantly reduced in the SiGe film grown on the SOI substrate.[7]

SOI has been used exclusively, thus far, as the compliant substrate implementation for GaN growth. Yang *et al.* proposed and demonstrated the formation of SiC compliant substrates by carbonizing the Si surface of a SOI film. An advantage of this approach is that a large area growth of a SiC template for subsequent GaN deposition is formed. Yang showed that at a temperature of 900 °C, an incident flux of

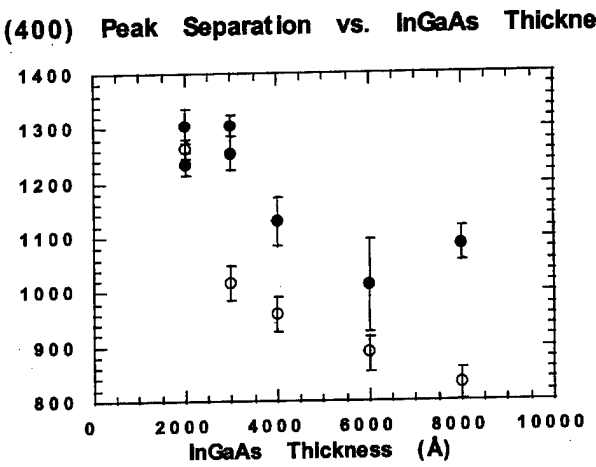


FIG. 3. X-ray diffraction peak separation vs thickness for InGaAs on compliant substrates (closed circles) and reference substrates (open circles) (after Ref. 12).

carbon can partially or completely convert 50 nm of Si to SiC.[8] Cao and co-workers have reported reduced strain in GaN on SOI compared to growth on sapphire. The authors infer that this results from compliancy of the SOI substrate.[9]

A GaAs-based bonded substrate removal technology was first proposed and demonstrated by Carter-Coman *et al.*[10,11] This process is shown in Fig. 1. Several different bonding technologies utilize similar process flows. The first step in the process is to grow an appropriate wafer for fabrication of the thin substrate portion of the compliant substrate. This wafer contains an etch stop layer (e.g., AlGaAs or GaInP for GaAs) that enables removal of the epitaxial film from the substrate. After growth, the wafer is bonded (surface side) to the mechanical handle wafer. The substrate is removed by lapping and etching.

Major differences in compliant substrate implementations result from the utilization of different types of bonds. For Carter-Coman's process, an In/Ga, bond is utilized. The exact nature of this bond is unknown. A reduced relaxation rate, as inferred by the separation between the substrate and film angle in x-ray measurements, has been observed for low-misfit InGaAs on GaAs compliant substrates.[12] This is shown in Fig. 3. Kuech and co-workers are developing an approach to engineer the viscosity of the bond material as a function of temperature by using doped glasses.[13] In addition, Zhang *et al.* have developed a process utilizing oxidized AlAs or AlGaAs (believed to be porous) as the bond layer.[14] Finally, the recent exciting results of Lo and coworkers with the compliant universal (CU) substrate utilize a twist wafer bond.[15–17] The process flow for this is shown in Fig. 4. For the twist bond, the substrate and handle wafer are purposely misoriented by an angle of 10°–45° (Ref. 17) to produce an array of screw dislocations at the bonded interface. The spacing between these dislocations decreases as the twist angle increases. A number of demonstrations of high-misfit materials on GaAs twist-bonded substrates have been made. As an example, Fig. 5 shows a cross-sectional TEM micrograph of 650 nm of InSb (14.7% mismatch to GaAs)

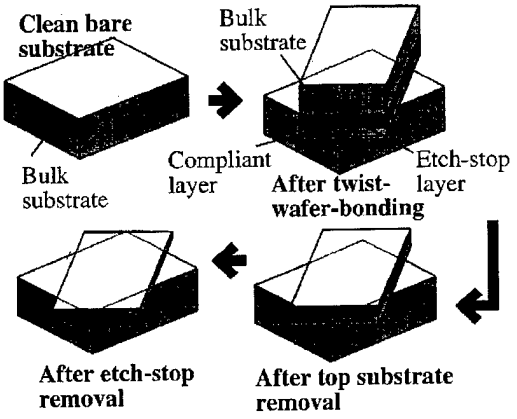


FIG. 4. Process flow for twist-bonded compliant substrates (after Ref. 16).

on a CU substrate. In this case, the thickness of the CU substrate is 3 nm. Clear evidence of a dramatically reduced threading dislocation density ($<10^6$ cm^{-2}) is seen for the growth on the compliant substrates compared to that on reference substrates.

Finally, a new approach to the production of improved GaN is under development by Doolittle *et al.*[18] This approach involves the production of thin-film GaN substrates. GaN is first grown on lithium gallate (LGO). Lithium gallate possesses a relatively close lattice match, but a poor thermal match to GaN. The GaN is then bonded to an appropriate mechanical handle wafer, and the LGO is removed by wet etching. This substrate is not compliant in the standard sense. The GaN substrate is used for GaN-based material overgrowth. However, the bond between the GaN thin film and the handle wafer must have a low viscosity just as in a compliant substrate to reduce defect production due to the coefficient of thermal expansion mismatches between the GaN and the handle wafer.

IV. MODEL DEVELOPMENT

Models describing a critical thickness for the growth of materials on compliant substrates have been developed. These have been based on the Matthews and Blakeslee force balance model,[2,19] or on a metastability model similar to that of Dodson and Tsao.[19–21] All of these models predict that, for some complaint substrate and mismatch strain, an epitaxial layer can be grown that does not have a critical thickness (see Fig. 6). This means that a (theoretically) infinite epitax-

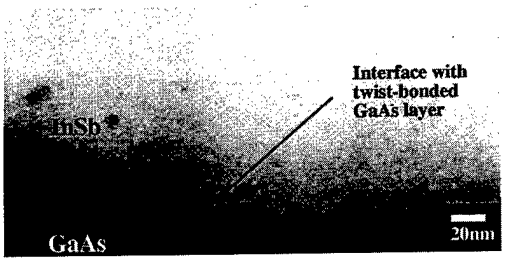


FIG. 5. TEM micrograph of InSb on a GaAs twist-bonded compliant substrate (after Ref. 16).

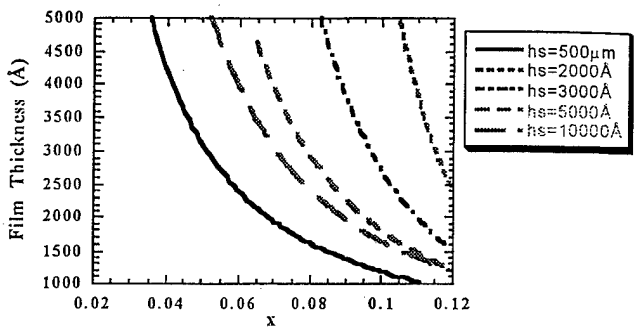


FIG. 6. Calculation of critical thickness for InGaAs (x=In concentration) on a GaAs compliant substrate as a function of substrate thickness (after Ref. 21).

ial film can be grown free of dislocations resulting from mismatch strain. Figure 7 shows the results of a metastable model of the growth of low-misfit InGaAs on GaAs. The model was developed using relaxation rate data for the same composition of InGaAs grown on reference substrates.[21]

Tan and Gösele have offered explanations of the observed behavior of reduced threading dislocation density enabled by growth on the CU substrate.[22] They propose that the overlayers on the CU substrates are partially or fully relaxed by misfit dislocations at the substrate–overlayer film interface. This situation results from a misfit dislocation network that is arranged in an orderly fashion, with very few threading segments in the overlayer. They refer to this as a *correlated* misfit dislocation process, which consists of two different possible mechanisms. First, misfit dislocations can be nucleated from the screw dislocation nodes at the substrate–handle wafer interface. This, in turn, plastically deforms the substrate. Second, dislocation half loops nucleate at the film surface at weakly bonded or high energy positions. For the growth on a twist-bonded substrate, these positions are not random, such as growth on a normal substrate, but are positioned above the screw dislocations. This increases the chance of interaction and annihilation of threading dislocation, which in turn can dramatically reduce the density of these segments.

V. ISSUES AND CONTINUED DEVELOPMENT

Current research and development in the area of compliant substrates have necessarily focused on the process development issues in fabricating a near ideal compliant substrate. The bonds, whether twist wafer bonds, metal layers, or engineered glasses, need to be well understood, particularly as a function of the temperature and stresses experienced in the growth process. Additional factors are introduced by the condition of the prepared surface of the thin substrate, in which the surface morphology of the substrate is related to the particular etch stop used in the fabrication process and the etching process.[23] The surface morphology, threading defect density, and the ability to protect and then *in situ* clean the surface prior to growth are critical factors that affect subsequent relaxation during growth. It may be difficult to truly have a control bulk sample that mimics all these features of the compliant substrate. Recent analysis indicates that the development of the surface morphology plays a role in the growth of mismatched films on compliant substrates. Seaford *et al.*[17] showed evidence via atomic force microscopy (AFM) data that the surface morphology is quite different on reference and CU substrates for InSb growth. The development of a nonplanar surface morphology reduces the strain in the mismatched overlayer. For the growth on the CU substrate, undulations with a depth of approximately 100 nm are observed. Similar surface ripples and undulations have been observed by Shen and Brown.[24]

VI. SELECTIVE AREA GROWTH VIA COMPLIANT SUBSTRATES

Patterning at the bond interface can be used to engineer a lateral strain profile at the surface of a growing, mismatched overlayer. By exploiting the relationship between strain-dependent growth kinetics and this strain profile, lateral material property variations can be achieved *without any surface patterning*. This approach is called strain-modulated epitaxy (SME).[11,25] Strain-dependent growth kinetics that can be exploited include the production of self-assembled quantum dots and inter- and intralayer growth rates and anisotropies, among others. Figure 7 shows AFM images of

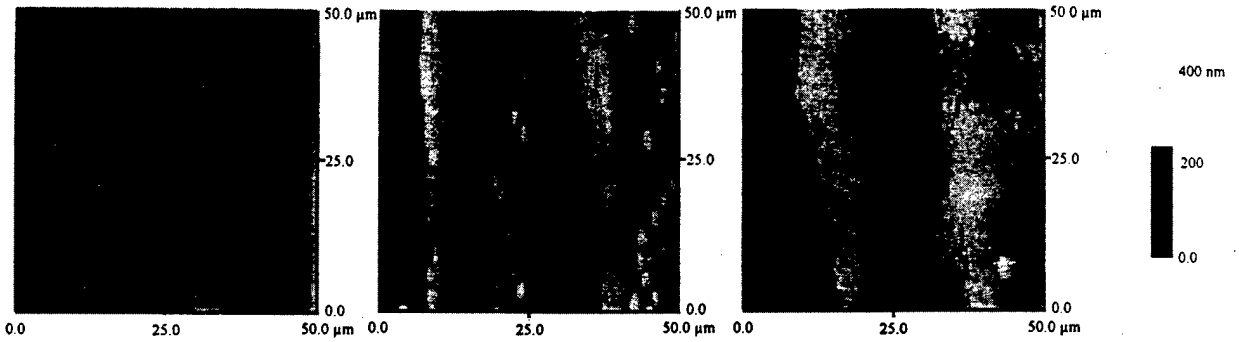


FIG. 7. AFM images of InGaAs on GaAs bottom-patterned compliant substrates as a function of growth temperature for (a) 520, (b) 560, and (c) 580 °C (after Ref. 25).

surface features or mounds that evolve as a function of growth temperature for 250 nm of $In_{0.07}Ga_{0.93}As$ growth on bottom-patterned GaAs compliant substrates. A clear spatial preference for mound formation is observed to correlate with 10 μm stripes formed at the bond interface.

VII. CONCLUSIONS

Compliant substrates offer a promising new approach to strain management in mismatched epitaxial films. The current emphasis is on implementation of various proposed compliant substrate approaches. Clear evidence of a reduced dislocation density in films grown on compliant substrates exists. Some of the data, however, are inconsistent with a simple model of compliant substrate behavior.

ACKNOWLEDGMENTS

The author would like to acknowledge Carrie Carter-Coman (now at Hewlett–Packard Optoelectronics Division), Jeng-Jung Shen, Alan Doolittle, Nan Marie Jokerst, Robert Bicknell-Tassius (now at the Jet Propulsion Laboratory), Paul Kohl, and Robert Metzger, all at Georgia Tech. She would also like to acknowledge Thomas F. Kuech at The University of Wisconsin–Madision, Matt Seaford at Wright Laboratories Materials Laboratory, Y.-H. Lo of Cornell University, and Wen Wang of Columbia University for providing information and insight.

[1]A compliant substrate may also be compliant due to its softness with respect to the epitaxial film.
[2]Y. H. Lo, Appl. Phys. Lett. **59**, 2311 (1991).
[3]Some of these issues were specified by T. F. Kuech in a private communication.
[4]C. L. Chua, W. Y. Hsu, C. H. Lin, G. Christenson, and Y. H. Lo, Appl. Phys. Lett. **64**, 3640 (1994).
[5]A. R. Powell, S. S. Iyer, and F. K. LeGoues, Appl. Phys. Lett. **64**, 1857 (1994).
[6]They did not, however, refer to the substrate as "compliant."
[7]W. I. Wang, The 16th North American MBE Conference (to be published).
[8]Z. Yang, F. Guarin, I. W. Tao, W. I. Wang, and S. S. Iyer, J. Vac. Sci. Technol. B **13**, 789 (1995).
[9]J. Cao, D. Pavlides, A. Eisenbach, A. Phillippe, C. Bru-Chevallier, and G. Guillot, Appl. Phys. Lett. **71**, 3880 (1997).
[10]C. Carter-Coman, A. S. Brown, N. M. Jokerst, D. E. Dawson, R. Bicknell-Tassius, Z. C. Feng, K. C. Rajkumar, and G. Dagnall, J. Electron. Mater. **25**, 1044 (1996).
[11]C. Carter-Coman, A. S. Brown, R. Bicknell-Tassius, N. M. Jokerst, F. Fournier, and D. E. Dawson, J. Vac. Sci. Technol. B **14**, 2170 (1996).
[12]C. Carter-Coman, R. Bicknell-Tassius, A. S. Brown, and N. M. Jokerst, Appl. Phys. Lett. **70**, 1754 (1997).
[13]T. F. Kuech (private communication).
[14]C. Zhang, D. I. Lubyshev, W. Cai, J. E. Neal, D. L. Miller, and T. S. Mayer, The 1997 International Symposium on Compound Semiconductors.
[15]F. E. Ejeckham, Y. H. Lo, S. Subramanian, H. Q. Hou, and B. E. Hammons, Appl. Phys. Lett. **70**, 1685 (1997).
[16]F. E. Ejeckham, M. L. Seaford, Y. H. Lo, H. Q. Hou, and B. E. Hammons, Appl. Phys. Lett. **71**, 776 (1997).
[17]M. L. Seaford, D. H. Tomich, K. G. Eyink, W. V. Lampert, and F. E. Ejeckham, in Ref. 14.
[18]W. A. Doolittle, T. Kropenwicki, C. Carter-Coman, S. Stock, P. Kohl, N. M. Jokerst, R. A. Metzger, S. Kang, K. Lee, G. May, and A. S. Brown, in Ref. 7.
[19]L. B. Freund and W. D. Nix, Appl. Phys. Lett. **69**, 173 (1996).
[20]D. Teng and Y. H. Lo, Appl. Phys. Lett. **62**, 43 (1993).
[21]C. Carter-Coman, R. Bicknell-Tassius, A. S. Brown, and N. M. Jokerst, Appl. Phys. Lett. **71**, 1344 (1997).
[22]T. Y. Tan and U. Gösele, Appl. Phys. A: Mater. Sci. Process. **64**, 631 (1997).
[23]C. Carter-Coman, R. Bicknell-Tassius, R. G. Benz, A. S. Brown, and N. M. Jokerst, J. Electrochem. Soc. **144**, L29 (1997).
[24]J. J. Shen and A. S. Brown (unpublished data).
[25]C. Carter-Coman, A. S. Brown, A. Metzger, N. M. Jokerst, J. Pickering, and L. Bottomley, Appl. Phys. Lett. **71** (1997).

Interface states induced in GaAs by growth interruption during an *in situ* process

F. Wakaya,[a] T. Matsubara, M. Nakayama, J. Yanagisawa, and Y. Yuba
Department of Physical Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560, Japan

S. Takaoka and K. Murase
Department of Physics, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560, Japan

K. Gamo
Department of Physical Science, Graduate School of Engineering Science and Research Center for Materials Science at Extreme Conditions, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560, Japan

(Received 21 January 1998; accepted 21 May 1998)

Density and level of interface states are accompanied by growth interruption, which is inevitable in an *in situ* process using molecular beam epitaxy (MBE). This MBE process coupled with focused ion beam systems is investigated by means of the Hall measurements and capacitance–voltage measurements. The experimental results are compared to those of the self-consistent calculation. These results suggest that interface states are located at a shallow level from the conduction band edge. © *1998 American Vacuum Society.* [S0734-211X(98)09304-4]

I. INTRODUCTION

Recently, transport properties not only of a single quantum dot and quantum wire but also of quantum dots and wires strongly coupled through tunnel barriers are attracting the attention of researchers. The coupled quantum dots[1,2] are used for investigation of photon-assisted tunneling and artificial molecules. The coupled quantum wires[3,4] are interesting for the switching characteristics of coherent electron waves.

In spite of such interesting features, fabrication technology for the coupled zero-dimensional (0D) and one-dimensional (1D) quantum structures has not yet become well established. Using molecular beam epitaxy (MBE), we can realize coupled two-dimensional (2D) structures, i.e., coupled quantum wells with abrupt potential barriers in the growth direction. Therefore, a combination of a lateral-directional fabrication technology with the MBE seems to be promising for the fabrication of the coupled 0D and 1D quantum structures. One of the promising lateral-directional fabrication technologies is laterally selective doping using *in situ* low-energy focused ion beam (FIB) implantation. Under such circumstances, we have developed an *in situ* fabrication system with MBE and FIB.[5]

One of the problems in the FIB-MBE *in situ* fabrication system is the carrier depletion caused by the growth interruption, which is inevitable in the system. Although we do not need the growth interruption if we use a FIB column attached to the MBE chamber,[6,7] which is an *in situ* process in a narrow sense, it is difficult to obtain a fine-focused beam in this case. Therefore the separate chambers for MBE and FIB with the ultrahigh vacuum transfer tunnel seem to be promising at present for the fabrication of coupled 0D and 1D quantum structures.

The purpose of this work is to investigate the levels and densities of the interface state at the growth-interrupted interface, which could provide information on the origins of the interface state.

II. EXPERIMENTS

In the FIB-MBE *in situ* fabrication system, there are two major origins of the carrier depletion; the interface states caused by the growth interruption and the damage caused by the FIB implantation. In order to focus on the former, we utilized the δ-doped Si using the MBE to simulate the low-energy FIB implanted Si in the present article.

In order to investigate the effects of the growth interruption experimentally, we prepared the wafers shown in Fig. 1 and performed the Hall measurements. The distance between the growth-interrupted and the δ-doped layers was 2 nm, which was on the order of the projected range for the low-energy FIB. Although there were the n^+ cap layers (Si:

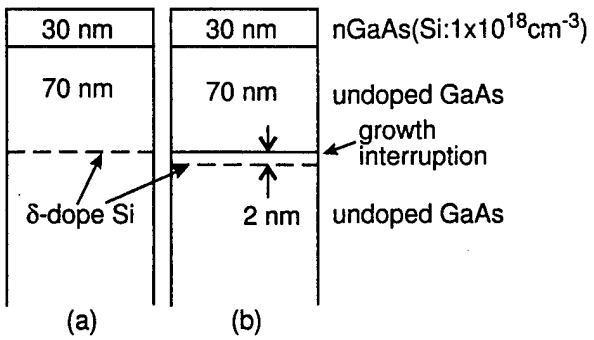


FIG. 1. Sample structures for Hall measurements.

[a]Electronic mail: wakaya@ee.es.osaka-u.ac.jp

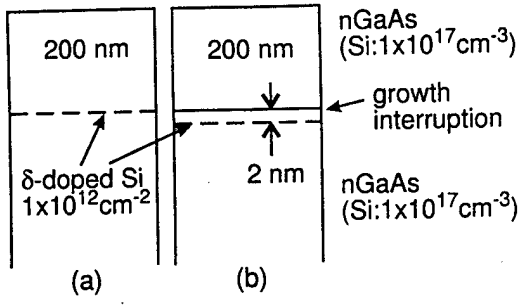


FIG. 2. Sample structures for $C-V$ measurements.

1×10^{18} cm^{-3}) to avoid severe surface depletion, the wafers with only the cap layer, i.e., without the δ-doped layer were not conductive, suggesting that there were no free electrons in the cap layers. The growth temperature was 580 °C. The growth interruptions were performed at room temperature for 3 h in vacuum at 7×10^{-9} Torr.

Capacitance–voltage ($C-V$) measurements provide information not only on the free-carrier profile but also on the levels and the densities of the interface states.[8–10] For the $C-V$ measurement, we prepared the wafer shown in Fig. 2. The major difference from the structure shown in Fig. 1 is the doping. The layers around the δ-doped layer were slightly doped (1×10^{17} cm^{-3}) in case of the wafers for the $C-V$ measurements. The growth temperature was 580 °C in this case also. Growth was interrupted for 3 h in a vacuum of 1×10^{-9} Torr.

III. RESULTS AND DISCUSSION

A. Dopant Si concentration dependence of electron density

The results of the Hall measurements for the wafer structures shown in Figs. 1(a) and 1(b) with several Si concentrations are shown in Fig. 3. The difference in the electron densities between the continuously grown and growth-interrupted wafers reflects the interface-state density at the growth-interrupted interface. As shown in Fig. 3 this difference depends on the Si concentration.

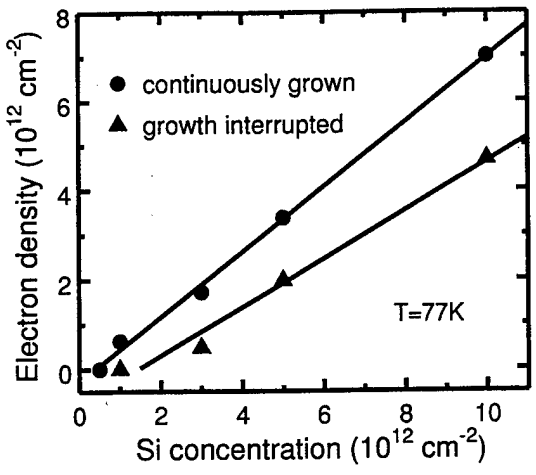


FIG. 3. Experimentally obtained electron density as a function of the dopant Si concentration. Structures of the continuously grown and growth-interrupted wafers are schematically shown in Figs. 1(a) and 1(b), respectively.

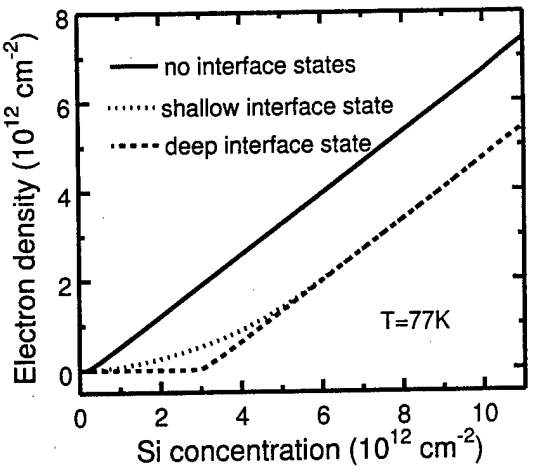


FIG. 4. Calculated electron density as a function of the dopant Si concentration. Solid, dotted, and dashed lines correspond to the cases of no interface states, shallow (5 meV) interface state, and deep (700 meV) interface state, respectively.

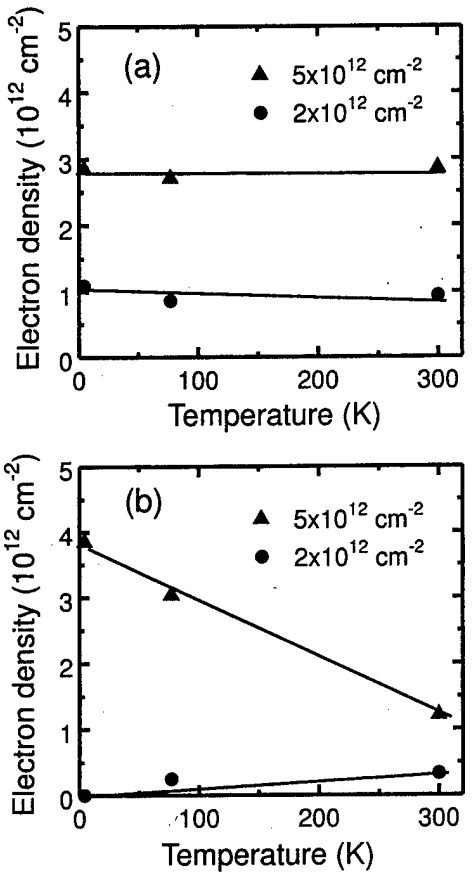


FIG. 5. Experimentally obtained electron density of (a) the continuously grown and (b) the growth-interrupted wafers. Triangles and circles denote the dopant Si concentrations of 5×10^{12} and 2×10^{12} cm^{-2}, respectively.

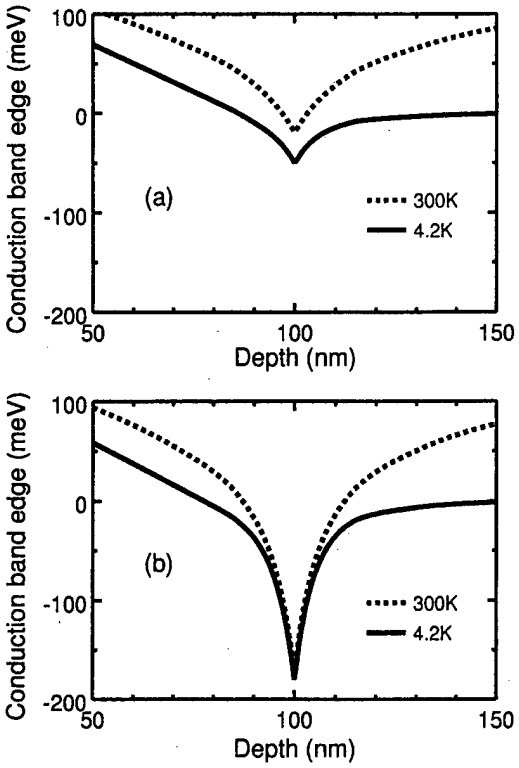


FIG. 6. Band diagram calculated by self-consistent method. Dopant Si concentrations are: (a) 1×10^{12} and (b) 5×10^{12} cm^{-2}.

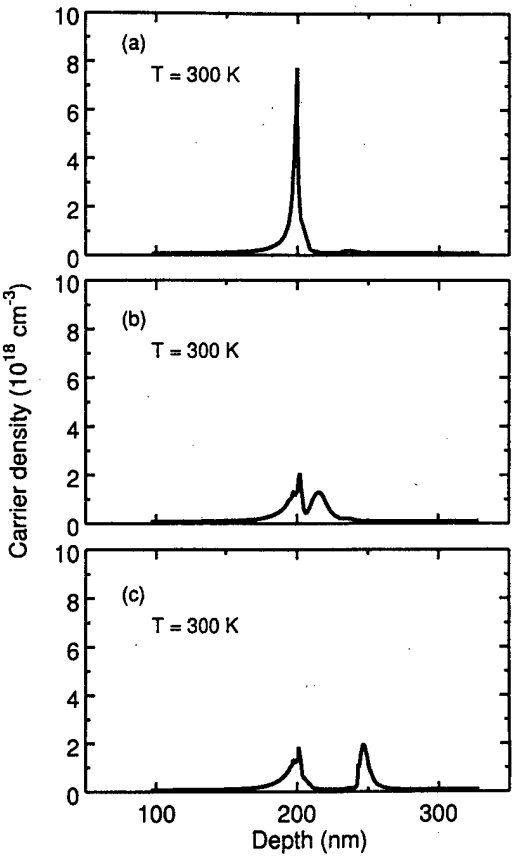


FIG. 7. Carrier profiles obtained from the calculated $C-V$ characteristics. Assumed wafer structure is shown in Fig. 2(b). Interface-state level is assumed to be (a) 100 meV, (b) 400 meV, and (c) 700 meV. Deep interface state makes the apparent carrier located at the deeper position than that of the δ-doped layer.

In order to discuss the origin of this dependence, we calculated the free-carrier density in the same structure as used in the experiment by a self-consistent method using the Schrödinger and the Poisson equations by assuming several interface-state levels. The results obtained are shown in Fig. 4. The dashed and dotted lines in the figure are the calculated electron densities for the structures with the interface state at 700 and 5 meV below the conduction-band minimum, respectively. Although the interface state should be a continuous function for both the depth and the energy,[8] we assumed the δ function state density for both of them in the present calculation. The interface-state density and the activation ratio of the doped Si were assumed to be 2×10^{12} cm^{-2} and 70%, respectively.

The calculated results clearly show that the difference of the electron densities between the continuously grown and the growth-interrupted wafers should be constant if the interface-state level is deep. In case of the shallow interface state, on the other hand, the difference can depend on the doped Si concentration. This is because the number of electrons captured at the interface state depend on the Fermi level, i.e., doping density, especially in case of shallow states.

Comparing the experimentally observed results and the calculated results, we speculated about that the observed dependence of the amount of the depletion on the doped Si concentration was due to the shallow interface-state level.

B. Temperature dependence of electron density

The temperature dependence of the electron density obtained by the Hall measurement is shown in Fig. 5. Here Figs. 5(a) and 5(b) correspond to the continuously grown and growth-interrupted wafers, respectively. The triangles and circles denote results observed for the doped Si concentration of 5×10^{12} and 2×10^{12} cm^{-2}, respectively.

In this growth run, the continuously grown wafer has the same structure as that shown in Fig. 1(a), while the growth-interrupted wafer has 68 nm undoped GaAs instead of the 70 nm undoped GaAs shown in Fig. 1(b). The growth was interrupted for 3 h in the vacuum of 1×10^{-9} Torr.

As shown in Fig. 5, the electron density was almost constant in the case of continuously grown wafers, while it depended on the temperature for the growth-interrupted wafers. Moreover, the electron density decreased for increasing temperature in the case of the wafer with the dopant Si of 5×10^{12} cm^{-2} with growth interruption. Although the origin of these temperature dependences is not clear at present, these phenomena should reflect the interface-state level and density. If the interface state is located at a deep level, it should capture the electrons in any case, with the result that the electron density should not depend on the temperature. Therefore, the observed results shown in Fig. 5(b) seem to suggest that the interface-state level is shallow.

Figure 6 shows the calculated band diagram of the structures shown in Fig. 1(a). The dopant Si was assumed to be 1×10^{12} and 5×10^{12} cm^{-2} for Figs. 6(a) and 6(b), respec-

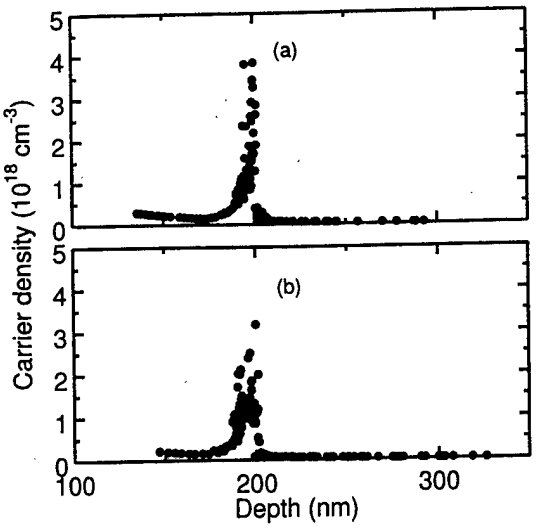


FIG. 8. Carrier profiles obtained from the experimental results of the $C-V$ measurements. Wafer structures for (a) the continuously grown and (b) the growth-interrupted wafers are shown in Figs. 2(a) and 2(b), respectively.

tively. The origin of the vertical axis is the Fermi level. The figure clearly shows that the conduction band minimum around the δ-doped layer hardly depends on the temperature in the case of heavy doping while it depends on the temperature more in the case of light doping. This dependence on the temperature is a general property of semiconductors as often seen in bulk semiconductors with heavy doping (degenerated) and light doping. Although these band diagrams were calculated for the structure without interface states, the temperature dependence of the conduction-band minimum in the structure with the interface states may be similar to that shown in this figure and could provide an explanation for the observed results shown in Fig. 5(b) as follows. In the case of low density of dopant Si, the Fermi level decreases (i.e., the conduction band minimum increases) with increasing temperature, with the result that the number of electrons captured at the interface state decreases. This leads to an increase of the free electron number with increasing temperature as observed in Fig. 5(b). In the heavy doping case, on the other hand, the Fermi level is almost constant. The number of the electrons captured at the interface state is therefore determined by the Fermi distribution function. In this case, the free carrier density may decrease with increasing temperature.

C. $C-V$ method

In order to see the effects of interface states on the carrier profile, we performed a self-consistent calculation using the Poisson equation and the Fermi distribution function and calculated $C-V$ characteristics. The carrier density obtained from the calculated $C-V$ characteristics is shown in Fig. 7. The wafer structure used in the calculation is shown in Fig. 2(b). At the growth-interrupted interface, the interface state density of 2×10^{12} cm^{-2} was assumed at (a) 100 meV, (b) 400 meV, and (c) 700 meV below the conduction band minimum. In the case of Fig. 8(a) there is only one peak at a depth of the δ-doped layer, while there is an additional peak at a deeper position in the cases of (b) and (c). These additional peaks can be explained as follows. The first peak, which appears at the δ-doped layer (200 nm), corresponds to the bias voltage at which the depletion layer reaches the δ-doped layer. When the negative bias voltage becomes larger, the depletion layer extends to the deeper position. At a certain bias voltage, the electrons captured at the growth-interrupted interface state detrap to the conduction band because the Fermi level becomes lower than the interface-state level. During this event, the depletion layer does not extend with increasing negative gate voltage. As a result, the second carrier peaks in Figs. 7(b) and 7(c) appears at a deeper position than the δ-doped layer and growth-interrupted interface. The carriers corresponding to the second carrier peaks in Fig. 7 cannot be detected by the Hall measurement, because there are no free carriers corresponding to the second carrier peaks in the zero-biased wafers. These calculated results mean that we can expect two carrier peaks if the interface state is deep, while one peak should be observed if the interface state is as shallow as 100 meV.

Figures 8(a) and 8(b) show the carrier profiles of the wafers shown in Figs. 2(a) and 2(b), respectively, which were obtained from the $C-V$ measurement. Although there were some differences in the peak height and width between the carrier profiles of the continuously grown and growth-interrupted wafers, we observed only one peak in both cases. This suggest that the interface state is located at a shallow level from the conduction-band edge, which is consistent with the results of the Hall measurements.

IV. SUMMARY

We have investigated the density and the level of the interface state caused by the growth interruption which is inevitable in the FIB-MBE *in situ* process. The experimental results of the Hall measurements and the $C-V$ measurements were compared to the results of the self-consistent calculation. Both the obtained results of the Hall measurement and the $C-V$ measurement suggest that the interface state is located at a shallow level from the conduction-band edge.

[1]T. Fujisawa and S. Tarucha, Superlattices Microstruct. **21**, 247 (1997).
[2]T. Fujisawa and S. Tarucha, Jpn. J. Appl. Phys., Part 1 **36**, 4000 (1997).
[3]J. A. del Alamo and C. C. Eugster, Appl. Phys. Lett. **56**, 78 (1990).
[4]N. Tsukada, A. D. Wieck, and K. Ploog, Appl. Phys. Lett. **56**, 2527 (1990).
[5]K. Gamo and S. Namba, Proceedings of the 2nd International MicroProcess Conference, Kobe, 1989, p. 293.
[6]J. H. Thompson, D. A. Ritchie, G. A. C. Jones, E. H. Linfield, J. E. F. Frost, A. C. Churchill, G. W. Smith, D. Lee, M. Houlton, and C. R. Whitehouse, Surf. Sci. **267**, 69 (1992).
[7]P. J. A. Sazio, J. H. Thompson, G. A. C. Jones, E. H. Linfield, D. A. Ritchie, M. Houlton, and G. W. Smith, J. Vac. Sci. Technol. B **14**, 3933 (1996).
[8]L. He, H. Hasegawa, T. Sawada, and H. Ohno, J. Appl. Phys. **63**, 2120 (1988).
[9]L. C. Kimerling, J. Appl. Phys. **45**, 1839 (1974).
[10]S. Loualiche, A. Nouaihat, and G. Guillot, Solid-State Electron. **25**, 577 (1982).

Influence of sulfur interlayers on the Mg/GaAs(100) interface formation

St. Hohenecker,[a] T. U. Kampen, and D. R. T. Zahn
Institut für Physik, TU Chemnitz, D-09107 Chemnitz, Germany

W. Braun
BESSY GmbH, Lentzeallee 100, D-14195 Berlin, Germany

(Received 21 January 1998; accepted 20 May 1998)

The modification of clean GaAs(100) surfaces by *in situ* deposition of molecular sulfur was investigated by soft x-ray photoemission spectroscopy. Upon S treatment of the clean GaAs(100) sample at 435–455 °C in ultrahigh vacuum the formation of a three monolayer thick gallium sulfide-like compound is observed, which exhibits a (2×1) low-energy electron diffraction pattern. Due to the S modification on *n*-GaAs a reduction of the band bending by 0.35 eV is achieved, while the band bending on *p*-GaAs is increased by 0.17 eV. The subsequent Mg evaporation leads to the formation of a metal/semiconductor contact with a reacted magnesium sulfide-like compound at the interface. After 1 nm Mg deposition the Schottky barrier height of the S-modified Mg/*n*-GaAs(100) contact amounts to 0.44 eV, which is 0.18 eV lower than without S modification, while the Mg/*p*-GaAs(100) Schottky contact exhibits an increase in the Schottky barrier height by 0.30 eV in comparison to the value of the unmodified Schottky contact (0.55 eV). © *1998 American Vacuum Society.* [S0734-211X(98)08504-7]

I. INTRODUCTION

Chalcogen atoms play an important role in the passivation of semiconductor surfaces. As a result there are many investigations dealing with the effects of sulfur modification of GaAs(100) surfaces. Different S containing materials were applied to achieve a S passivation of the GaAs(100) surface, such as $(NH_4)_2S_x$,[1,2] S_2Cl_2,[3,4] and Na_2S[5] in aqueous solutions but also in alcoholic solutions.[6] There are also many experimental results obtained by the *in situ* S preparation of GaAs(100) samples using H_2S[24] or an electrochemical cell.[7] Different methods were applied to investigate the properties of the S passivated GaAs(100) surfaces like scanning tunneling microscopy (STM),[7] low-energy electron diffraction (LEED),[7] reflection high-energy electron diffraction (RHEED),[2] photoluminescence (PL) measurements,[6] x-ray photoelectron spectroscopy (XPS),[1–4,6,7] and Auger electron spectroscopy (AES).[3,4,7] Today there is hardly any comparative study dealing with the sulfur modification of *n*- and *p*-doped GaAs(100) surfaces, which is important for the understanding the passivating properties of sulfur on GaAs(100). It is further interesting to examine the influence of the sulfur modification on the Schottky barrier formation of metal/GaAs(100) contacts, which also was done only in a few cases for Au,[8] Fe,[9] Pd,[10,11] and Al[11] on S passivated GaAs(100). The latter is, however, of great technological relevance since it may provide new opportunities of interface engineering, in particular if the sensitivity of the barrier height to properties of the metal can be influenced to a larger extent than by metal deposited on the unmodified GaAs(100) surface. Here first experiments were carried out for the Mg/GaAs(100) system for both types of doping.

[a]Author to whom correspondence should be addressed; electronic mail: s.hohenecker@physik.tu-chemnitz.de

II. EXPERIMENT

The sulfur modification of GaAs(100) surfaces and subsequent metal deposition was studied using high resolution soft x-ray photoemission spectroscopy (SXPS) at the TGM2 beamline of the synchrotron light source BESSY at Berlin. The beamline was instrumented with a VG ADES 400 electron spectrometer giving a combined resolution of light and photoelectrons of approximately 0.3 at 60 eV photon energy. This could be determined from the measurement of the Fermi edge on a clean molybdenum surface.

Homoepitaxial *n*- and *p*-type GaAs(100) layers ($n,p = 10^{18}$ cm^{-3}) grown by molecular beam epitaxy (MBE) and subsequently capped by an amorphous As layer served as substrates in this study. These samples were decapped by gentle annealing to 380 °C in ultrahigh vacuum (base pressure $<2\times10^{-10}$ mbar) leading to an As rich surface as can be judged from the SXPS data. S was evaporated onto these surfaces by thermal decomposition of tindisulfide from a Knudsen celllike oven. This yields a saturation coverage as could be determined from the SXPS measurements. The sulfur treatment was carried out at elevated temperatures (435–455 °C) leading to a (2×1) reconstructed surface.

Thereafter Mg was evaporated onto these modified GaAs(100) samples with increasing thickness up to approximately 1 nm nominal thickness. A Mg wire (3 N) wrapped around a tungsten filament, which was resistively heated, served as a Mg source. The thickness of the Mg film was monitored by a quartz crystal microbalance in close vicinity to the sample surface. After each step of modification photoelectron spectra were recorded of the different participating elements. The photon energies for the photoemission measurements of the different core levels [Ga 3*d* (60 eV), As 3*d* (79 eV), S 2*p* (195 eV), and Mg (89 eV)] were always chosen to achieve maximum surface sensitivity. To determine the binding energy of the Ga 3*d* and As 3*d* core level exactly

TABLE I. Fit parameters for the deconvolution of the core level spectra, which are held fixed during the fitting procedure.

	Ga 3*d*	As 3*d*	S 2*p*	Mg 2*p*
Lorentzian width/meV	100	100	100	100
Branching ratio	1.68	1.50	2.00	2.00
Spin-orbit Splitting/meV	440	690	1180	260
Asymmetry factor	0.166	...	...	0.280

after each spectrum the Fermi level position was measured on a clean piece of molybdenum. Thereafter the measured spectra were fitted using spin-orbit split Voigt profiles to separate the contributions of chemical shifted components from the bulk.

III. RESULTS AND DISCUSSION

A. Chemistry

In the decomposition of the core level spectra the spin-orbit splitting, the branching ratio and the Lorentzian width were always kept constant throughout the fits of the differently modified GaAs(100) surfaces. These values (Table I) were first derived from the clean surface and thereafter optimized to provide the best fit for the complete series. The other parameters such as binding energy, intensity, background (Shirley background), and Gaussian width were variable during the fits. The asymmetry parameters of the Doniach–Sunjic line shape of metallic components were achieved from fitting the core level spectra in which the metallic component is most prominent. In the following the discussion of the chemistry will be related to *p*-type GaAs(100) samples since *n*-type samples reveal very similar behavior. After the decapping the clean GaAs(100) sample exhibits an As-rich $c(4\times4)$ or (1×1) reconstructed surface as can be stated from the shape of the As 3*d* core level spectrum. The deconvolution of this spectrum as shown in Fig. 1(a) shows three different components. The As 1 component is assigned to bulk As, while As 2 and As 3 are surface components. The binding energy of the bulk component amounts to 40.98 eV (all binding energies are given for the $d_{5/2}$ components of the spin-orbit split core levels relative to the Fermi level), while the surface components are shifted by 0.67 eV towards higher (As 2) and 0.48 eV towards lower binding energy (As 3). These shifts are in good agreement with the data given in Ref. 12. According to the literature[13–17] the As 2 component is assigned to As in an pure As surrounding as it occurs for the As dimers on the $c(4\times4)$ reconstructed GaAs(100) surface. The As 3 component is attributed to threefold coordinated As which is caused by the missing dimer rows of this surface.[15,18]

In Fig. 2(a) the Ga 3*d* core level spectrum exhibits only a bulk (Ga 1) and one surface component (Ga 2). The binding energy of the bulk component amounts to 19.12 eV while the surface component is shifted by 0.6 eV towards higher binding energy. An explanation for the Ga surface component

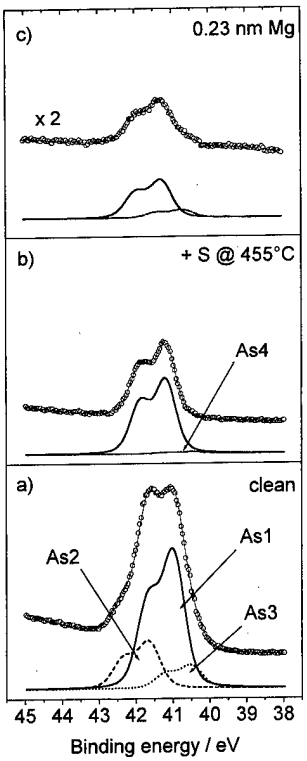


FIG. 1. As 3*d* core level spectra measured at 79 eV photon energy: (a) As 3*d* spectrum of the clean, As-rich GaAs(100) surface after the decapping at 380 °C; (b) after S modification at 455 °C sample temperature; (c) after deposition of 0.23 nm Mg.

may be that it originates from Ga at the interface of the bulk to the As covered surface where it is bound to the threefold coordinated As atoms in the second layer.[15]

The S treatment of the GaAs(100) surface changes the As 3*d* and Ga 3*d* core level spectra dramatically. In Fig. 1(b) one can see that both As surface components disappeared and only the bulk component (As 1) and a small interface component (As 4) remain in contrast to the room temperature S treatment of GaAs(100) samples, where the formation of arsenic sulfides can be observed.[19] The excess As is removed from the surface by the formation of arsenic sulfide which evaporates from the GaAs surface at temperatures above 180 °C.[19] The shift of the bulk component will be discussed later. The As 4 component is shifted by 0.75 eV relative to the bulk component towards higher binding energy, which is by 0.25 eV higher than observed by others.[7] A shift to higher binding energy is in agreement with a similar model for (2×1) reconstructed Se modified GaAs(100) surfaces of Pashley (Fig. 3, Ref. 20) where threefold coordinated As atoms occur at the interface of bulk GaAs to the gallium selenide-like layer caused by Ga vacancies. There is also a reduction of the Gaussian width by approximately 60 meV in the As 3*d* components which can be explained by a more homogeneous distribution of the As atoms near this S modified surface.

The Ga 3*d* core level spectrum in Fig. 2(b) exhibits two new intense surface components (Ga 4, Ga 5). These two components are distinct from the former Ga 2 surface com-

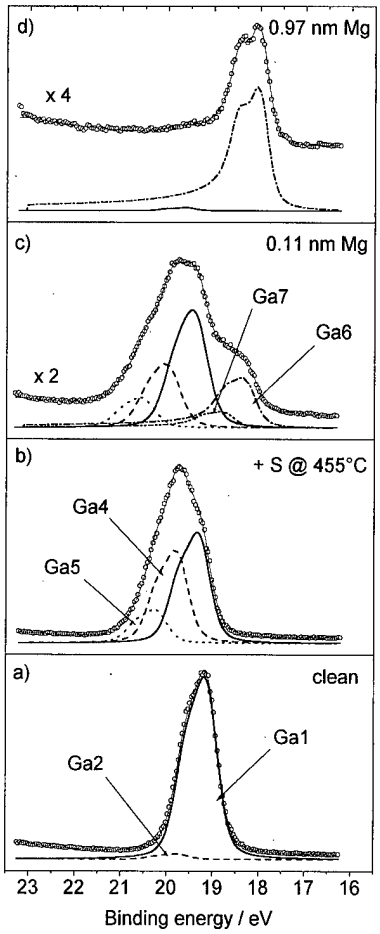


FIG. 2. Ga 3*d* core level spectra measured at 60 eV photon energy: (a) Ga 3*d* spectrum of the clean, As-rich GaAs(100) surface after the decapping at 380 °C; (b) after S modification at 455 °C sample temperature; (c) after deposition of 0.11 nm Mg; (d) after deposition of 0.97 nm Mg.

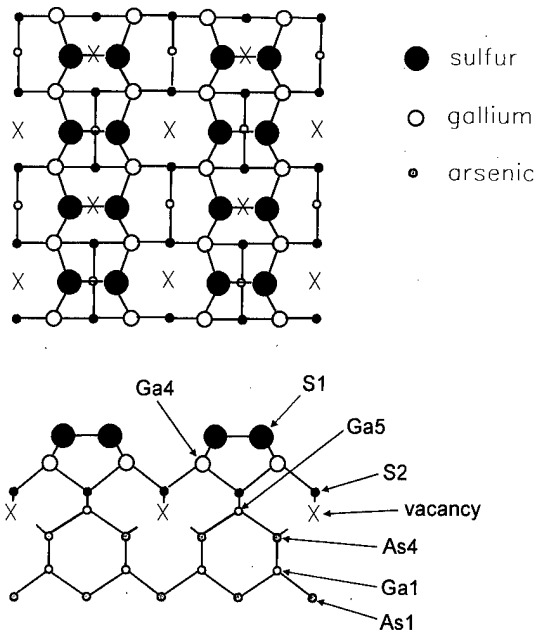


FIG. 3. Model for the (2×1) reconstructed S modified GaAs(100) surface in coincidence with the model proposed by Pashley (see Ref. 20) for (2×1) reconstructed Se modified GaAs(100).

ponent, because of their higher intensity and the distinct shifts in binding energy relative to the bulk component (Ga 1). Here the Ga 4 and Ga 5 components are shifted by 0.47 and 0.93 eV, respectively. These two shifted Ga components originate from Ga–S bonds which have been formed during the S modification of the sample. From angle resolved core level photoemission experiments[21] it is known that at higher surface sensitivity the Ga 4 component increases much more in intensity than the Ga 5 component. This leads to the conclusion that the Ga 4 emission originates from atoms nearer to the surface of the sample than Ga 5. According to Pashley's[20] model this means that the Ga 4 component represents Ga atoms surrounded by S while the Ga 5 is derived from Ga atoms at the interface of the gallium sulfide-like layer and the GaAs bulk (Fig. 3). From this it can be established that an exchange reaction between S and As took place during the S modification.

Calculating the ratio of the total integrated intensities of As 3*d* to Ga 3*d* leads to a value of 0.84 for the clean surface and to a value of 0.58 for the integrated intensities of the bulk components (As 1, Ga 1). After the S modification the first ratio decreases to 0.24 while the latter remains almost the same at 0.51. This is a further indication that all excess As is removed from the clean surface and a new gallium sulfide layer is formed on top of the sample.

The fitting of the S 2*p* spectra is quite difficult due to the poor resolution of the 1800 l/mm grid at 195 eV photon energy, which amounts to approximately 1 eV. Nevertheless two distinct components (S 1, S 2) can be recognized in Fig. 4(a), which is compatible with the results for Se on GaAs(100) found earlier.[21] In accordance with these results the S 1 component is attributed to S atoms on the surface forming dimers and S 2 originates from subsurface S, which is surrounded by Ga atoms as next neighbors. This coincides with electronegativity (EN) arguments, since S has a high EN (Pauling EN: 2.58), so that the electron charge density of S 2 should be higher than for S 1, resulting in a lower binding energy (higher kinetic energy). The energy separation of these two components amounts to 0.92 eV in contrast with 0.50 eV as observed by Moriarty *et al.*[7]

Looking at the As 3*d* spectra during increasing Mg coverages one observes that the As 4 component remains on the lower binding energy side [Fig. 1(c)]. But the shift of this component decreases from 0.75 eV on the S modified surface to a saturation value of 0.55 eV at 0.05 nm of nominal Mg coverage. This is interpreted in terms of the formation of a Mg_3As_2-like compound from the fact that this As 4 component cannot be attributed either to pure As–As bonds or to As–S bonds because both bonds would cause a shift to larger binding energy as can be concluded from the differences in electronegativity. The integrated intensity of the As 4 component shown in Fig. 5 also saturates at a nominal Mg coverage of approximately 0.05 nm, which coincides with the saturation value of the binding energy shift of the As 4 component as mentioned above. This saturation indicates that the Mg–As compound segregates at the surface, because its in-

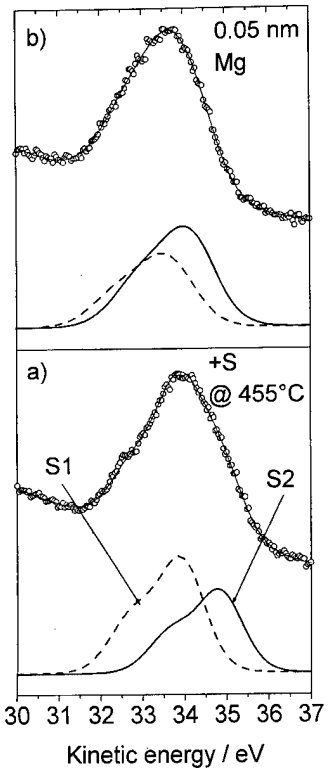


FIG. 4. S 2*p* core level spectra measured at 195 eV photon energy; (a) after S modification of the GaAs(100) sample kept at 455 °C; (b) after deposition of 0.05 nm Mg.

tensity does not decrease upon further Mg deposition. In contrast, the As 1 component exhibits a strong dependence on increasing Mg coverage as can be seen in Fig. 5. From the exponential decrease the escape depth of the photoelectrons from the As 3*d* core level is estimated to be 0.5 nm, which coincides very well with the minimum in the escape depth of photoelectrons at 40 eV kinetic energy which is detected here. From this and the linear shape of the attenuation curve it can be concluded that deposition of Mg causes the formation of a homogeneous layer with negligible islanding.

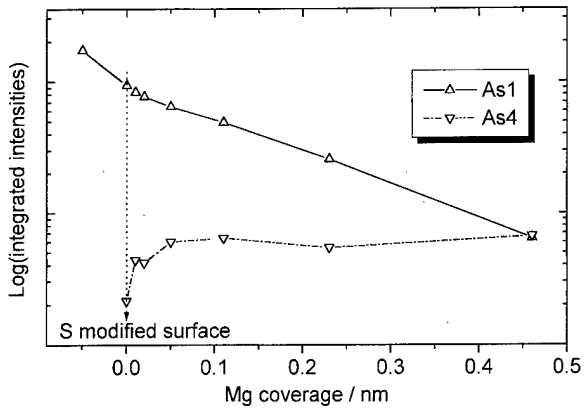


FIG. 5. Integrated intensities of the deconvoluted components in the As 3*d* spectra at different modification steps of the GaAs(100) surface (the first points indicate the clean surface).

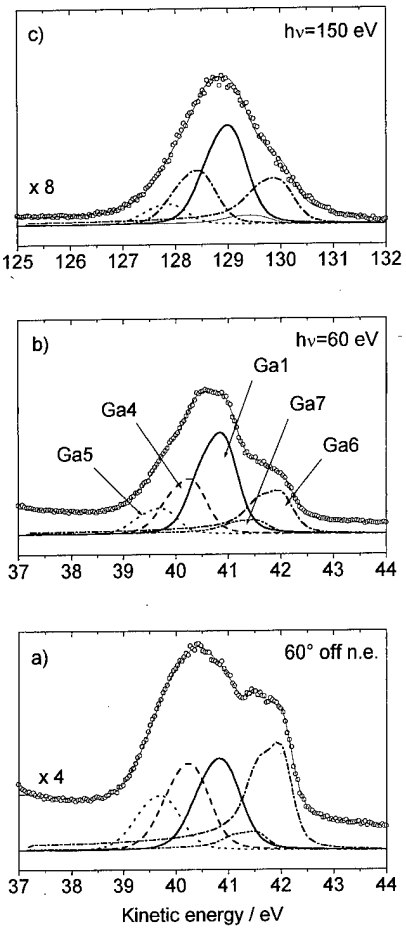


FIG. 6. Ga 3*d* core level spectra for 0.11 nm Mg coverage measured: (a) at 60 eV photon energy and 60° off the normal axis of the sample; (b) at 60 eV photon energy in the direction of the surface normal; (c) at 150 eV photon energy normal to the sample surface.

The Mg deposition also has a strong influence on the Ga 3*d* core level as can be seen in Fig. 2(c). Two new components occur in the Ga 3*d* spectrum on the lower binding energy side, which are shifted by 0.66 eV (Ga 7) and 1.12 eV (Ga 6) towards lower binding energy. They are assigned to metallic Ga in agreement with earlier investigations, where a component in the Ga 3*d* core level spectrum shifted by 0.9 eV relative to the bulk component was attributed to Ga metallic clusters.[15] To distinguish between the bulk and surface components angle resolved spectra were recorded. In Fig. 6(a) a Ga 3*d* core level spectrum was measured at 60° off the normal direction of the sample, which leads to an increased surface sensitivity. It is obvious that the components Ga 4, Ga 5, Ga 6, and Ga 7 are increased in intensity relative to Ga 1 in comparison to the spectrum measured in normal emission in Fig. 6(b). At higher photon energy the escape depth of the photoelectrons is increased, which leads to a higher bulk sensitivity in the spectra. In Fig. 6(c) the Ga 1 component is the most prominent feature in the spectrum, which is in agreement with the interpretation that it is assigned to the bulk.

At higher Mg coverages the Ga 6 component becomes more pronounced in the spectrum and an asymmetric line

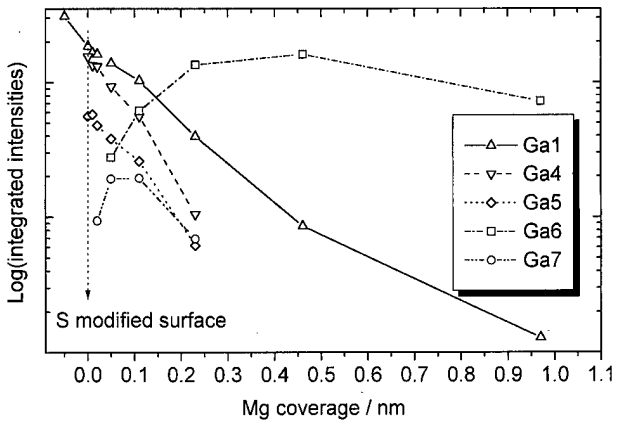


FIG. 7. Integrated intensities of the deconvoluted components in the Ga 3*d* spectra at different modification steps of the GaAs(100) surface (the first point indicates the clean surface).

shape can be clearly observed, which confirms the metallic character of this component (Doniach–Sunjic line shape). The Ga 7 component, which can clearly be distinguished in Figs. 6(a) and 6(b) is attributed to Ga atoms which are dissolved from the GaAs/Ga_xS_y interface. This component is no longer observed at Mg coverages higher than 0.23 nm (Fig. 7), which leads to the conclusion that these Ga atoms now contribute to the Ga clusters represented by Ga 6. The Ga 4 and Ga 5 components also disappear at this Mg coverage as can be seen in Fig. 7. From a comparison of their decay with the decay of the Ga bulk component (Ga 1), which decreases in intensity much stronger, it can be concluded that the intensities of these components are not only attenuated by the Mg layer on the surface but also by the dissolution of the Ga_xS_y compound. The dissolution of the gallium sulfide-like compound is also indicated by the binding energy shift of the two S components in the S 2*p* spectra relative to each other and also by the change in the relative intensity ratio of S 1 to S 2. This is a sign for a change in the chemical environment of the two S components, most likely being due to the formation of a MgS-like compound.

Further proof of this conclusion can be observed in the Mg 2*p* core levels [Fig. 8(a)], where a broad feature (Mg 1) occurs in the spectrum. The Mg 1 component is fitted by only one component of 950 meV Gaussian width. That is done because no consistent deconvolution could be achieved with the application of two Mg components at low Mg coverages up to 0.11 nm nominal thickness. This is due to the fact that the Mg 2*p* core level has only a very small spin-orbit splitting (Table I) and there are also no asymmetries or shoulders visible in the core level emission. So the broad Mg 1 peak covers the two proposed chemical environments of Mg, the MgS-like and the Mg_3As_2-like compounds, mentioned above. First at 0.23 nm Mg coverage a slight asymmetry is observed in the Mg 2*p* spectrum assigned to a second component (Mg 2), which can be clearly separated at 0.46 nm Mg deposition as seen in Fig. 8(b). The Gaussian width of this new component is much smaller than for the Mg 1 component and it also exhibits an asymmetry due to its

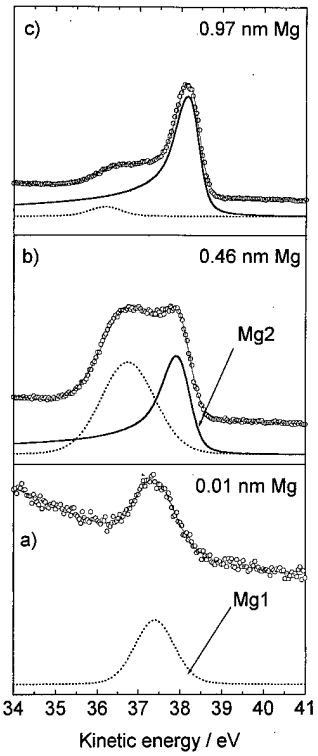


FIG. 8. Mg 2*p* core level spectra measured at 89 eV photon energy after the deposition of: (a) 0.01 nm Mg, (b) 0.46 nm Mg, and (c) 0.97 nm Mg onto the S modified GaAs(100) surface.

metallic character (Doniach–Sunjic line shape). This metallic component becomes even more prominent after 0.97 nm deposition of Mg [Fig. 8(c)] and it covers the MgS-like compound as can be evaluated from the intensity ratio of Mg 2 to Mg 1. The occurrence of metallic Mg at 0.23 nm Mg coverage shows that all the S bound in the gallium sulfide-like layer on top of the GaAs(100) sample is consumed by the MgS formation, which coincides with the disappearance of the gallium sulfide related components (Ga 4, Ga 5) in the Ga 3*d* core level spectra. The observation of metallic Mg at 0.23 nm Mg coverage coincides also with the occurrence of a Fermi edge distinguished in the valence band spectra of the sample, which is another indication for the existence of metallic Mg on top of the sample.

B. Electronic properties

The electronic properties of the S modified Mg/GaAs(100) Schottky contact can be derived from the position of the Fermi level relative to the valence band maximum. The measurement of the Fermi level is quite difficult in the case of adsorbate deposition, because no sharp valence band edge can be observed at higher adsorbate coverages. To determine the Fermi level position at higher Mg coverages first the Fermi level of the clean *n*- and *p*-doped GaAs(100) surfaces in relation to the valence band maximum was carefully determined leading to 0.48 and 0.37 eV, respectively (Fig. 9). In a second step the binding energies of the Ga $3d_{5/2}$ (*n*-GaAs: 19.19 eV, *p*-GaAs: 19.12 eV) and As $3d_{5/2}$ (*n*-GaAs: 41.09 eV, *p*-GaAs: 40.98 eV) components were

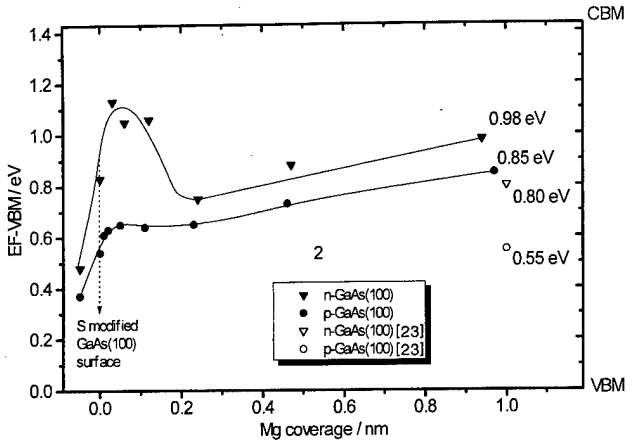


FIG. 9. Position of the Fermi level on *n*- and *p*-type GaAs(100) at different modification steps, in particular Mg coverages. The value finally achieved is compared to the data determined by $I-V$ and $C-V$ measurements of unmodified Mg/GaAs(100) Schottky contacts (see Ref. 23).

measured in relation to the Fermi level position. From these values the energies of the Fermi level stated above are subtracted leading to 18.73 eV (Ga $3d_{5/2}$) and 40.61 eV (As $3d_{5/2}$) coincident for *n*- and *p*-doped GaAs(100), which are approximately 0.10 eV lower than the values of Eastman *et al.*[18] In the following the binding energies of Ga 3*d* and As 3*d* are measured relative to the Fermi level so that the subsequent subtraction of the binding energies relative to the Fermi level gives the position of the Fermi level relative to the valence band maximum.[22] Only the values derived from the shift of the Ga 3*d* bulk component (Ga 1) are shown because they are in agreement with the values derived from the As 3*d* bulk component (As 1).

After the S modification of the sample the Fermi level is shifted by 0.35 and 0.17 eV towards the conduction-band minimum (CBM), respectively. Considering now the condition on the clean *n*-GaAs(100) surface, it can be assumed that there are acceptorlike surface states in the band gap, which are only partially filled by electrons from the valence band, because they cross the Fermi level caused by the upward band bending. This surface charge is compensated by the space charge of the immobile, ionized bulk donors, which are also lifted above the Fermi level by the band bending. The S modification causes the formation of gallium sulfide with sulfur nearest to the surface. Due to the larger EN of S (2.58, Pauling) relative to Ga (EN_{Ga}: 1.81) a surface dipole is induced with higher electron density at the S atoms on top of the sample. The "positive" Ga atoms at the interface can now compensate the negative charge of the acceptorlike states in the band gap, which induces a reduction in band bending. On the other hand this surface dipole induces an additional charge at the interface to the donorlike states which must be compensated by an increased band bending. Due to the deposition of Mg an exchange takes place between Mg and Ga leading to the formation of MgS, which has a much lower EN (EN_{Mg}: 1.31) than Ga, so that the large overshoot of the Fermi level 0.3 eV can be explained by the formation of a much stronger dipole interlayer, which leads to a further reduction in band bending. The Fermi level remains at approximately 1.05 eV up to 0.23 nm Mg coverage and drops thereafter to 0.75 eV. This point coincides with the occurrence of metallic Ga and Mg in the related core level spectra, which leads to the conclusion that the charge of the interface dipole is compensated by the metallic film. The overshoot on the *p*-doped GaAs is much smaller, which needs further investigation. At 1.0 nm Mg coverage the Fermi level position amounts to 0.98 eV on the *n*- and 0.85 eV on the *p*-type sample, which is 0.18 and 0.30 eV higher than the values observed for unmodified Mg/GaAs(100) Schottky contacts by current–voltage ($I-V$) and capacitance–voltage ($C-V$) measurements for *n*- and *p*-doped samples, respectively.[23] Calculating the sum of the barrier heights

$$\Phi_{Bn}+\Phi_{Bp}=0.44\ \text{eV}+0.85\ \text{eV}=1.29\ \text{eV}$$

leads to a value, which is closer to the band gap of GaAs [1.42 at room temperature (RT)] than the value observed for the unmodified Schottky contact, which amounts to 1.17 eV.[23]

IV. SUMMARY

The S modification of clean GaAs(100) surfaces at 435–455 °C was investigated by SXPS, which leads to the conclusion that the excess As disappears from the surface and a gallium sulfide-like layer is formed. The deposition of Mg thereafter leads to the formation of a MgS-like compound which is accompanied by the occurrence of metallic Ga atoms. From this and the fact that the chemical environment of the S changes it is concluded that the gallium sulfide is dissolved and the remaining S is consumed by the formation of MgS. Additionally the formation of a small amount of a Mg_3As_2 like compound is observed. First at 0.23 nm of Mg thickness a metallic Mg component is observed. The determination of the Fermi level shows a reduction in band bending of 0.35 eV on *n*-doped GaAs while it is increased by 0.17 eV on *p*-doped GaAs. This Fermi level movement in the same direction for *n*- and *p*-type GaAs is indicative of a shifting of the density of interface states within the band gap rather than removing the states from the gap. The Schottky barrier heights finally achieved on *n*- and *p*-type GaAs(100) amount to 0.44 and 0.85 eV, respectively, in comparison to 0.62 and 0.55 eV determined by electrical measurements[23] for unmodified Mg/GaAs(100) contacts.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support by the "Bundesministerium für Bildung und Forschung" under Grant No. 05 622OCA 3. They also thank D. Westwood from the Department of Physics and Astronomy of the University of Wales, College of Cardiff for preparing the substrates and the BESSY staff for their help during the measurements.

[1]J. Shin, K. M. Geib, C. W. Wilmsen, and Z. Lilliental-Weber, J. Vac. Sci. Technol. A **8**, 1894 (1990).

[2]H. Sugahara, M. Oshima, H. Oigawa, H. Shigekawa, and Y. Nannichi, J. Appl. Phys. **69**, 4349 (1991).
[3]Z. S. Li, W. Z. Cai, R. Z. Su, G. S. Dong, D. M. Huang, X. M. Ding, X. Y. Hou, and X. Wang, Appl. Phys. Lett. **64**, 3425 (1994).
[4]X. Wang, X.-Y. Hou, Z.-S. Li, and X.-Y. Chen, Surf. Interface Anal. **24**, 564 (1996).
[5]V. N. Bessolov, A. F. Ivankov, and M. V. Lebedev, J. Vac. Sci. Technol. B **13**, 1018 (1995).
[6]V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, J. Vac. Sci. Technol. B **14**, 1 (1996).
[7]P. Moriarty, B. Murphy, L. Roberts, A. A. Cafolla, G. Hughes, L. Koenders, and P. Bailey, Phys. Rev. B **50**, 14237 (1994).
[8]J. R. Waldrop, J. Vac. Sci. Technol. B **3**, 1197 (1985).
[9]G. W. Anderson, M. C. Hanf, P. R. Norton, M. Kowalewski, K. Myrtle, and B. Heinrich, J. Appl. Phys. **79**, 4954 (1996).
[10]R. Berrigan, Y. Watanabe, T. Scimeca, and M. Oshima, Jpn. J. Appl. Phys., Part 1 **31**, 3223 (1992).
[11]M. Oshima, T. Scimeca, M. Sugiyama, S. Maeyama, H. Oigama, N. Nannichi, and H. Hashizume, Appl. Surf. Sci. **70/71**, 496 (1993).
[12]M. Larive, G. Jezequel, J. P. Landesman, F. Solal, J. Nagle, B. Lépine, A. Taleb-Ibrahimi, G. Indlekofer, and X. Marcadet, Surf. Sci. **304**, 298 (1994).
[13]J. H. Neave, P. K. Larsen, J. F. van der Veen, P. J. Dobson, and B. A. Joyce, Surf. Sci. **133**, 267 (1983).
[14]I. M. Vitomirov, A. D. Raisanen, A. C. Finnefrock, R. E. Viturro, L. J. Brillson, P. D. Kirchner, G. D. Pettit, and J. M. Woodall, J. Vac. Sci. Technol. B **10**, 1898 (1992).
[15]G. LeLay, D. Mao, A. Kahn, Y. Hwu, and G. Margaritondo, Phys. Rev. B **43**, 14301 (1991).
[16]J. F. van der Veen, P. K. Larsen, J. H. Neave, and B. A. Joyce, Solid State Commun. **49**, 659 (1984).
[17]P. K. Larsen, J. H. Neave, J. F. van der Veen, P. J. Dobson, and B. A. Joyce, Phys. Rev. B **27**, 4966 (1983).
[18]D. E. Eastman, T.-C. Chiang, P. Heimann, and F. J. Himpsel, Phys. Rev. Lett. **45**, 656 (1980).
[19]St. Hohenecker, D. Drews, M. Lübbe, D. R. T. Zahn, and W. Braun, Appl. Surf. Sci. **123/124**, 120 (1998).
[20]M. D. Pashley and D. Li, J. Vac. Sci. Technol. A **12**, 1848 (1994).
[21]St. Hohenecker, T. Werninghaus, D. R. T. Zahn, and W. Braun, J. Electron Spectrosc. (accepted).
[22]E. A. Kraut, R. W. Grant, J. R. Waldrop, and P. Kowalczyk, Phys. Rev. Lett. **44**, 1620 (1980).
[23]J. R. Waldrop, J. Vac. Sci. Technol. B **2**, 445 (1984).
[24]L. Geelhaar, R. A. Bartynski, F. Ren, M. Schnoes, and D. N. Buckley, J. Appl. Phys. **80**, 3076 (1996).

Direct imaging of the evolving Au/InSb(111) *B* interface

Tetsuya Mishima[a] and Jun Nakamura
Kagami Memorial Laboratory for Materials Science and Technology, Waseda University, Nishiwaseda, Shinjuku-ku, Tokyo 169-0051, Japan

Kenji Tsukada, Masayasu Nishizawa, and Toyoaki Eguchi
Department of Materials Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-0072, Japan

Toshiaki Osaka
Department of Materials Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-0072, Japan and Kagami Memorial Laboratory for Materials Science and Technology, Waseda University, Nishiwaseda, Shinjuku-ku, Tokyo 169-0051, Japan

(Received 21 January 1998; accepted 11 May 1998)

In situ high-resolution transmission electron microscopy in the profile geometry has been used to observe the evolving features of the Au/InSb(111) *B*-(2×2) interface. During Au deposition in the range between 0 monolayer (ML) and ~1 ML coverage, the outermost Sb-trimer layer of the InSb(111) *B*-(2×2) substrate changes in contrast, presumably revealing that deposited Au atoms are partially captured into it. At ~2 ML coverage, an unknown phase emerges on the outermost layer, beyond which it continues to grow epitaxially in an island state, causing partial disruption of the substrate. The phase is identified as Au_9In_4 alloy with a γ-brass structure determined from a digital Fourier transform diffractogram and a transmission electron diffraction pattern. The epitaxial relationship of Au_9In_4 with the substrate is given by (111) InSb‖(111) Au_9In_4 and $[1\bar{1}0]$ InSb‖$[1\bar{1}0]$ Au_9In_4. The high resolution-profile transmission electron microscopy images of this alloy agree well with the results calculated by the multislice method. © *1998 American Vacuum Society.* [S0734-211X(98)05604-2]

I. INTRODUCTION

Metal/III–V compound semiconductor systems have been extensively studied because of the great interest for higher speed electronic devices. As is well known, the properties of the heterostructures in such systems are dominated by the interfaces formed between the metal and the substrate. However, it is hard for most of the systems to form ideally abrupt interfaces, because metals frequently attack III–V substrates.[1] For example, at low coverages and even at room temperature, the metals easily react with species constituting the substrates, except for systems such as Au/GaAs, which forms an abrupt interface,[1,2] and one or both of the III–V species are not scarce to segregate toward the growth front. Thus, in order to design desired interfaces in the systems, we have to serve detailed features obtained by pursuing dynamically formation processes of the interfaces as a precise interface control.

Although there are several analytical methods for *in situ* observation of thin film growth behavior, most of them are not suitable for the real time observation of an interface which is being embedded. High resolution transmission electron microscopy in the profile geometry (HR-profile TEM) is one of the methods, which enables us to directly image the evolving interface on an atomic scale. Recently, we observed the reconstructed structures characteristic of InSb{111} *A*,*B*-(2×2) surfaces by using this method and confirmed their previously proposed structure models.[2] In the present study, we apply the HR-profile TEM to observe *in situ* the Au/InSb(111) *B*-(2×2) interface formed by molecular-beam epitaxy (MBE) at room temperature.

II. EXPERIMENT

A detailed description of the apparatus used in this study has been given in our previous paper.[3] A 200 kV conventional HR-TEM (JEOL JEM-2010) was converted to an ultrahigh vacuum (UHV) *in situ* HR-TEM (henceforth it is referred to as UHV-TEM) equipped with a bakable stainless-steel (SUS304) specimen chamber, to which two MBE units are attached. The ultimate base pressure in the chamber was less than 3×10^{-10} Torr and the pressure during In, Sb, and Au evaporation was below 1×10^{-9} Torr. Samples were prepared for the profile observation from InSb single crystals (nondoped) by crushing them with an agate mortar under purified ethanol. The resulting fragmentary specimens were then mounted onto a holey carbon microgrid. For the plan-view observation, InSb(111) *B* wafers (nondoped) with diameters of 3 mm were mechanically thinned and then chemically etched with a mixture of lactic acid-HNO_3(10:1) solution to form a round hole about 0.1 mm in diameter with electron-transparent peripheries. In the specimen chamber, the native oxide was removed from the (111) *B* surface by heating the specimens at 460 °C by impinging Sb_4 molecules at a rate of 3.5×10^{14} atoms/cm^2 min. Immediately after the cleaning, several nm of InSb were homoepitaxially grown on the specimen surface at 300 °C with In_1 (2.8×10^{14} atoms/cm^2 min) and Sb_4 (1.4×10^{15} atoms/cm^2 min) molecular beams. The homoepitaxially treated surfaces give a HR-profile TEM image corresponding to a T_4 site Sb-

[a]Electronic mail: mishima@mn.waseda.ac.jp

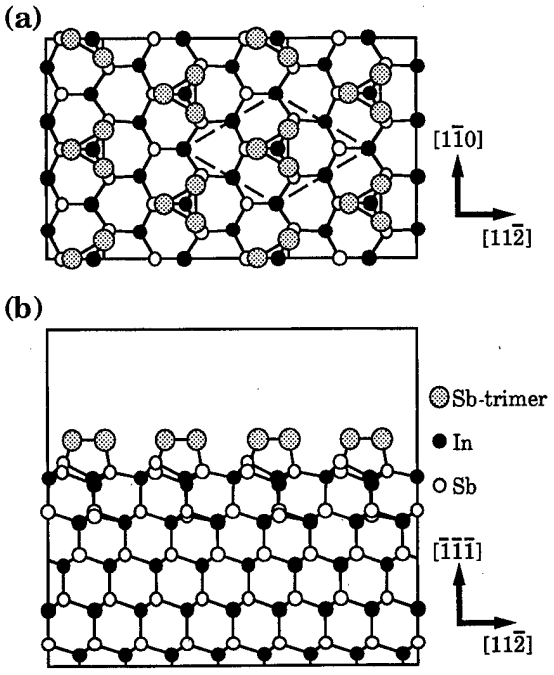


FIG. 1. (a) Top view of the T_4 site Sb-trimer model for the InSb(111) B-(2×2) surface. A (2×2) unit cell is outlined by dashed lines. (b) Side view of (a).

trimer structure proposed by Nakada *et al.*[4] [Fig. 1(a)]. Subsequently, Au was evaporated by heating a W basket onto the room-temperature substrate at a rate of ~1.1 $\times 10^{14}$ atoms/cm^2 min, corresponding to ~0.1 monolayer (ML)/min. The deposition rate and film thickness were monitored with a quartz-crystal oscillator. The HR-profile TEM images were taken along the [1$\bar{1}$0] direction and recorded on a video tape recorder by an on-line TV camera (Gatan 622SC) attached to the UHV-TEM. The typical electron-beam current densities at the specimens were ~1 A/cm^2. While observing the magnified images, no changes in image detail arising from electron beam irradiation were detected. The multislice simulation[5] for profile images was performed using the commercial MACTEMPAS program. In this simulation, a spherical and a chromatic aberration coefficient of 0.9 and 1.0 mm for the UHV-TEM are used. Furthermore, in analyzing a digital Fourier transform diffractogram, we used the NIH IMAGE program.

III. RESULTS AND DISCUSSION

Figures 2(a)–2(e) show a series of HR-profile TEM images of the Au/InSb(111) B interface, taken during Au deposition on the clean InSb(111) B-(2×2) surface in the range of 0–~3 ML coverage at room temperature. A typical HR-profile TEM image of the InSb(111) B-(2×2) substrate prior to the Au deposition is shown in Fig. 2(a). The substrate surface is characterized by a profile of "sawtooth" type with a ×2 periodicity, which closely resembles the [110] projected Sb-trimer model shown in Fig. 1(b). In Fig. 2(b), below ~0.5 ML coverage, changes in contrast at the first layer composed of the Sb trimers are clearly seen, a part

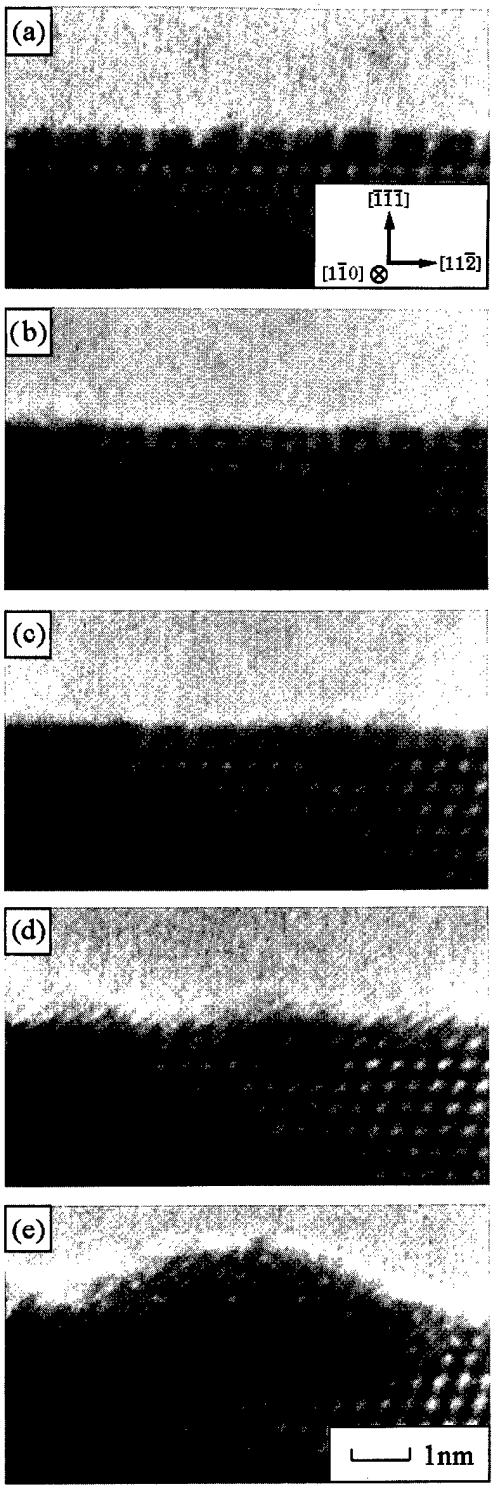


FIG. 2. Series of high resolution profile images of the evolving Au/InSb(111) B interface in the range between 0 and ~3 ML Au coverage, taken at a defocus value of −36 nm: (a) the InSb(111) B-(2×2) substrate before Au deposition. Profiles of the Sb trimers are clearly seen. (b) Below ~0.5 ML, the first layer of Sb trimers changes in contrast. (c) Beyond ~0.5 ML, the first layer is converted to the (1×1) structure. (d) At ~2 ML coverage, an Au_9In_4 alloy phase is formed on the outermost layer, partially disrupting its (1×1) structure. (e) At ~3 ML coverage, the Au_9In_4 phase becomes an island and has an epitaxial relationship with the substrate.

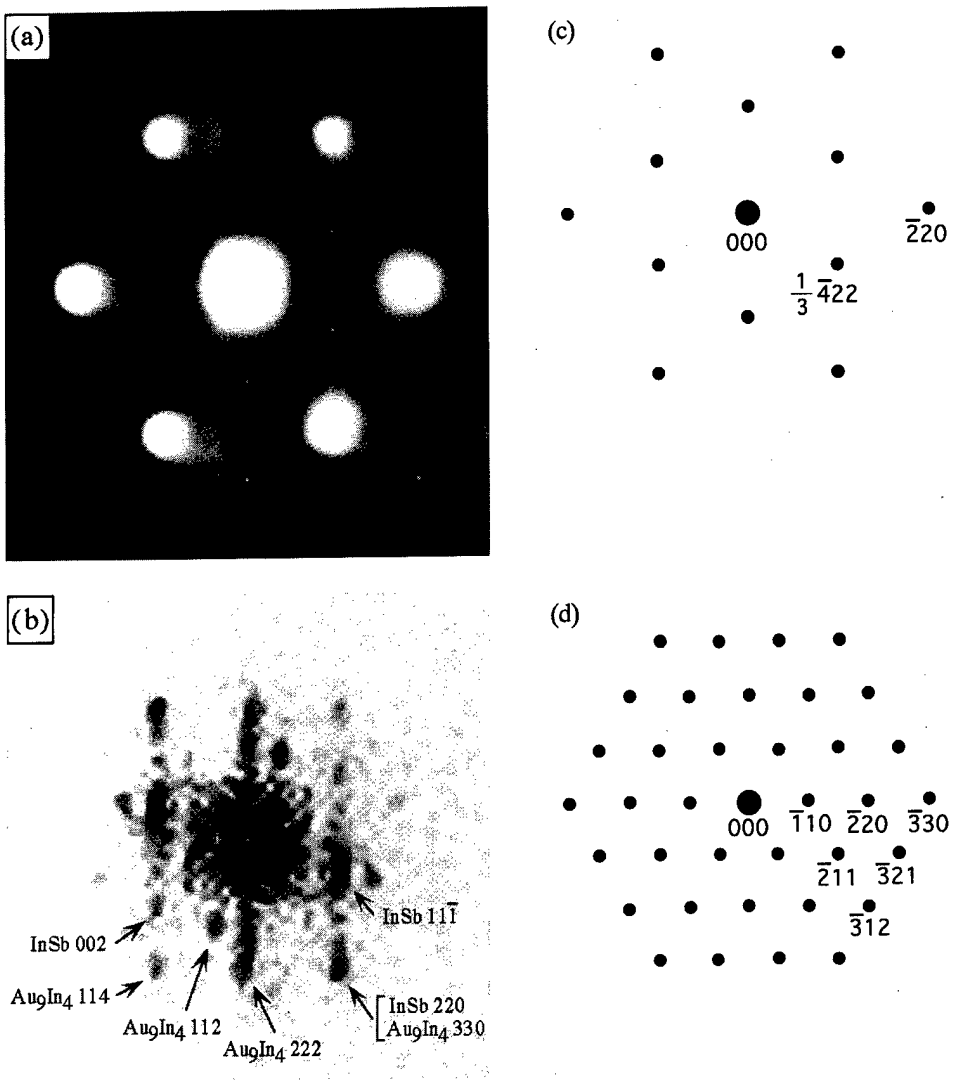


FIG. 3. (a) TED pattern taken from the specimen which includes the Au/InSb(111) *B* interface at room temperature. (b) Digital Fourier transform diffractogram computed for the lattice image of Fig. 2(e). Schematic drawings of diffraction patterns for (c) InSb and (d) Au_9In_4, shown from the [111] direction.

of them having no trace. This suggests at this stage that one of the elements constituting the substrate, Sb, begins to segregate on the Au deposited surface, because it is known that in combinations of InSb with I_b group elements such as Au, Al, and Cu, Sb easily segregates on these metal surfaces.[1] Since Sb has five valence electrons, it would be triply bonded to three atoms in the metal surfaces, leaving a lone pair state at the surface. Such a surface is expected to be become highly stable. However, we consider that the deposited Au atoms change the energy configuration of the stable Sb–Sb bonds in the trimer to weaken or to break them. On the other hand, the increase of diffraction contrast in the whole Sb-trimer layer, which is visible in Fig. 2(b), is attributed to the enhancement of a scattering factor caused by capturing the Au in the partially bond-weakened Sb-trimer layer. Beyond ~0.5 ML, as seen in Fig. 2(c), a large part of the Sb-trimer structure changes to a period of (1×1). Further Au deposition (e.g., ~2 ML) leads to the formation of a lattice fringe having a distance of ~0.23 nm, parallel to the $[00\bar{1}]$ direction of the substrate, and to the partial disruption of the (1×1) structure, as shown in Fig. 2(d). This *d* spacing is very close to that of the {220} plane of InSb. These findings indicate that this newly formed lattice is in a given epitaxial relationship with the substrate.

X-ray photoemission spectroscopy studies of Au/InSb(110) similar to the present system evidenced disruption at the interface and intermixing of deposited Au atoms with the substrate surface at low coverage.[6] Thus, although we have to consider the ~0.23 nm lattice fringe from which of Au–Sb, Au–In, or Au–(In, Sb) alloy stems, its identification will be made later. With an increase of Au coverage, the alloy grows in an island state. Shown in Fig. 2(e) is the profile image of an island at ~3 ML Au coverage, in which the spacing of the lattice fringe is also ~0.23 nm, remaining unchanged up to ~5 ML coverage. Using a ~5 ML Au/InSb specimen under otherwise almost identical conditions with the HR-profile TEM case, we observed its transmission electron diffraction (TED) pattern in the plan-view geometry. At first glance, the TED pattern, shown in Fig. 3(a), very closely resembles that of a $(\sqrt{3}\times\sqrt{3})$-$R30°$ structure which often emerges in a variety of metal/semiconductor systems. However, the island profile in Fig. 2(e) strongly suggests that this

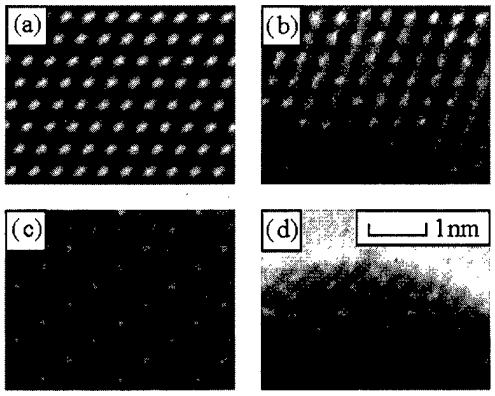


FIG. 4. (a) Simulated and (b) observed image for the InSb substrate. (c) Simulated and (d) observed image for the Au_9In_4 island. The observed images are a part of Fig. 2(e). On the basis of the defocus value of −36 nm, the crystal thickness is estimated to be 5 nm for the substrate and 4 nm for the island.

pattern stems not from a $(\sqrt{3}\times\sqrt{3})$ adsorption structure of Au on InSb(111) *B* but from the island itself. By examining the $(\sqrt{3}\times\sqrt{3})$-type TED pattern and a digital Fourier transform diffractogram [Fig. 3(b)] computed for the lattice image of Fig. 2(e), we determined the island phase to be composed of Au_9In_4 alloy with a γ-brass structure.[7] The epitaxial relationship between the Au_9In_4 island and the substrate is given by (111) InSb‖(111) Au_9In_4 and $[1\bar{1}0]$ InSb‖$[1\bar{1}0]$ Au_9In_4. Figures 3(c) and 3(d) show schematic diffraction patterns of InSb and Au_9In_4 depicted from the [111] direction, respectively. By overlaying these two patterns, we obtain a pattern similar to the $(\sqrt{3}\times\sqrt{3})$-type TED pattern of Fig. 3(a). The lattice mismatch between InSb(220) ($d=0.229$ nm) and Au_9In_4(330) ($d=0.232$ nm) is a very small value of 1.3%. This is also supported by the fact that no extra reflections produced by double diffraction are observed in the TED pattern [Fig. 3(a)].

Multislice simulations for the profile image of Fig. 2(e) were performed to confirm the presence of the Au_9In_4 phase. Figure 4(a) shows one of the through focal images simulated, from which the defocus value for Fig. 2(e) and the crystal thickness of the substrate are estimated. As seen in Fig. 4(a), the simulated image for a defocus value of −36 nm and a crystal thickness of 5 nm agrees well with the observed one [Fig. 4(b)]. Next, using this defocus value, we simulated images for Au_9In_4 ranging from 2 to 5 nm crystal thickness. As a result, the simulated image for an island thickness of 4 nm [Fig. 4(c)] was in reasonable agreement with the observed one [Fig. 4(d)].

The chemical reactivities of Au with the representative III–V compounds increase in the order of InSb>GaAs>InP, because their heats of compound formation are −30.5, −71, and −88.7 kJ mol^{-1}, respectively.[1] This trend elucidates Au to most easily react with the InSb, which is confirmed by the room-temperature interface disruption occurring only in this system and also by the formation of flat and abrupt interfaces in the other systems. However, here, the question arises as to why a phase of Au_9In_4 (cubic, $a_0=0.983$ nm) is initiated with Au deposition, because, taking into account a lattice mismatch of the Au–In alloy with the InSb substrate, $AuIn_2$ (cubic, $a_0=0.6157$ nm) fits the substrate better than Au_9In_4 (0.6% for the former and 1.3% for the latter), and because the $AuIn_2$ phase is more stable in the wide temperature range of room temperature and 540.7 °C. This question can be answered as follows: As mentioned above, no deposited Au atoms could meet In atoms up to ~0.5 ML coverage until the Sb segregated onto the growth front. This means that at the very early stage of Au deposition, the Au has time to become considerably rich compared with the In on the substrate side. Thus, it is most likely that the situation suitable for forming the Au-rich Au–In alloy and a relatively good lattice fit of 1.3% in the Au_9In_4/InSb system allows a phase of Au_9In_4 to emerge.

IV. SUMMARY

We studied the evolving interface of the Au/InSb(111) *B* system in Au coverages of 0–~5 ML at room temperature by using HR-profile TEM, multislice image simulations and TED in the plan-view geometry. The HR-profile TEM observation revealed that the first Sb-trimer layer of the InSb(111) *B*-(2×2) substrate gradually altered its (2×2) structure to the (1×1) period. The subsequent Au deposition led to partial disruption of the substrate and to island formation of Au_9In_4 by reaction of the deposited Au atoms with the substrate. The Au_9In_4 island with a γ-brass structure has the following epitaxial relationship with the substrate: (111) InSb‖(111) Au_9In_4 and $[1\bar{1}0]$ InSb‖$[1\bar{1}0]$ Au_9In_4. The HR-profile TEM images of this alloy agreed well with the results calculated by the multislice method.

ACKNOWLEDGMENTS

The authors would like to acknowledge useful discussions with S. P. Cho. They wish to thank Dr. Y. Matsui (NIRIM) and Dr. Y. Harada, T. Hata, Y. Kondo, and Dr. M. Takeguchi (JEOL Ltd.) for their technical support. This work was partially supported by a Grant-in-Aid for Scientific Research (B) from the Ministry of Education, Science, Sports and Culture, by a Waseda University grant for Special Research Projects and by a Grant-in-Aid for Research for the Future (Next-Generation Process Technologies; wafer-scale fabrication processes of quantum dots) from the Japan Society for the Promotion of Science.

[1]D. M. Hill, F. Xu, Z. Lin, and J. H. Weaver, Phys. Rev. B **38**, 1893 (1988).
[2]Z. Liliental-Weber, J. Vac. Sci. Technol. B **5**, 1007 (1987).
[3]T. Mishima and T. Osaka, Surf. Sci. **395**, L256 (1998).
[4]T. Nakada and T. Osaka, Phys. Rev. Lett. **67**, 2834 (1991).
[5]J. M. Cowley and A. F. Moodie, Acta Crystallogr. **10**, 609 (1957).
[6]Y. Shapira, F. Boscherini, C. Capasso, F. Xu, D. M. Hill, and J. H. Weaver, Phys. Rev. B **36**, 7656 (1987).
[7]J. K. Brandon, R. Y. Brizard, W. B. Pearson, and D. J. N. Tozer, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **33**, 527 (1977).

Microwave modulated photoluminescence used to measure surface recombination velocities

C. E. Inglefield,[a] M. C. DeLong, and P. C. Taylor
Physics Department, University of Utah, Salt Lake City, Utah 84112

W. A. Harrison
Department of Applied Physics, Stanford University, Stanford, California 94305

(Received 21 January 1998; accepted 26 May 1998)

Microwave modulated photoluminescence (MMPL) is a characterization technique in which a semiconducting sample is subjected to continuous optical pumping and chopped microwave electric fields. The signal normally detected in an MMPL experiment is the change in the photoluminescence (PL) spectrum due to the presence of the microwave electric field, which increases the kinetic energy of the free carriers. We have previously correlated the quenching of the PL signal, as induced by the microwaves, with nonradiative recombination at a surface/interface of the photoexcited volume. In this work, we determine quantitatively surface recombination velocities through a combined measurement of microwave induced changes in photoconductivity and in PL. From the change in the photoconductivity we infer a change in the diffusion constant of free carriers in the material. The change in diffusion constant, along with the change in luminescent intensity, uniquely determines the surface recombination velocity of the layer. Results for GaAs layers with bare surfaces are presented and the potential usefulness of the technique to other material systems, including the measurement of properties of buried interfaces, is discussed. © *1998 American Vacuum Society.* [S0734-211X(98)10704-7]

I. INTRODUCTION

A. Microwave modulated photoluminescence

In many cases, as with other modulation spectroscopies, more information about a sample is obtained from microwave modulated photoluminescence (MMPL) than from photoluminescence (PL) alone. In the specific case described here, MMPL is used to measure the surface recombination velocity (SRV) of a layer. Although the luminescence efficiency clearly depends on the nonradiative lifetime and hence the SRV, it is not possible to extract an SRV from a single PL measurement on a single layer. In what follows we will describe how the SRV may be extracted from the dependence of the PL and the sample conductivity on a perturbation caused by the presence of a microwave electric field.

In an MMPL apparatus, the sample is placed in the electric field maximum of a microwave cavity. The microwave cavity has optical access through which we excite photocarriers and collect luminescence. The microwaves are normally switched on and off while continuously illuminating the sample with above-band-gap light, and a lock-in amplifier is used to detect the change in the luminescence caused by the presence of the microwaves.

The direct effect of the microwaves is to increase the kinetic energy of the free carriers in the sample.[1] One type of MMPL signal arises from the fact that this energy may be transferred from the carriers to the sample lattice and cause sample heating.[2,3] The MMPL signal in this case is a manifestation of the dependence of the PL on sample temperature. In this work, such sample heating effects have been made negligible by immersing the sample in superfluid helium during the measurements.[2] All the MMPL signals discussed here are therefore manifestations of the dependence of the PL on free-carrier energy with the sample temperature held at 2 K. In the past, MMPL signals unrelated to any heating of the crystal lattice have been attributed to impact ionization[4] or to the dependence of the cross section for a recombination path on the free-carrier energy.[5] In the discussion that follows we will show that the MMPL signals we present in this work are a manifestation of an increase of the diffusion constant of the free carriers as their kinetic energy is increased by the presence of the microwave electric field.

B. Nonradiative recombination at surfaces and interfaces

The nonradiative lifetime of carriers due to recombination at a surface or interface has been shown[6] to be well approximated by the expression:

$$\tau_{\mathrm{nr}}=\frac{4d^2}{\pi^2 D}+\frac{d}{S}, \tag{1}$$

where S is the recombination velocity of the surface or interface, d is the thickness of the active layer (in our case of photogenerated carriers this is the light absorption length), and D is the diffusion constant of minority carriers within the active layer. Equation (1) differs by numerical factors from the expression found in Ref. 6 due to differences in sample geometry. The nonradiative lifetime and the radiative lifetime, τ_r, determine the luminescent efficiency, η, as follows:

[a]Electronic mail: colin@mail.physics.utah.edu

$$\eta = \frac{\tau_{nr}}{\tau_r + \tau_{nr}}. \quad (2)$$

In a previous study,[7] we established that the microwave electric fields produce changes in D which affect the PL efficiency. The amount by which the PL efficiency changes depends on the value of S. In this work we measure changes in luminescence efficiency and conductivity[8] to infer changes in D and τ_{nr} that uniquely determine S.

An established method for obtaining S using luminescence is to measure the dependence of the PL lifetime on the thickness of a layer, d, in Eq. (1).[6] This technique obviously requires the growth of layers with similar surface properties and different thicknesses, and it is fundamentally limited in accuracy by the ability to grow samples with the same surface properties. The technique described in this work is based on modulating the diffusion constant, D, in Eq. (1) for a single sample and does not require the growth of multiple samples for any single measurement of S.

At low temperatures (all the experiments discussed here were performed at 2 K) the mobility of free carriers in GaAs is expected to be determined (limited) by ionized impurity scattering.[9] The scattering time associated with this scattering mechanism is expected to increase with carrier temperature. Hence the mobility of free carriers is expected to increase with carrier temperature in our experiment. The microwave electric fields in our experiment cause an increase in the kinetic energy of the free carriers,[2] which produces a measurable decrease in the resistance of the sample. The increase in mobility also produces an increase in diffusion constant, as the two quantities are connected by the Einstein relation.[10] As shown in Eq. (1), the increase in diffusion constant causes a decrease in the nonradiative lifetime, which leads to a change in luminescent efficiency—another measurable quantity.

II. EXPERIMENT

Luminescence was excited with 5145 Å light from a Coherent Ar^+ ion laser with an intensity on the order of 0.3 W/cm^2. The GaAs (100) samples were placed in the electric field maximum of a 16 GHz transverse electric $(TE)_{001}$ microwave cavity. Microwave powers of 100 mW and less were used. Luminescence was dispersed with a 0.85 m Spex double monochromator and detected with a GaAs photomultiplier. The microwaves were chopped at the reference frequency of a lock-in amplifier so that microwave-induced changes in luminescence could be measured by standard lock-in techniques. Resistance was measured with an ohmmeter by attaching leads to the sample using In contacts on either side of the laser-illuminated area. The sample and cavity were immersed in superfluid He (2 K) during the measurements.

Samples were chosen to be lightly doped (impurity concentrations less than 10^{17} cm^{-3} for p type and less than 10^{15} cm^{-3} for n type) so that the impurities would not contribute significantly to the conductivity (with the laser on) at 2 K (i.e., the carriers were "frozen out").[11] Hence, all electrical measurements made can be attributed to properties of photogenerated carriers; the sample resistance was at least two orders of magnitude greater with the laser off than with the laser on. The importance of this choice of samples will be discussed in more detail later.

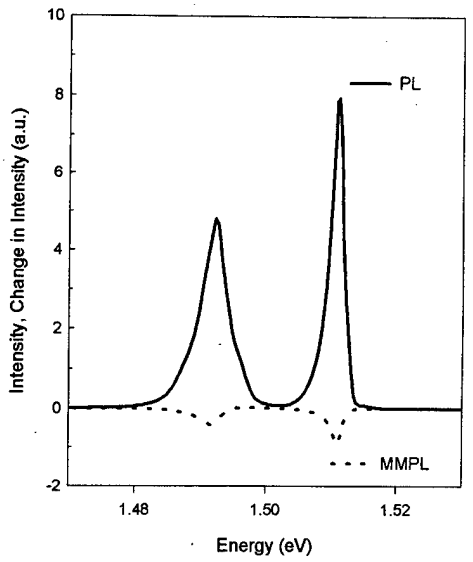


FIG. 1. PL and MMPL spectra from a GaAs sample. MMPL spectrum is the difference between PL spectra with the microwave electric fields on and off. Sign and relative magnitude of the MMPL spectra indicate that, in this case, the microwaves produce an approximately 10% decrease in PL intensity.

As alluded to above, microwave power and laser power were varied throughout the study despite the fact that the desired quantity (S) is not expected to depend on them. This was done in an effort to eliminate experimental artifacts while developing this new measurement technique, as well as to estimate the uncertainty in the values of S. The positions of the contacts on the samples was also changed for similar reasons. The motivations for, and effects of, these changes will be made clear in later sections.

III. RESULTS

The PL and MMPL spectra from a GaAs sample with a bare GaAs surface are shown in Fig. 1 in order to illustrate the effect the microwaves have on the near-gap luminescence. The same scale is used for both the PL and MMPL spectra. The MMPL spectrum is the difference, detected with the lock-in amplifier, between the PL spectrum with the microwave electric fields on and that with the microwaves off. A negative MMPL spectrum such as this one, by our convention, means that the luminescence intensity is decreased when the microwaves are switched on. The fact that the magnitude of the MMPL signal is approximately one tenth that of the PL signal indicates that the PL efficiency with the microwaves on is approximately nine tenths the efficiency with the microwaves off.

As discussed in the previous section, knowing this microwave-induced decrease in luminescent efficiency and a corresponding decrease in sample resistance allows us to calculate a surface recombination velocity. Such data are shown

TABLE I. Experimental data and calculated values of S for different microwave powers. Microwave power is shown in column one. ($P \approx 40$ mW.) Columns two and three show the changes in the two measured quantities due to the presence of the microwaves. Column four gives the calculated values for S. Uncertainty in the values is in the last digit except as noted.

Microwave power	R_{on}/R_{off}	η_{on}/η_{off}	S (cm/s)
P	0.58	0.89	$1.5e4 \pm 5e3$
$0.3P$	0.65	0.95	$6e3 \pm 2e3$
$0.1P$	0.71	0.98	$3e3$
$0.03P$	0.92	0.992	$4e3$

in Table I for the sample whose PL and MMPL spectra were depicted in Fig. 1. Several microwave powers and laser powers were used in addition to the conditions that produced the data for Fig. 1. As can be seen in Table I, the presence of microwave electric fields causes a decrease in sample resistance and a decrease in luminescent efficiency. Correspondingly, both the change in sample resistance and in efficiency become larger as the microwave power increases. As will be shown in Sec. IV the change in the sample resistance places a limit on the change in luminescent efficiency for any finite value of S. Correspondingly, any time our experiment produces a measurable difference between the two values depicted in columns 2 and 3 of Table I, we can estimate a surface recombination velocity for the luminescent layer using estimates for some bulk material parameters of the layer (i.e., the radiative lifetime and microwaves-off diffusion constant) and Eqs. (1) and (2). Changing the laser intensity by an order of magnitude has a negligible effect on the numbers shown in Table I. The values we obtain by our technique are on the order of previous estimates of low temperature SRV's.[12]

IV. DISCUSSION

A. Assumptions used to calculate S

Before explaining the details of the calculation of surface recombination velocities from the measured quantities, it is appropriate to explain and justify some of the approximations that will be made. First, we assume that the change in PL efficiency comes only from a change in the nonradiative lifetime and that the change in that lifetime comes only from a change in the diffusion constant for free carriers. This assumption is justified primarily with empirical evidence. The quenching of the PL signal caused by the microwave electric field has been shown to be associated with the presence of an exposed surface on a GaAs sample.[7] The change in PL affects the different luminescence processes equally, i.e., the MMPL signal in Fig. 1 does not change the relative amplitudes of the two peaks, meaning that any relative change between the efficiencies of the PL processes is small compared to the quenching effect of the surface. In addition, we have taken data at a variety of microwave and laser powers that give similar values of S, indicating we are probably not neglecting any large extraneous effects of the microwave electric fields. In particular the data shown in Table I converge at low microwave powers, as was the trend in all of our data. The deviations at high microwave powers could be due to a microwave-induced change in radiative rate. Since the energy dependence of radiative capture cross sections is approximately proportional to the fourth power of the energy,[13] while the change in mobility is proportional to the three halves power,[9] the effect on the radiative rates may become more important at high microwave powers. For this reason, the low-microwave-power values of S are used for discussion.

In principle the change in sample resistance could occur due to a change in mobility of a different carrier type than the carriers which determine the luminescence efficiency. That is, the D in Eq. (1) refers to the diffusion constant of minority carriers,[6] while the sample resistance would in general be determined by majority carriers.[14] For the carrier freeze-out conditions we use in this study, the free carriers that our measurements probe consist of equal numbers of electrons and holes. In the case of equal numbers of both carrier types, the resistance of the sample is expected to be determined by the more mobile carriers,[14] while the luminescence efficiency may be determined by the less mobile. However, we use the measured change in resistance to infer the change in D used in Eq. (1). This inference is possible because for carrier freeze-out conditions we expect the increase in carrier kinetic energy, and hence the relative change in mobility, to be essentially the same for both types of carriers. The increase in kinetic energy of the carriers will be proportional to the strength of the microwave electric field and the scattering length. As already stated, the dominant scattering mechanism is expected to be impurity scattering. For the carrier freeze-out conditions, the impurities consist of equal numbers (compensated) of positively charged donors and negatively charged acceptors, as well as uncompensated neutral acceptors. Due to this symmetry of the charge of the impurities, we assume that the scattering length, and hence the microwave-induced increase in kinetic energy, is the same for both electrons and holes. We assume that the only dependence of the resistance on the presence of the microwave electric fields is through the increase of carrier energy. Since the sample resistance will enter our calculations only in a ratio of the resistance with microwaves on to the resistance with microwaves off, the value of S will only depend upon the change in free carrier energy and all of our results should be independent of which carrier type we are measuring.

B. Exact solution for S

The two quantities measured in our experiment that are displayed in Table I are the ratio of the efficiencies with the microwaves on and off:

$$\frac{\eta_{on}}{\eta_{off}} = \left(1 + \frac{\tau_r}{\tau_{nr,on}}\right)^{-1} \left(1 + \frac{\tau_r}{\tau_{nr,off}}\right), \tag{3}$$

and the ratio of the sample resistances with the same conditions,

$$\frac{R_{on}}{R_{off}}=\frac{D_{off}}{D_{on}}. \tag{4}$$

The "off" and "on" subscripts refer to the absence and presence of the microwave electric fields, respectively. In Eq. (4), the Einstein relation[10] has been used to relate the diffusion constant and the resistance of the sample. As explained in Sec. IV A, any dependences of quantities other than D on the presence or absence of the microwaves will be ignored.

Using Eqs. (1), (3), and (4) it is possible to solve for the surface recombination velocity, S, in terms of the two dimensionless measured quantities η_{on}/η_{off} and R_{on}/R_{off}, or equivalently η_{on}/η_{off} and D_{on}/D_{off}. An exact solution based on Eq. (1) and the assumptions detailed in Sec. IV A is found by solving a quadratic equation for S,

$$S^2\left[\left(1-\frac{\eta_{on}}{\eta_{off}}\right)\frac{16d^4}{\pi^4 D_{off}}+\left(1-\frac{\eta_{on}}{\eta_{off}}\frac{D_{on}}{D_{off}}\right)\frac{\tau_r 4d^2}{\pi^2}\right]$$
$$+S\left[\left(1-\frac{\eta_{on}}{\eta_{off}}\right)\left(\frac{4d^3}{\pi^2}+\tau_r d D_{off}\frac{D_{on}}{D_{off}}+\frac{4d^3}{\pi^2}\frac{D_{on}}{D_{off}}\right)\right]$$
$$+\left[\left(1-\frac{\eta_{on}}{\eta_{off}}\right)D_{off}d^2\frac{D_{on}}{D_{off}}\right]=0, \tag{5}$$

and this method was used to find the values of S listed in Table I. Given that η_{on}/η_{off} is less than unity and S must be positive real, there must be a sign difference between the coefficients of the quadratic and linear terms in S in Eq. (5), and the following condition must be met:

$$\frac{\eta_{on}}{\eta_{off}}\frac{D_{on}}{D_{off}}>1, \tag{6}$$

or equivalently,

$$\frac{\eta_{on}}{\eta_{off}}>\frac{R_{on}}{R_{off}}. \tag{7}$$

Hence, for any physically meaningful (i.e., finite, real, positive) value of S, the condition (7) must be true, as it is for the data presented in Table I. Operationally this conclusion means that as the two ratios in Eq. (7) become close to each other such that their error bars overlap, S becomes immeasurably large in our experiment.

C. Approximate solutions for *S*

Although it is easy to solve Eq. (5) for any value of S, for illustrative purposes, approximate solutions will be presented that make clearer the dependence of S on the measured (modulated) quantities and on the material parameters in certain limiting cases.

The first approximate solution applies for the case of "bad" interfaces or surfaces. Quantitatively, this case applies when the nonradiative lifetime is small compared to the radiative lifetime and the following approximation is valid:

$$\eta\cong\frac{\tau_{nr}}{\tau_r}. \tag{8}$$

Note that for the nonradiative lifetime to be small compared to the radiative lifetime, and Eq. (8) to apply, both terms on the right hand side of Eq. (1) must be small compared to the radiative lifetime. Specifically, the "bad" interface approximation used here requires not only a large value of S but also a small layer thickness, or equivalently a large diffusion constant. When Eq. (8) is valid the ratio of the efficiencies is independent of the radiative lifetime.

$$\frac{\eta_{on}}{\eta_{off}}\cong\frac{\frac{4d^2}{\pi^2 D_{on}}+\frac{d}{S}}{\frac{4d^2}{\pi^2 D_{off}}+\frac{d}{S}}. \tag{9}$$

By rearranging terms one obtains the recombination velocity of a "bad" surface or interface:

$$S_{bad}\cong\frac{\pi^2 D_{off}}{4d}\frac{\left(1-\frac{\eta_{on}}{\eta_{off}}\right)}{\left(\frac{\eta_{on}}{\eta_{off}}-\frac{D_{off}}{D_{on}}\right)}. \tag{10}$$

The second condition for which an approximate solution is useful is the case of a "good" interface or surface which is defined by the following two conditions:

$$\frac{d}{S}\gg\frac{4d^2}{\pi^2 D} \tag{11}$$

and

$$\frac{d}{S}\gg\tau_r. \tag{12}$$

Using Eqs. (11) and (12) it is possible once again to write a simplified expression for the ratio of the efficiencies, Eq. (3).

$$\frac{\eta_{on}}{\eta_{off}}\cong 1-\frac{\tau_r}{\tau_{nr,on}}+\frac{\tau_r}{\tau_{nr,off}}. \tag{13}$$

Using Eq. (11) to obtain an approximate expression for the inverse of the nonradiative lifetime (not shown), the final expression for "good" interfaces becomes:

$$S_{good}=\frac{\pi}{2}\sqrt{\frac{1-\eta_{on}/\eta_{off}}{\frac{\tau_r}{D_{off}}(1-D_{off}/D_{on})}}. \tag{14}$$

A graph of the exact solution of the quadratic equation (5), and the approximate expression (14) is shown in Fig. 2 for material parameters typical of samples we studied for this work (i.e., a 1 μm thick GaAs sample). The figure shows the dependence of the surface recombination velocity on one of the measured quantities (η_{on}/η_{off}) with the other measured quantity (R_{on}/R_{off}) taken as 0.7. This value of (R_{on}/R_{off}) corresponds roughly to a microwave power of 4 mW. At low values of S, the two curves converge and Eq. (14) is valid. Equation (10) is not plotted because the required condition [Eq. (8)] is not valid for this choice of parameters. Specifically, in a 1 μm thick GaAs sample the nonradiative lifetime [Eq. (1)], which is determined by the diffusion constant at

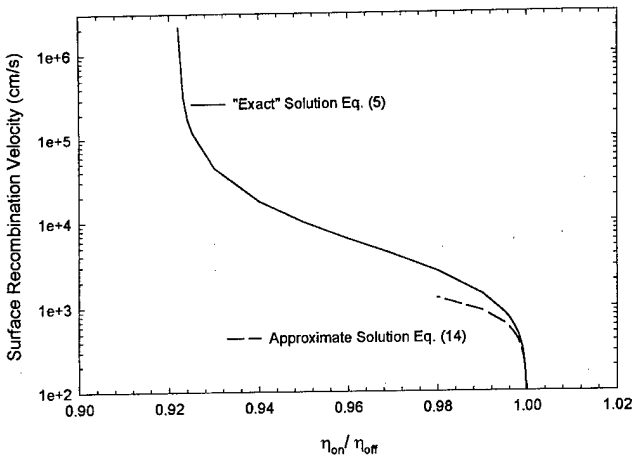


FIG. 2. Graph of the solution of the quadratic equation (5) and the approximate solution (14) for a value of R_{on}/R_{off} of 0.7. The two curves converge at small values of recombination velocity, S. Note that the error in the value of S we obtain, as determined by experimental error in the measurement of PL efficiency, becomes larger at values of S below 10^3 cm/s and above 10^5 cm/s. Range over which we can accurately measure S can be extended by changing the layer thickness and the microwave power.

large values of S, never becomes small compared to the radiative lifetime. In the GaAs system the use of Eq. (10) would require thinner layers.

D. Uncertainty in S

In Table I, the uncertainty in S, which comes from the experimental uncertainty in the two measured quantities, is $\pm 10^3$ cm/s for the low microwave powers and becomes larger at higher microwave powers. This increase in uncertainty in S comes from the fact that at high microwave powers, τ_{nr} [given by Eq. (1)] becomes insensitive to the value of S as the diffusion constant, D, becomes smaller. This uncertainty in S at low D (high microwave power) is equivalent to the problem of measuring large values of S accurately. The fundamental limit where τ_{nr} becomes insensitive to S is the same for our measurements and the technique described in Ref. 6 based on multiple layers with different thicknesses. It is also clear from Eq. (1) that one can measure higher values of S using thinner layers. For example, for a 100 Å layer it should be possible to measure values of S accurately up to 10^7 cm/s.

As mentioned at the beginning of this section there may be additional systematic errors with the technique described in this work associated with high microwave powers. The fact that we cannot increase the microwave power *ad infinitum* places a lower limit on the values of S that we can measure accurately. Whereas the technique in Ref. 6 becomes more accurate as τ_{nr} becomes insensitive to the value of D (in the limit of small D or thin layers), our technique requires measuring changes in the efficiency based on changes in D. For the GaAs system with thick layers (>1 μm) this means we probably cannot accurately measure values of S below 10^2 cm/s. In contrast, using the time-resolved, multiple layer technique, values of S as low as 1.5 cm/s have been reported.[15]

Despite the fact that in its present form the microwave modulated photoluminescence experiment is limited to low temperatures, relatively pure samples, and a slightly smaller range of measurable values of S, than the competing technique discussed, the ability to measure surface recombination velocity with only a single sample eliminates a potentially large source of error from the measurement, the variation of surfaces from sample to sample. It should be noted for clarity that microwave modulated photoluminescence can be used to measure recombination velocities in a variety of systems other than the one reported on here, including buried interfaces. The only restrictions on the layer to be measured are that it must have negligible conductivity without photoexcitation compared to with photoexcitation, it must have measurable PL, the layer must be contacted for electrical measurements, and the recombination velocity must be between the two limits discussed in the previous paragraphs.

V. CONCLUSIONS

We have demonstrated a new technique for measuring interfacial recombination velocities using changes in luminescence and conductivity caused by microwave heating of photogenerated carriers. We have obtained values for surface recombination velocity in reasonable agreement with previous low-temperature results.[12] The samples used in this work, for the sake of simplicity, all had a bare surface with relatively high recombination velocity, but the technique is not necessarily limited to that class of samples. Using the methods described here it is, in principle, possible to measure the recombination velocity of a surface or interface that terminates a luminescent layer in other sample geometries, including buried interfaces.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Contract No. N000149410941. The authors wish to thank D. S. Katzer at the Naval Research Laboratory, J. M. Olson at the National Renewable Energy Laboratory, and W. Duncan at Texas Instruments for supplying the samples, and W. Wingert at the University of Utah for help with sample preparation.

[1]I. Viohl, Ph.D. thesis, University of Utah, 1991.
[2]M. C. DeLong, I. Viohl, W. D. Ohlsen, P. C. Taylor, and J. M. Olson, Phys. Rev. B **43**, 1510 (1991).
[3]C. E. Inglefield, M. C. DeLong, P. C. Taylor, and W. A. Harrison, Phys. Rev. B **56**, 12434 (1997).
[4]M. C. DeLong, W. D. Ohlsen, I. Viohl, X. Yin, P. C. Taylor, D. Sengupta, G. E. Stillman, J. M. Olson, and W. A. Harrison, Phys. Rev. B **48**, 5157 (1993); F. P. Wang, B. Monemar, and M. Ahlström, *ibid.* **39**, 11195 (1989); H. Weman, M. Godlewski, and B. Monemar, *ibid.* **38**, 12525 (1988).
[5]M. C. DeLong, W. D. Ohlsen, I. Viohl, X. Yin, P. C. Taylor, D. Sengupta, G. E. Stillman, J. M. Olson, and W. A. Harrison, Phys. Rev. B **48**, 5157 (1993).

[6]*Minority Carriers in III–V Semiconductors: Physics and Applications*, edited by R. K. Ahrenkiel and M. S. Lundstrom (Academic, San Diego, CA, 1993), Chap. 2.
[7]C. E. Inglefield, M. C. DeLong, P. C. Taylor, J. F. Geisz, and J. M. Olson, J. Vac. Sci. Technol. B **15**, 1201 (1997).
[8]If the microwave induced change in the nonradiative lifetime is large enough, a decrease in the total number of free carriers present results in a decrease in conductivity. In our calculations, this additional mechanism for changing sample resistance is taken into account in an iterative fashion.
[9]D. C. Look, *Electrical Characterization of GaAs Materials and Devices* (Wiley, Chichester, 1989), pp. 76–8.
[10]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College Press, Philadelphia, PA, 1976), p. 602.
[11]B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer, Berlin, 1984).
[12]S. Myhajlenko, D. B. Davito, R. A. Puechner, and J. L. Edwards, Jr., Philos. Mag. B **69**, 553 (1994).
[13]I. Viohl, Ph.D. thesis, University of Utah, 1991, p. 123.
[14]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981), p. 31.
[15]J. M. Olson, R. K. Ahrenkiel, D. J. Dunlavy, Brian Keyes, and A. E. Kibbler, Appl. Phys. Lett. **55**, 1208 (1989).

Native extended defects in $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures

B. Müller, S. Heun, R. Lantier, S. Rubini, J. J. Paggel,[a] L. Sorba,[b] A. Bonanni, M. Lazzarino, B. Bonanni, and A. Franciosi[c]
Laboratorio Nazionale TASC-INFM, Area di Ricerca, Padriciano 99, I-34012 Trieste, Italy

E. Napolitani, F. Romanato,[d] and A. Drigo
INFM and Dipartimento di Fisica, Universita' di Padova, I-35131 Padova, Italy

J.-M. Bonard and J.-D. Ganière
Institut de Micro- et Optoélectronique, Département de Physique, Ecole Polytechnique Fédérale, CH-1015 Lausanne, Switzerland

L. Lazzarini and G. Salviati
Istituto MASPEC del CNR, Parma, Italy

(Received 21 January 1998; accepted 11 May 1998)

Lattice-matched $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterojunctions can be fabricated by molecular beam epitaxy on GaAs(001) 2×4 surfaces in a wide range of compositions provided that a suitable strain relaxation profile is achieved within the ternary III–V buffer layer. We focus here on the structural properties of the resulting II–VI/III–V heterostructures and discuss the distribution of native defects, including misfit and threading dislocations, stacking faults, and surface corrugations. © *1998 American Vacuum Society.* [S0734-211X(98)05304-9]

I. INTRODUCTION

Heterovalent heterostructures, i.e., junctions comprised of semiconductors belonging to different families (e.g., II–VI/III–V or III–V/IV) have been attracting increasing attention in recent years for both basic and applied reasons.[1,2] On the fundamental side, the physics of these systems is in principle qualitatively different from that of the more conventional isovalent heterostructures.[1] Ideally abrupt heterovalent heterojunctions with polar orientation would be charged, and unstable.[3,4] Elimination of the resulting long-range electrostatic field may occur through the formation of point defects, which, in turn, may modify the band alignment. As far as applications are concerned, most materials currently being considered for optoelectronic devices in the short-wavelength visible range (e.g., the wide-gap II–VI materials or III–V nitrides) are grown on heterovalent substrates, and this emphasizes the importance of understanding and controlling the properties of heterovalent heterostructures.

Strain further complicates the issue, since the heterovalent substrates are seldom lattice matched to the active epilayers, and strain affects *per se* the electronic structure of the heterojunctions,[1,5] and may lead to the formation of extended as well as point defects.[1,5] Substantial progress towards the development of lattice-matched heterovalent heterostructures has been made in the II–VI/III–V area, in parallel with the recent demonstration of blue–green coherent emitters.[6] Because of the lack of II–VI substrates of suitable quality and doping, most such lasers have been fabricated by molecular beam epitaxy (MBE) on GaAs wafers. ZnSe, with a lattice parameter $a_0=5.6687$ Å,[7] exhibits a 0.27% in-plane mismatch with GaAs, but is lattice matched to $In_{0.038}Ga_{0.962}As$,[8] so several authors have attempted to achieve lattice-matched interfaces by incorporating $In_{0.04}Ga_{0.96}As$ buffer layers in blue–green lasers.[9,10]

We propose $Zn_{1-y}Cd_ySe/In_{1-x}Ga_xAs$ heterojunctions as a new prototype family of II–VI/III–V heterovalent heterojunctions that can be fabricated by MBE in lattice-matched form in a wide range of compositions on GaAs wafers, provided that a suitable strain-relaxation profile is achieved in the ternary III–V layers. We review here the conditions under which lattice matching is achieved, and the nature and spatial distribution of the dominant extended defects which are found in the structure.

An additional motivation of our work is the goal of eliminating strain from the active layer of current II–VI based blue–green lasers. $Zn_{1-y}Cd_ySe$ alloys are widely utilized as active layers in laser diodes grown on GaAs, and the large resulting compressive strain appears to promote device degradation.[11,12] Since high-quality $In_xGa_{1-x}As$ wafers with $x\geq0.02$ are not available because of homogeneity problems in high-x bulk crystals, fabrication of $In_xGa_{1-x}As$ buffer layers with optimized concentration and strain relaxation profile on GaAs seems at present the only viable option. We present here results for $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures of interest for both blue ($y=0$ and 0.15) and blue–green ($y\sim0.25$) emitters.

II. EXPERIMENTAL DETAILS

All heterostructures were grown by solid-source MBE in a system which includes interconnected chambers for the growth of II–VI and III–V materials and an analysis chamber with monochromatic x-ray photoemission spectroscopy (XPS) capabilities. After thermal removal of the native oxide

[a]Present address: Dept. of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801.
[b]Also with Istituto ICMAT del CNR, Monterotondo, I-00016 Rome, Italy.
[c]Also with Dipartimento di Fisica, Universita' di Trieste, I-34127 Trieste, Italy and Dept. of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455.
[d]Present address: Laboratorio Nazionale TASC-INFMe Sincrotrone Trieste S.C.p.A., Area di Ricerca Padriciano 99, I-34012 Trieste, Italy.

from semi-insulating GaAs(001) wafers, ~0.5-μm-thick GaAs buffer layers were initially grown at 600 °C. $In_xGa_{1-x}As$ epilayers 1–2 μm thick were subsequently grown at 490–500 °C with III/V beam pressure ratios (BPRs) in the 1/15–1/30 range, as determined from an ion gauge positioned at the sample location. Typical growth rates were between 1 and 1.3 μm/h. The indium content in the alloy was calibrated *in situ* by means of XPS, and *ex situ* by x-ray diffraction (XRD), photoluminescence (PL) spectroscopy, or Rutherford backscattering spectrometry (RBS).[13]

When required, graded composition $In_xGa_{1-x}As$ layers in which x was varied superlinearly with a parabolic composition profile were designed, as described in Sec. III, and fabricated by gradually increasing the indium cell temperature during growth. The composition profile of the graded layers was tested *a posteriori* by RBS and secondary ion mass spectrometry.

On the resulting $In_xGa_{1-x}As$ surface, $Zn_{1-y}Cd_ySe$ overlayers were deposited at 250–290 °C with a II/VI BPR of 1 or 0.4. Under the growth conditions employed the layers exhibited a $c(2\times2)$ (BPR=1) or (2×1) (BPR=0.4) surface reconstruction, as determined by reflection high energy electron diffraction (RHEED). The cadmium concentration in the alloy was also calibrated *in situ* by XPS and *ex situ* by RBS.

The extended defect density for each sample was probed using plan-view and cross-sectional transmission electron microscopy (TEM) micrographs. Thin foil samples were obtained by ion milling the structures at room temperature using 5 keV Ar^+ ions at an angle of incidence of 4° until perforation. Plan-view images obtained using a Philips EM430ST TEM operated at 300 keV allowed us to sample the overall heterostructure in sections as thick as 2 μm. Observations in thin regions of the plan-view thin foils allowed us to focus on the structural quality of the II–VI epilayer alone, while thicker regions yielded information on the whole sequence of epilayers down to the III–V buffer/substrate interface. Cross-sectional micrographs were recorded using either a JEOL 2000 FX TEM operated at 200 keV or a Philips EM430ST TEM operated at 300 keV.

The residual strain in the epilayers was determined either through XRD using a double-crystal diffractometer[14] or from the measured channeling directions using ^{4}He beams with an energy of 2.0 or 4.0 MeV.[13,15,16] Throughout this article the measured elastic deformation will be quantified in terms of the diagonal components of the strain tensor $\epsilon_\parallel$ and $\epsilon_\perp$.[17]

The surface morphology of the samples was examined using Nomarski optical microscopy and atomic force microscopy (AFM). The AFM images were obtained using a commercial instrument operated in contact mode with Si_3N_4 tips on cantilevers with a spring constant of 0.06–0.12 N/m, and typical scan rates in the 0.5–2.0 Hz range.

III. RESULTS AND DISCUSSION

A. Dislocations

In Table I we summarize the nominal parameters of $In_xGa_{1-x}As$ and $Zn_{1-y}Cd_ySe$ bulk alloys that would be lattice matched at room temperature for $y=0$, 0.15, and 0.25. The lattice parameters of the $Zn_{1-y}Cd_ySe$ alloys were obtained by linear interpolation between the lattice parameter of ZnSe (Ref. 18) and that measured in epitaxial, relaxed CdSe epilayers grown with a zinc blende structure on GaAs.[19] The lattice parameters of the $In_xGa_{1-x}As$ alloys were also obtained through interpolation from literature values.[20] The parallel decrease in the band gap with an increasing lattice parameter shown in Table I defines a series of potentially lattice-matched heterojunctions in which the band gap difference remains virtually unchanged while becoming an increasing fraction of the band gap of each semiconductor constituent. This should produce a substantial modulation of confinement effects in the heterostructure series.

TABLE I. Nominal parameters of $In_xGa_{1-x}As$ and $Zn_{1-y}Cd_ySe$ bulk alloys that would be lattice matched at room temperature for $y=0$, 0.15, and 0.25. Column 1: Lattice parameter; column 2: Cd content in the II–VI alloy; column 3: direct band gap of the II–VI alloy: column 4: In content in the III–V alloy; column 5: direct band gap of the III–V alloy; column 6: band gap difference between lattice-matched II–VI and III–V alloys.

	$Zn_{1-y}Cd_ySe$		$In_xGa_{1-y}As$		
a_0 (Å)	y	E_G (eV)	x	E_G (eV)	ΔE_G (eV)
5.669	0	2.70	0.04	1.38	1.32
5.730	0.15	2.55	0.19	1.22	1.33
5.771	0.25	2.44	0.29	1.11	1.33

Implementation of such lattice-matched heterostructures on GaAs wafers requires that controlled relaxation is obtained within $In_xGa_{1-x}As$ epilayers of a practical thickness (1–2 μm). For sufficiently low values of x this can be achieved using homogeneous composition layers. For example, in Fig. 1(a) we show a cross-sectional TEM micrograph of a $ZnSe/In_xGa_{1-x}As/GaAs$ heterostructure incorporating a 2-μm-thick ternary buffer layer with homogeneous composition $x=0.05$. At the interface between the ternary III–V layer and GaAs, an important dislocation network can be seen. However, these defects are confined near the interface, and no threading defects are observed within the $In_xGa_{1-x}As$ ternary layer or the ZnSe overlayer. From plan-view studies we estimate an upper limit of less than 2×10^4 cm^{-2} for the threading dislocation (TD) density in the II–VI overlayer.

When the mismatch between the II–VI/III–V heterostructure and the GaAs substrate becomes too high, homogeneous composition $In_xGa_{1-x}As$ layers become less suitable for the purpose. In Fig. 1(b) we show a cross-sectional TEM micrograph of a $Zn_{0.85}Cd_{0.15}Se/In_{0.19}Ga_{0.81}As/GaAs$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick, homogeneous composition $In_{0.19}Ga_{0.81}As$ layer. Although the vast majority of the misfit dislocations (MDs) are at the $In_xGa_{1-x}As/GaAs$ interface, the larger dislocation interaction, as compared to Fig. 1(a), is seen to be accompanied by the emergence of a number of TDs, which affect both the bulk of the ternary III–V layer and the II–VI over-

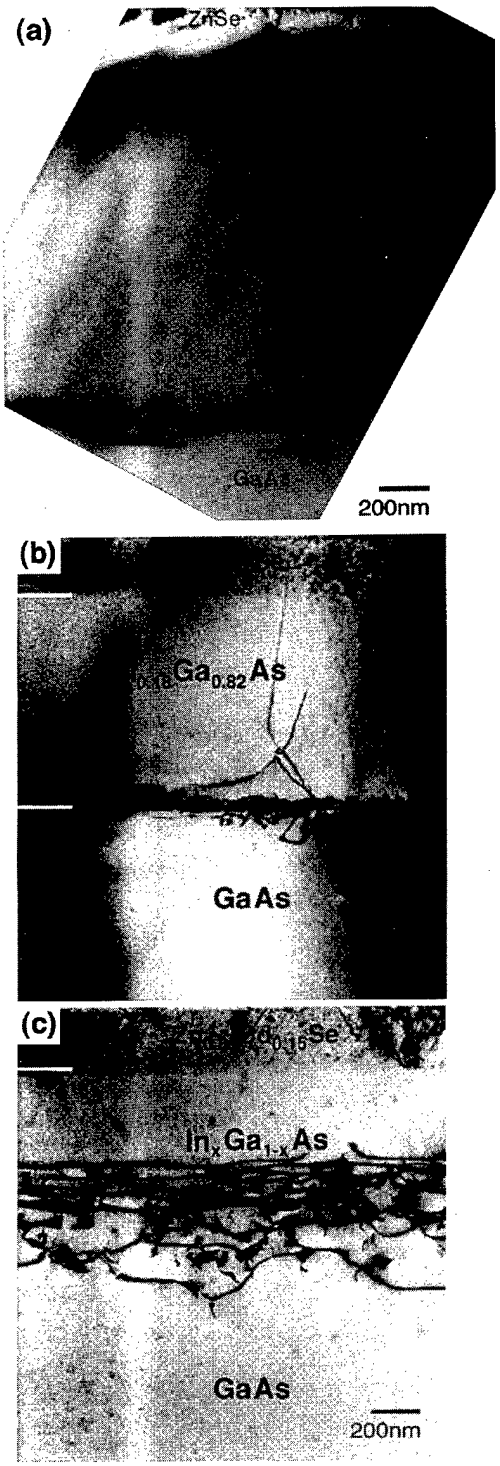


FIG. 1. (a) TEM micrograph of a $ZnSe/In_{0.05}Ga_{0.95}As/GaAs(001)$ heterostructure which includes a 100-nm-thick ZnSe epilayer and a 2-μm-thick, homogeneous composition $In_{0.05}Ga_{0.95}As$ layer. The cross-sectional image was obtained in two-beam diffraction conditions near the $[1\bar{1}0]$ zone axis. (b) Cross-sectional TEM micrograph of a $Zn_{0.85}Cd_{0.15}Se/In_{0.19}Ga_{0.81}As/GaAs$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick, homogeneous composition $In_{0.19}Ga_{0.81}As$ layer. (c) Cross-sectional TEM micrograph of a $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As/GaAs$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick, graded composition $In_xGa_{1-x}As$ layer in which the In concentration exhibited a parabolic, superlinear profile from $x=0$ at the bottom to $x_0=0.23$ at the top.

layer. Dislocation interaction is known to enhance the nucleation of additional dislocations since the already-present threading segments are prevented from gliding further.[21]

To reduce the number of TDs and to avoid the three-dimensional (3D) growth effects which hinder fabrication of high-quality $In_xGa_{1-x}As$ layers with higher In concentrations on GaAs,[16,22,23] we employed $In_xGa_{1-x}As$ epilayers in which the In concentration was varied gradually from $x=0$ at the interface with the substrate to the desired $x=x_0$ at the surface, i.e., at the II–VI/III–V interface. As discussed elsewhere,[15,24] we selected a parabolic, superlinear concentration profile:

$$x(t)=x_0\left[1-\left(1-\frac{t}{T}\right)^2\right], \tag{1}$$

where t it the distance from the GaAs substrate and T is the total thickness of the $In_xGa_{1-x}As$ layer. The profile in Eq. (1) was selected for its simplicity, relatively flat behavior in the near-surface region, which makes it less sensitive than other profiles to variations in composition and layer thickness, and for its expected property of confining MDs away from the surface while minimizing dislocation interaction.[24]

The corresponding residual strain at the film surface in the direction parallel to the interface $\epsilon_\parallel$ (Ref. 17) can be calculated from the following expression obtained through a semiempirical model of strain relaxation:[24,25]

$$\epsilon_\parallel^5=\left(\frac{9K}{4}\right)^2\frac{x_0}{T^2}\left(\frac{a_0(\mathrm{InAs})}{a_0(\mathrm{GaAs})}-1\right), \tag{2}$$

where $K=0.0037\pm0.0007$ nm is an empirical constant, and a_0 (InAs) and a_0 (GaAs) are the equilibrium (unstrained) lattice parameters of the binary parent compounds. Based on the predictions of Eq. (2), the In concentration x_0 at the surface of the graded layer must exceed the values of x in Table I to compensate for the partial character of the strain relaxation within any layer of finite-size thickness. For example, in order to obtain a surface lattice parameter[17] that matches that of a bulk $Zn_{0.85}Cd_{0.15}Se$ alloy using an $In_xGa_{1-x}As$ graded layer with $T=1100$ nm, the surface composition should be $x_0=0.245$.

In Fig. 1(c) we show a cross-sectional TEM micrograph of a $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As/GaAs$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick, graded composition $In_xGa_{1-x}As$ layer in which the In concentration exhibited a parabolic, superlinear profile of Eq. (1) from $x=0$ at the bottom to $x_0=0.23$ at the top. MDs are seen to be distributed throughout a region approximately 500 nm wide, gradually relaxing most of the misfit within this region, and leaving the remaining portion of the buffer unaffected, in agreement with theoretical predictions.[24–26] Within our experimental sensitivity, no TDs are observed.

The $Zn_{1-y}Cd_ySeSe/In_xGa_{1-x}As$ heterostructures discussed in what follows will include homogeneous composition ternary III–V layers only for $y=0$, while samples with $y=0.15$ and 0.25 will always be implemented using graded composition $In_xGa_{1-x}As$ layers with target structural parameters derived from Eqs. (1) and (2).

TABLE II. Experimental structural and compositional parameters of $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures grown by MBE on GaAs(001) substrates. The data in the first row ($y=0$) were obtained from XRD. The data in the second and third rows were obtained from RBS and channeling studies. Column 1: Cd content in the II–VI overlayer; column 2: in-plane strain within the II–VI overlayer; column 3: In content in the near-surface region of the III–V layer; column 4: in-plane residual strain in the near-surface region of the III–V layer; column 5: type and thickness of the $In_xGa_{1-x}As$ layer employed. The *h* and *g* labels denote homogeneous and graded composition layers, respectively.

$Zn_{1-y}Cd_ySe$		$In_xGa_{1-x}As$		
y	$\epsilon_\parallel$ (%)	x_0	$\epsilon_\parallel$ (%)	Type/thickness (nm)
0	-0.014 ± 0.007	0.051 ± 0.002	-0.097 ± 0.016	h/2000
0.150 ± 0.04	-0.29 ± 0.04	0.214 ± 0.004	-0.40 ± 0.04	g/920
0.249 ± 0.002	-0.14 ± 0.04	0.340 ± 0.004	-0.48 ± 0.04	g/915

In Fig. 1 no MDs are observed within the experimental sensitivity at any of the II–VI/III–V interfaces, indicating that the II–VI overlayers grow pseudomorphically on the III–V layers. To quantify how close a lattice-match was routinely achieved, we list in Table II the structural parameters measured for three $Zn_{1-y}Cd_ySeSe/In_xGa_{1-x}As$ heterostructures with $y=0$, 0.15, and 0.25. To obtain structural information on both the II–VI overlayer and the near-surface region of the graded composition III–V epilayer within the *same* heterostructure, we followed a two-step growth process. First, a graded III–V ternary epilayer was grown on a suitable GaAs wafer and protected with an As cap layer. The wafer was then removed from the MBE system and cleaved into different sections. Some sections were used in RBS studies of composition and strain. Others were introduced again in the MBE system and used as substrates for $Zn_{1-y}Cd_ySe$ growth after thermal removal of the As cap layer. In view of the RBS sampling depth and the relatively flat profile in the near-surface region [see Eq. (1)], the RBS-derived values of x will be taken as experimental determinations of the surface composition x_0.

The structural parameters of the $In_xGa_{1-x}As$ layers in Table II indicate that a substantial amount of residual strain is present in the III–V layers, as expected because of the finite thickness of any practical ternary buffer layer. For the graded composition layers, the experimental values of $\epsilon_\parallel$ (-0.40% and -0.48%) are very similar to the theoretical predictions (-0.42% and -0.46%, respectively) obtained from Eq. (2) using the experimental values of x_0 and T. The implication is that the semi-empirical model[15,24,25] does remarkably well in predicting the strain relaxation in the graded layers. The target values of x_0 and T (i.e., those required for lattice match to the II–VI overlayer) are achieved within 5%–10% through the growth procedure employed, with a trend toward lower thicknesses (by up to 10% of the target T) and lower In concentrations (by up 2%–5% of the target x_0) as compared to the optimum design values.

The structural parameters of the II–VI overlayers in Table II indicate that they grow pseudomorphically on the III–V epilayers, so that the strain present largely reflects the deviations of the ternary III–V parameters from the design values. For the $ZnSe/In_xGa_{1-x}As$ heterostructure, the low strain achieved ($\epsilon_\parallel=-0.014\%\pm0.007\%$) indicates that the excess In concentration over the value in Table I (5% as opposed to 4%) compensates almost exactly for the partial character of the strain relaxation within the III–V layer. The resulting expected critical thickness for the II–VI overlayer exceeds 300 μm,[14] suggesting that for all practical purposes lattice matching has been achieved.

The strain in the II–VI overlayer is somewhat larger for $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As$ and $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As$ heterostructures, reflecting the larger deviations of the ternary III–V parameters from the design values in the graded-composition layers. To put these results in perspective, two points should be noted: (i) Table II was compiled using a subset of samples for which RBS studies were performed on both the III–V and the II–VI layers, and are not necessarily those for which the best lattice matching was achieved. For example, on a total set of six $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures with $y=0.15-0.25$, the minimum strain encountered in the II–VI overlayer was -0.06% and the maximum -0.29%. (ii) The lattice matching achieved should be assessed vis-à-vis the applications envisioned. For example, the mismatch between $Zn_{0.75}Cd_{0.25}Se$ and GaAs is 2.0%, so that a -2.0% strain exists in the active layer of any corresponding blue–green laser fabricated on GaAs. If the same laser were to be fabricated on the corresponding graded layer in Table II, the strain would be decreased by a factor of 14.3.

In the absence of MDs and TDs in the II–VI epilayers, the main type of extended defect that remains is native stacking faults on {111} planes. Several such defects are clearly observed in Fig. 1, giving rise to the characteristic type of V-shaped contrast in the cross-sectional image. Stacking faults are known to form at the II–VI/III–V interface and to propagate upward during growth along {111} fault planes. Although some isolated Frank stacking faults exist on $(1\bar{1}1)$ and $(\bar{1}11)$ fault planes, we have shown[27,28] that the majority of such defects are Shockley stacking fault pairs on (111) and $(\bar{1}\bar{1}1)$ fault planes, bounded by partial dislocations exhibiting Burgers vectors $(1/6)[\bar{1}2\bar{1}]$ and $(1/6)[\bar{2}11]$ for the former, and $(1/6)[\bar{2}1\bar{1}]$ and $(1/6)[\bar{1}21]$ for the latter.

A discussion of the stacking fault issue, together with a review of the extensive recent literature on this subject, can be found elsewhere.[27,28] Here we remind the reader that in view of the presence of the partial dislocations bounding the

stacking faults, the II–VI overlayers in Fig. 1 cannot be considered, strictly speaking, dislocation free. However, such stacking faults do not contribute significantly to strain relaxation, and in any case their concentration can be reduced below 10^4 cm^{-2} by controlling the interface nucleation stage, either by acting on the Zn/Se BPR (Se-rich growth conditions yield lower stacking fault densities[27,28]) or by predosing the surface with submonolayer coverages of Zn.[29,30]

B. Surface properties

Figure 2 shows the surface morphology of the three types of $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures examined in this study as sampled by Nomarski optical microscopy. Clearly visible in all three cases is a regular pattern of parallel lines on the surface with a separation of approximately 1 μm. These striations run throughout the sample along the [1$\bar{1}$0] direction. Weaker lines are also visible in the [110] perpendicular surface direction, with lower density and therefore larger spacing yielding a characteristic cross-hatched pattern.

Nomarski optical micrographs of the corresponding III–V buffer layers can be found elsewhere[14–16] and are not reproduced here for sake of brevity. The similarity of the cross-hatched patterns from the II–VI overlayers and the III–V substrates, and the fact that the striations are absent from the surface of pseudomorphic ZnSe layers grown on GaAs[14] suggest that the surface morphology of the $Zn_{1-y}Cd_ySe$ overlayers replicates that of the underlying $In_xGa_{1-x}As$ epilayers.

Surface corrugations arranged to give a cross-hatched pattern have been observed for both $In_xGa_{1-x}As$ epilayers[31–33] as well for as $Si_{1-x}Ge_x$ epilayers[21,34] during growth on lattice-mismatched substrates when the critical thickness was exceeded. In general, their origin has been phenomenologically related to the misfit dislocation network at the epilayer–substrate interface,[31] based on their orientation, and on the fact that they were not observed in layers which do not contain substantial numbers of glissile misfit dislocations.[31]

The similar orientation of the surface cross-hatched pattern and the MD dislocation network near the $In_xGa_{1-x}As/GaAs$ is illustrated by a comparison of the results of Fig. 2 with selected plan-view TEM micrographs in Fig. 3. In Fig. 3(a) we show the MD network observed in a $ZnSe/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a 2-μm-thick $In_xGa_{1-x}As$ layer with homogeneous composition $x=0.043$. Figure 3(b) depicts the corresponding MD network in a $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with graded composition profile and surface composition $x_0=0.33$.

In both micrographs, a relatively ordered, asymmetric network of misfit dislocations is observed. Most of the MDs are found to be mixed dislocations with Burgers vectors of the $a/2\langle 110\rangle$ type, at 60° to the dislocation line.[15,35] The dislocation lines appear as straight lines in the micrographs be-

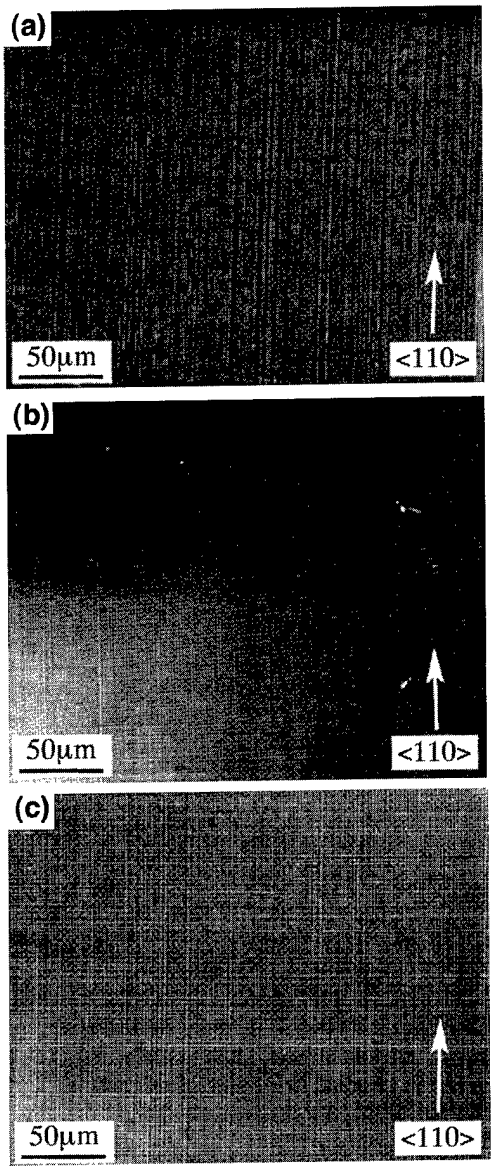


FIG. 2. (a) Surface morphology of a $ZnSe/In_xGa_{1-x}As(001)$ heterostructure grown on GaAs incorporating a 300-nm-thick II–VI overlayer and a 2-μm-thick $In_xGa_{1-x}As$ layer with homogeneous composition $x=0.043$, as measured with a Nomarski optical microscope. (b) Surface morphology of a $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.23$. (c) Surface morphology of a $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.33$.

cause they lie at the intersection of {111} slip planes with the (001) heterointerface and are in [110] and [1$\bar{1}$0] directions. The apparent asymmetry in the contrast associated with the network stems from the fact that along one of the two orthogonal $\langle 110\rangle$ directions, the dislocations are more uniformly distributed, while in the other there is a marked tendency for periodic bunching. This was attributed[15,35] to the different velocities of group III and group V dislocations in the zinc blende structure.[36]

The difference in the overall number of MD density be-

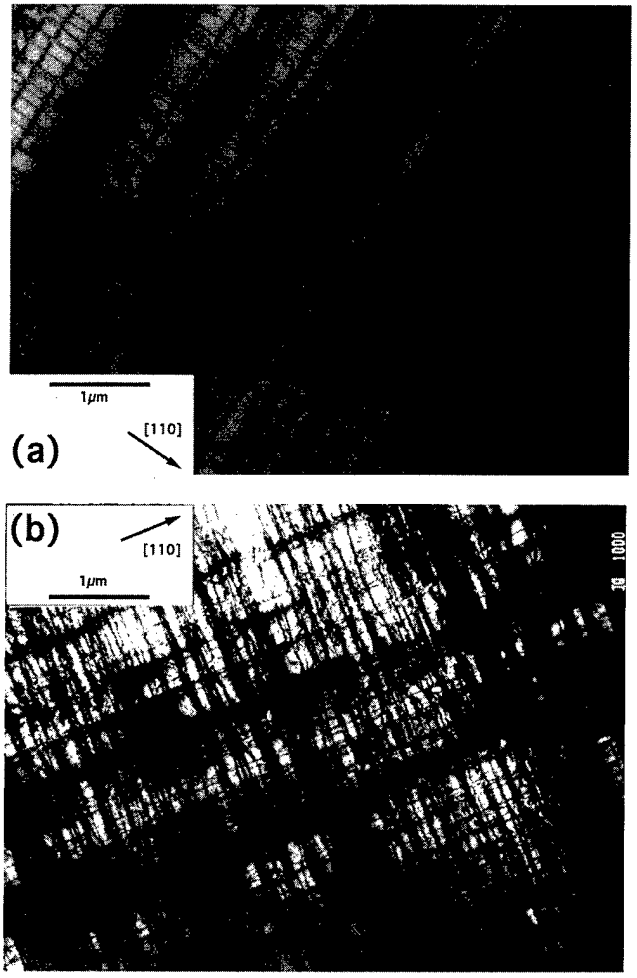


FIG. 3. TEM plan-view micrographs of $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures grown on GaAs, emphasizing the MD dislocation network near the $In_xGa_{1-x}As/GaAs$ interface. (a) $ZnSe/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a 2-μm-thick $In_xGa_{1-x}As$ layer with homogeneous composition $x=0.043$. (b) $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.33$.

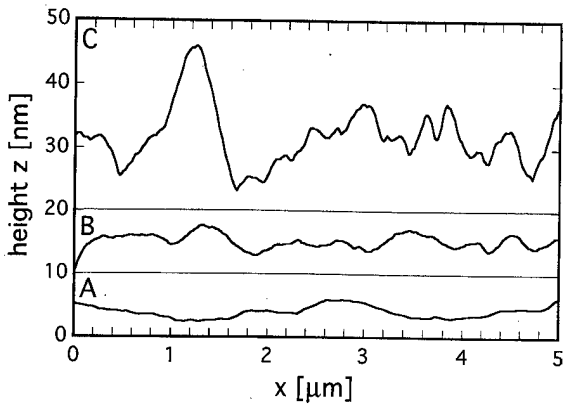


FIG. 4. AFM line scans for the type of samples depicted in Figs. 1 and 2, recorded in the direction perpendicular to the major surface corrugations, i.e., the [110] direction. (a) $ZnSe/In_xGa_{1-x}As(001)$ heterostructure grown on GaAs incorporating a 300-nm-thick II–VI overlayer and a 2-μm-thick $In_xGa_{1-x}As$ layer with homogeneous composition $x=0.043$. (b) $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.23$. (c) $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.33$.

tween Figs. 3(a) and 3(b) is simply a manifestation of the different lattice mismatch with the GaAs wafer, which is 0.27% for the sample in Fig. 3(a) and 2.0% for the sample in Fig. 3(b), yielding a 7.4-fold increase in the areal dislocation density. We emphasize, however, that while in the former case all of the dislocations are at the interface between the homogeneous composition ternary layer and the GaAs substrate [see the corresponding cross-sectional image in Fig. 1(a)], in the latter the larger number of MDs are distributed within a ~500-nm-thick region of the ternary layer due to the graded concentration profile.

A more quantitative analysis of the surface corrugations for the three heterostructures in Fig. 2 was performed by AFM. In Fig. 4 we show line scans recorded in the direction perpendicular to the main corrugations, i.e., along the [110] direction. The line profile in Fig. 4(a) was obtained from a $ZnSe/In_xGa_{1-x}As(001)$ heterostructure grown on GaAs incorporating a 300-nm-thick II–VI overlayer and a 2-μm-thick $In_xGa_{1-x}As$ layer with homogeneous composition $x=0.043$. The amplitude of the corrugations is of the order of 3 nm, and the period larger than 1 μm. A more precise determination of the period as an average over several line scans from different regions of the same sample would yield 1.4 ± 0.4 μm.[14] The measured root mean square (rms) roughness of the surface is of the order of 1 nm.

The line scan in Fig. 4(b) was recorded from a $Zn_{0.85}Cd_{0.15}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.23$. The amplitude of the corrugations is still of the order of 4–5 nm, but the period decreases to about 0.5 μm. Correspondingly, the rms roughness of the surface is 1.7 ± 0.4 nm.[15] The line profile in Fig. 4(c) was obtained from a $Zn_{0.75}Cd_{0.25}Se/In_xGa_{1-x}As(001)$ heterostructure incorporating a 300-nm-thick II–VI overlayer and a ~1-μm-thick $In_xGa_{1-x}As$ layer with a graded composition profile and surface composition $x_0=0.33$. The amplitude of the corrugations is of the order of 10 nm, and the period 0.3–0.5 μm. Correspondingly, the rms roughness of the surface is of the order of 5 nm.[16] Much higher values of surface roughness are observed for heterostructures incorporating homogeneous as opposed to graded composition $In_xGa_{1-x}As$ layers, despite the lower value of the In concentration at the surface of the ternary III–V layer.[14–16]

C. Relation between dislocations and surface morphology

In general, the formation of a MD network is explained with one of two types of mechanisms. One mechanism[35] assumes that a number of TDs derive from an imperfect sub-

strate or from surface inhomogeneities of the substrate. In the presence of a misfit between the epilayer and substrate the TD lines are bent by the stress field within the film. TD segments confined at the interface would generate the necessary MDs. This is an efficient mechanism of strain relief which does not require the nucleation of new dislocations. However, since the $In_xGa_{1-x}As$ layers in the heterostructures under consideration contain a negligible TD density, we suggest that the second type of mechanism is dominant in our case.

In this alternate picture,[37,38] dislocation loops (DLs) starting at the surface of the growing film may expand by gliding along a {111} plane and provide the MD as they reach the interface, together with a pair of TDs which are annihilated upon reaching the edge of the epilayer. This implies two consequences: (i) Every MD at the interface would correspond to a monoatomic slip step on the epilayer surface. The direction of the step (arbitrary labeled up or down) would depend on the glide plane ({111} or $\{\bar{1}11\}$) and thus on the Burgers vector of the dislocation line ($a/2\langle 101\rangle$ or $a/2\langle 01\bar{1}\rangle$); (ii) dislocations with $a/2\langle 101\rangle$-type Burgers vectors contain a vertical component which produces a small tilt in the epilayer relative to the substrate. The existence of a small epilayer tilt is well established in literature,[39] and is especially important for $In_xGa_{1-x}As$ layers grown on misoriented substrates.[40] Bauer and co-workers observed both slip steps correlated to interface dislocations and a finite epilayer tilt in a recent scanning tunneling microscopy study of strained EuTe layers on PbTe.[41]

With such considerations in mind, we can explore a possible quantitative correlation between dislocation slip steps and surface corrugations. The average spacing between dislocation lines in our $In_{0.05}Ga_{0.95}As$ buffers was of the order of 100 nm [112 ± 7 nm for the sample in Fig. 3(a)]. Our AFM measurements yielded an average spacing of the major surface corrugations of approximately 1.5 μm [1.38 ± 0.4 μm for the sample in Fig. 4(a)]. This gives about 15 dislocations and slip steps per crossed-hatch period. Assuming a series of 7–8 "up" steps followed by 7–8 "down" steps per period, and a monoatomic step height of 0.28 nm, we estimate a height modulation of 2.1 nm, i.e., of the same order of magnitude as that observed in Fig. 4(a).[14]

The presence of series of up steps and down steps rather than a random distribution is not predicted *a priori* by the DL mechanism. None of the {111} gliding planes is *a priori* preferred so that MDs with $a/2\langle 101\rangle$ or $a/2\langle 01\bar{1}\rangle$ Burgers vectors, and hence the up and down steps on the surface, should be distributed randomly. DLs, however, are not the only MD nucleation sources once the dislocation density becomes high enough for effective interaction between MD lines to take place.

For example, Hagen and Strunk[42] proposed that the repulsive interaction between two dislocations with the same Burgers vector and crossed dislocation lines along the [011] and $[01\bar{1}]$ directions would generate a right angle and a rounded right angle segment. The latter can act as a starting point for a new MD line with the *same* Burgers vector.[43] Following the glide along a {111} plane, a slip step at the surface is generated, but since the Burgers vector remains the same, the direction of the steps (up or down) also remains the same. So the original random distribution of MDs originating from the DL mechanism, followed by MD multiplication to generate bunches of MDs with the same Burgers vectors would account for the observed surface morphology through a purely plastic mechanism.

The above picture readily explains the smoother surface morphology of heterostructures incorporating a graded as opposed to a homogeneous composition $In_xGa_{1-x}As$ layer, since in the graded layers the spacing of the dislocation lines in the direction perpendicular to the interface is much larger than in the homogeneous composition buffer, greatly reducing the importance of the MD multiplication mechanism. Qualitatively, the reduced dislocation interaction in the graded buffer may explain why the roughness of the surface depicted in Fig. 4(b) is still of the same order of magnitude as that depicted in Fig. 4(a) despite the threefold increase in lattice mismatch.

IV. CONCLUSIONS

We have determined growth protocols and design parameters for the implementation of lattice-matched heterovalent $Zn_{1-y}Cd_ySe/In_xGa_{1-x}As$ heterostructures on GaAs(100) in the composition range of interest for blue–green optoelectronic applications ($0<y<0.3$). Except for the smallest Cd and In concentrations explored, minimization of the threading dislocation density and surface roughness requires the use of graded composition III–V ternary layers. In all cases, the In concentration within the III–V ternary layer has to be selected so that the partial character of strain relaxation is appropriately compensated. Final strains at or below the 0.1% level for the II–VI epilayers are routinely achieved by molecular beam epitaxy.

The resulting II–VI epilayers are dislocation free, except for the partial Frank and Shockley dislocations bounding the stacking faults originating at the II–VI/III–V interface. The surface morphology of the II–VI overlayers is found to exhibit a crossed-hatch pattern of surface corrugations with amplitudes in the 3–10 nm range and periods of 0.3–1.5 μm. Its origin maybe purely plastic, and follow from the formation of slip steps at the surface during dislocation multiplication at the epilayer–substrate interface.

ACKNOWLEDGMENTS

This work was supported in part by the Consiglio Nazionale delle Ricerche of Italy under the MADESS project and by the Commission of the European Communities. The authors thank G. Bauer, R. Cingolani, M. Lomascolo, and Y. Zhuang for useful discussions and for providing them with their results prior to publication.

[1]A. Franciosi and C. G. Van de Walle, Surf. Sci. Rep. **25**, 1 (1996).
[2]L. J. Brillson, in *Handbook on Semiconductors*, edited by P. T. Landsberg (North-Holland, Amsterdam, 1992), Vol. I, p. 281.

[3]W. Harrison, J. Vac. Sci. Technol. **16**, 1492 (1979); K. Kunc and R. M. Martin, Phys. Rev. B **24**, 3445 (1981); R. G. Dandrea, S. Froyen, and A. Zunger, *ibid.* **42**, 3213 (1990).
[4]S. Baroni, R. Resta, and A. Baldereschi, in *Proceedings of the XIX International Conference on the Physics of Semiconductors*, edited by W. Zawadzki (Institute of Physics, Polish Academy of Sciences, Wroclaw, 1988), p. 525; M. Peressi, S. Baroni, R. Resta, and A. Baldereschi, Phys. Rev. B **43**, 7347 (1991).
[5]J. Tersoff and C. G. Van de Walle, Phys. Rev. Lett. **59**, 946 (1987); R. G. Dandrea and C. B. Duke, J. Vac. Sci. Technol. B **11**, 1553 (1993).
[6]See, for example, Proc. SPIE **2346**, 1 (1994).
[7]J. Petruzzello, B. L. Greenberg, D. A. Cammack, and R. Dalby, J. Appl. Phys. **63**, 2299 (1988).
[8]S. Adachi, J. Appl. Phys. **53**, 8775 (1982).
[9]H. Jeon, J. Ding, A. V. Nurmikko, H. Luo, N. Samarth, J. K. Furdyna, W. A. Bonner, and R. E. Nahory, Appl. Phys. Lett. **57**, 2413 (1990); W. Xie, D. C. Grillo, R. L. Gunshor, M. Kobayashi, H. Jeon, J. Ding, A. V. Nurmikko, G. C. Hua, and N. Otsuka, *ibid.* **60**, 1999 (1992); H. Jeon, J. Ding, A. V. Nurmikko, W. Xie, D. C. Grillo, M. Kobayashi, R. L. Gunshor, G. C. Hua, and N. Otsuka, *ibid.* **60**, 2045 (1992).
[10]D. Hervé, R. Accomo, E. Molva, L. Vanzetti, J. J. Paggel, L. Sorba, and A. Franciosi, Appl. Phys. Lett. **67**, 2144 (1995).
[11]D. Herve, J. M. Bonard, L. Vanzetti, J. J. Paggel, L. Sorba, J. D. Ganiere, E. Molva, and A. Franciosi, J. Cryst. Growth **159**, 600 (1996).
[12]J.-M. Bonard, J.-D. Ganiere, L. Vanzetti, J. J. Paggel, L. Sorba, A. Franciosi, D. Herve, and E. Molva, J. Appl. Phys. **83**, 1945 (1998).
[13]A. V. Drigo, Mikrochim. Acta **114–115**, 89 (1994); A. Armigliato, M. Servidori, F. Cembali, R. Fabbri, R. Rosa, F. Corticelli, D. Govoni, A. V. Drigo, M. Mazzer, F. Romanato, S. Frabboni, R. Balboni, S. S. Iyer, and A. Guerrieri, Microsc. Microanal. Microstruct. **3**, 363 (1992).
[14]S. Heun, J. J. Paggel, L. Sorba, S. Rubini, A. Bonanni, R. Lantier, M. Lazzarino, B. Bonnani, A. Franciosi, J.-M. Bonard, J.-D. Ganiere, Y. Zhuang, and G. Bauer, J. Appl. Phys. **83**, 2504 (1998).
[15]B. H. Müller, R. Lantier, L. Sorba, S. Heun, S. Rubini, M. Lazzarino, A. Franciosi, E. Napolitani, F. Romanato, A. Drigo, L. Lazzarini, and G. Salviati (unpublished).
[16]B. H. Müller, R. Lantier, L. Sorba, S. Rubini, M. Lazzarino, S. Heun, A. Bonanni, G. Bratina, A. Franciosi, E. Napolitano, F. Romanato, and A. Drigo, J. Cryst. Growth **185**, 21 (1998).
[17]It is, therefore, $\epsilon_\parallel=(a_\parallel-a_0)/a_0$, $\epsilon_\perp=(a_\perp-a_0)/a_0$, where $a_\parallel$ and $a_\perp$ are the measured lattice parameters in the directions parallel and perpendicular to the interface, respectively, and a_0 is the equilibrium (unstrained) lattice parameter. In a biaxial strain field the two values are connected by the simple relations, $\epsilon_\perp=-\alpha\epsilon_\parallel$, $\alpha=2(C_{12}/C_{11})=2(\nu/1-\nu)$, where C_{12} and C_{11} are the elastic stiffness constants and ν is Poisson's ratio.
[18]J. Petruzzello, B. L. Greenberg, D. A. Cammack, and R. Dalby, J. Appl. Phys. **63**, 2299 (1988).
[19]N. Samarth, H. Lou, J. K. Furdyna, S. B. Qudri, Y. R. Lee, A. V. Ramdas, and N. Otsuka, Appl. Phys. Lett. **54**, 2680 (1989). These authors also reported a direct band gap of 1.67 eV at 300 K.
[20]*Semiconductors—Group IV Elements and III–V Compounds*, edited by O. Madelung, Data in Science and Technology Series (Springer, Berlin, 1991).
[21]E. A. Fitzgerald, Y. H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y. J. Mii, and B. E. Weir, Appl. Phys. Lett. **59**, 811 (1991).
[22]S. Guha, A. Madhukar, and K. J. Rajkumar, Appl. Phys. Lett. **57**, 2110 (1990).
[23]J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre', and O. Vatel, Appl. Phys. Lett. **64**, 196 (1994), and references therein.
[24]A. Bosacchi, A. C. De Riccardis, P. Frigeri, S. Franchi, C. Ferrari, S. Gennari, L. Nasi, G. Salviati, A. V. Drigo, and F. Romanato, J. Cryst. Growth **175/176**, 1009 (1997); L. Lazzarini, C. Ferrari, S. Gennari, A. Bosacchi, S. Franchi, F. Romanato, M. Berti, A. V. Drigo, and G. Salviati, Inst. Phys. Conf. Ser. **157**, 149 (1997).
[25]A. V. Drigo, A. Aydinli, A. Carnera, F. Genova, C. Rigo, C. Ferrari, P. Franzosi, and G. Salviati, J. Appl. Phys. **66**, 1975 (1989).
[26]J. Tersoff, Appl. Phys. Lett. **62**, 693 (1993).
[27]S. Heun, J. J. Paggel, S. Rubini, L. Sorba, A. Franciosi, J.-M. Bonard, and J.-D. Ganiere, Appl. Phys. Lett. **70**, 237 (1997); S. Heun, J. J. Paggel, L. Sorba, S. Rubini, A. Franciosi, J.-M. Bonard, and J.-D. Ganiere, J. Vac. Sci. Technol. B **15**, 1279 (1997).
[28]J.-M. Bonard, J.-D. Ganiere, S. Heun, J. J. Paggel, S. Rubini, L. Sorba, and A. Franciosi, Philos. Mag. Lett. **75**, 219 (1997).
[29]B. J. Wu, G. M. Haugen, J. M. DePuydt, L. H. Kuo, and L. Salamanca-Riba, Appl. Phys. Lett. **68**, 2828 (1996).
[30]L. H. Kuo, K. Kimura, S. Miwa, Y. Yasuda, and T. Yao, Appl. Phys. Lett. **69**, 1408 (1996); L. H. Kuo, K. Kimura, Y. Yasuda, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, *ibid.* **68**, 2413 (1996).
[31]R. Beanland, M. Aindow, T. B. Joyce, P. Kidd, M. Lourenço, and P. J. Goodhew, J. Cryst. Growth **149**, 1 (1995).
[32]K. H. Chang, R. Gibala, D. J. Srolovitz, P. K. Bhattacharya, and J. F. Mansfield, J. Appl. Phys. **67**, 4093 (1990).
[33]S. F. Yoon, J. Vac. Sci. Technol. B **11**, 562 (1993).
[34]S. Y. Shiryaev, F. Jensen, and J. W. Petersen, Appl. Phys. Lett. **64**, 3305 (1994).
[35]J.-M. Bonard, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 1996.
[36]M. S. Abrahams, J. Blanc, and C. J. Buiocchi, Appl. Phys. Lett. **21**, 185 (1972).
[37]J. W. Matthews, J. Vac. Sci. Technol. **12**, 126 (1975).
[38]J. W. Matthews, A. E. Blakeslee, and S. Mader, Thin Solid Films **33**, 253 (1976).
[39]J. M. Kang, C. S. Son, M.-S. Kim, Y. Kim, S.-K. Kim, and C. S. Kim, Appl. Phys. Lett. **67**, 641 (1995).
[40]F. Riesz, J. Cryst. Growth **140**, 213 (1994).
[41]G. Springholz, G. Bauer, and V. Holy, Surf. Sci. **365**, 453 (1996).
[42]W. Hagen and H. Strunk, Appl. Phys. **17**, 85 (1978).
[43]E. A. Fitzgerald, G. P. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall, J. Appl. Phys. **65**, 2220 (1989).

In situ determination of in-plane strain anisotropy in ZnSe(001)/GaAs layers using reflectance difference spectroscopy

K. Hingerl[a)]
Profactor GmbH, A-4400 Steyr, Austria

T. Yasuda
Joint Research Center for Atom Technology (JRCAT)–NAIR, Tsukuba 305, Japan

T. Hanada
Joint Research Center for Atom Technology (JRCAT)–NAIR, Tsukuba 305, Japan and Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

S. Miwa, K. Kimura, and A. Ohtake
Joint Research Center for Atom Technology (JRCAT)–ATP, Tsukuba 305, Japan

T. Yao
Joint Research Center for Atom Technology (JRCAT)–NAIR, Tsukuba 305, Japan and Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

(Received 21 January 1998; accepted 26 May 1998)

In this article we show *in situ* and *ex situ* experimental reflectance difference spectroscopy (RDS) data of ZnSe(001) grown on GaAs, where special attention is focused on the resonancelike structure at the band gap of ZnSe. This feature was reported previously and interpreted as either the transition from the valence band of ZnSe to a quantum well state at the ZnSe/GaAs interface or it was attributed to interface states. We report a new interpretation of this spectral feature which considers anisotropic strain in the ZnSe/GaAs heterostructures. By straining the ZnSe(001) epilayers mechanically, it was observed that the resonance feature changed sign and magnitude, proving that the height of this in-plane anisotropy feature is a measure for the in-plane strain. In-plane strain induces a splitting and shift in energy of the light and heavy hole valence bands. The transition between the light hole valence band and the conduction band is not polarization dependent, however the transition between the heavy hole valence band and the conduction band, as well as the transition between the spin split-off band and the conduction band shows polarization dependence. This difference between the reflectances parallel and perpendicular to the in-plane strain direction can be measured *ex situ* and *in situ* with RDS. © *1998 American Vacuum Society.*
[S0734-211X(98)10804-1]

I. INTRODUCTION

A continuing goal of epitaxial crystal growth is the production of high quality material and *in situ* monitoring techniques are extremely helpful to reach this target. Optical probes as, e.g., reflectance difference spectroscopy (RDS) or spectroscopic ellipsometry (SE) are best suited to be applied simultaneously with crystal growth, because they are noninvasive and nondestructive. Within the several last years understanding of information delivered by RDS and by kinetic RD data has grown considerably, however full exploitation of the power of these optical techniques needs further investigation particularly where heteroepitaxial systems are concerned. RDS, which measures in normal incidence the difference of the reflectances between the eigenaxes along the (110) and the $(\bar{1}10)$ directions of the sample, was only recently applied to investigate heterostructures,[1,2] where optical anisotropies are due to various origins. For the ZnSe/GaAs system, three possible major contributions were identified in Ref. 1 by a mathematical procedure as contributing to the measured RDS signal: (1) the interface induced anisotropy, (2) the surface induced anisotropy, and (3) a (complex) difference in the dielectric functions of the epilayer along the principal in-plane axes (dichroism).

In Ref. 1, it was demonstrated that *in situ* RDS can simultaneously measure the surface and the interface induced optical anisotropies, and that the spectral features occurring for epilayer thicknesses between 30 and 400 Å are almost independent of the layer thickness.

The ZnSe/GaAs electronic and structural interface properties, are of high scientific interest because this interface exhibits the effects of heterovalency as well as very close lattice match. This makes this system an ideal candidate to investigate structural properties such as strain due to lattice mismatch $[(a_{sub}-a_{epi})/a_{epi}]$. This number is 2.7×10^{-3} at room temperature (RT) and 3.03×10^{-3} at 250 °C, where the increase is due to different thermal expansion coefficients. Also, misfit dislocations and the value of the critical thickness (1500 Å) are important structural properties.[3] From the technological point of view, this interface is most critical to the lifetime of blue/green semiconductor lasers, which is inversely proportional to the density of defects originating at this interface.[4]

Effects of residual strain on the properties of thin ZnSe

[a)]Electronic mail: kurt.hingerl@pr-steyr.ac.at

films have been extensively characterized by photoluminescence (PL) as well as by x-ray diffraction.[3] It turned out that up to the critical thickness of 1500 Å the films are tetragonally distorted with two equal in-plane lattice constants and a different out-of-plane lattice constant. To our knowledge, there has been no observation of in-plane strain anisotropy in epitaxial ZnSe/GaAs layers up to now.

In this article we show *in situ* and *ex situ* experimental RDS data of ZnSe(001) grown on GaAs, where attention is focused on the resonancelike structure near the band gap of ZnSe (2.5–2.67 eV, depending on the temperature). This feature was already published in Ref. 5 by Yang *et al.* and interpreted as the crossover transition from the valence band of ZnSe to a quantum well state at the ZnSe/GaAs interface. They changed their interpretation in a later publication,[6] where they claimed the feature comes from interface states. This shows that the interpretation is rather tentative. The same type of resonance feature can, for example, be seen in the RDS spectra of 7, 14, and 36 nm thick ZnSe(001) epilayers (Fig. 1)[1] as well as in SE data.[7] We thus report a new interpretation of this spectral feature which considers anisotropic in-plane strain in the ZnSe/GaAs heterostructures. We also present data taken *in situ* at different epilayer thicknesses and with different interface reconstructions due to different surface terminations. We also present an interpretation of the origin of the in-plane strain anisotropy. In our opinion this strain is an indication that the length of the projection of Zn–Se backbonds onto the (001) plane is different from the length of the projection of Se–Zn backbonds onto the (001) plane. However, the physics of the stress causing this effect is not clear to us at the moment.

II. EXPERIMENT

The homemade RDS system is similar to the one reported by Aspnes *et al.*,[8] except for the use of a calcite Glan–Taylor prism which allows measurements in the spectral range between 1.5 and 5.8 eV. The monochromator used has a focal length of 0.1 m. In RDS, the measured quantity is the difference between the normal incidence reflectances for two orthogonal polarization directions in the surface plane. The results are displayed as $\Delta\tilde{r}/\tilde{r}=\Delta r/r+i\Delta\theta$ with $\tilde{r}=r\exp(i\theta)$. We define $\Delta\tilde{r}$ as $\tilde{r}_{\bar{1}10}-\tilde{r}_{110}$, where the subscripts denote the incident polarization vectors.

The *in situ* experiments were performed in a dual chamber molecular beam epitaxy (MBE) system equipped with *in situ* RDS and reflection high-energy electron diffraction (RHEED). After growth of an undoped homoepitaxial GaAs layer on a thermally cleaned GaAs(001) substrate, the substrate was transferred to the II–VI growth chamber. Prior to growth the surface showed an As stabilized (2×4) reconstruction, as confirmed by RHEED and RDS measurements. Because it is known,[9] that a low density of stacking faults is obtained in samples with a Zn treatment of the As stabilized (2×4) GaAs surface, this procedure was employed prior to growth. The *in situ* RDS data reported hereafter were taken, if not indicated otherwise, under "no flux" conditions. *In situ* data under Se or Zn flux can be found in Refs. 1 and 2. The substrate temperature was 250 °C, measured with a pyrometer, and the growth rate was 0.4 μm/h.

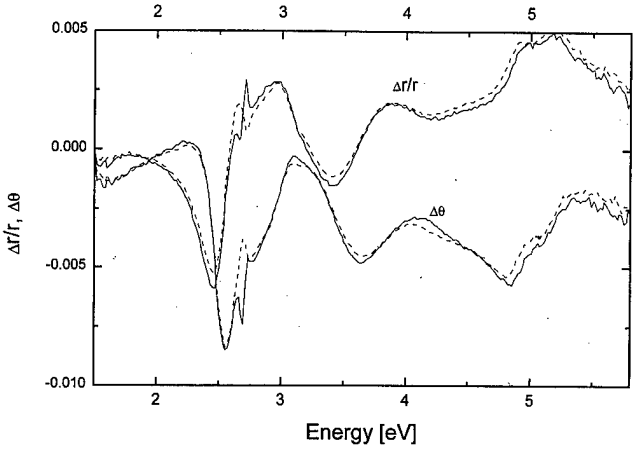


FIG. 1. Comparison of *ex situ* RDS data ($\Delta r/r$ and $\Delta\theta$) for a 1340 Å thick (001) ZnSe epilayer on GaAs. The dashed curve displays the spectra of the external unstressed layer, and the solid line displays the spectra of the mechanically stressed layer.

The *ex situ* RDS data were taken at RT on an optical bench, with the sample mounted within focus of the RDS setup. The ZnSe sample was strained mechanically for a few periods. No evidence of plastic deformation was found by straining it for a few cycles; the measured spectra were vitually the same.

III. RESULTS AND DISCUSSION

A. *Ex situ* measurements

Figure 1 shows the *ex situ* RDS ($\Delta r/r$ and $\Delta\theta$) data of a 1340 Å thick ZnSe layer measured, nominally unstrained and under tensile strain. The length dilatation $\Delta l/l$ along the [110] direction, applied by bending the sample, is crudely estimated from the radius of curvature of the reflected light to be about 3×10^{-4}. It can be seen, that the most significant differences occur at the critical points of the ZnSe band structure, at the E_0 (2.7 eV) and $E_0+\Delta_0$ (3.1 eV) transition, and in the E_1 (5 eV) region. The unstrained sample shows at 2.7 eV a broader structure than the strained one, which is probably due to slight thickness and strain variations in the sample. Applying an external stress, one stronger than the built-in stress, results in sharper features. Taking into account the fact that the polarization selection rules for the transition between the light hole (lh) band and the conduction band (cb) do not change with the applied stress[10] (the energy position does!), it is reasonable to assume that the resonance feature can only be due to the polarization dependence of the transition between the heavy hole band and the conduction band.

Comparing the RDS data of the stressed and the unstresses samples, one can see that without external stress there is either a small amount of tensile strain along the [$\bar{1}$10] direction or compressive strain along the [110] direction.

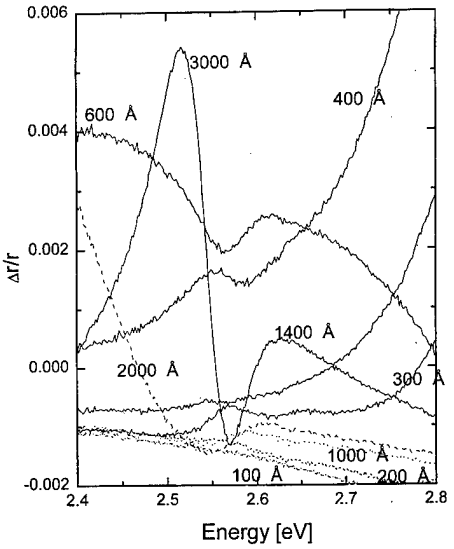


FIG. 2. *In situ* $\Delta r/r$ data taken at 100, 200, 300, 400, 600, 1000, 1400, 2000, and 3000 Å. The in-plane strain anisotropy can be seen between 2.56 and 2.62 eV. Please note that this resonance is superimposed by an increasing or decreasing background due to interference fringes. At 400 Å, e.g., there is a positive background, and at an energy of 2.56 eV the resonance is added to the fringe. At 600 Å the resonance is on a diminishing interference fringe, reversing the observed feature, although neither the strain nor the stress situation changed. For better discernibility the 100 Å curve is dashed dotted, the 200 Å curve is dashed, the 1000 Å curve short dashed, and the 2000 Å curve long dashed.

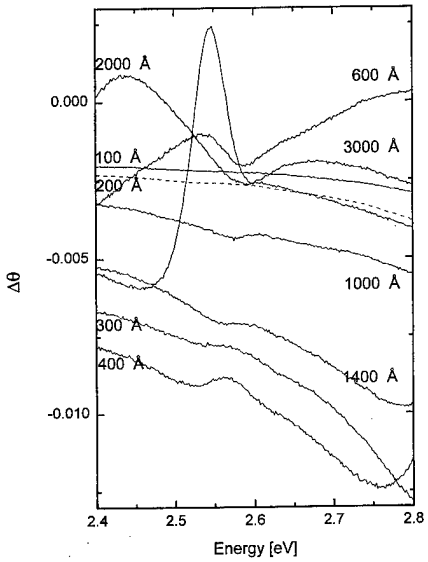


FIG. 3. *In situ* $\Delta\theta$ data taken at 100, 200, 300, 400, 600, 1000, 1400, 2000, and 3000 Å. The in-plane strain anisotropy can be seen between 2.56 and 2.62 eV. Please refer to the discussion in the text on the resonance shape and note the difference in the absolute scale between Fig. 2 and what is shown here. For better discernibility the 200 Å curve is dashed.

Both possibilities result in the same resonance signature. Because the lh−>cb transition cannot be observed by RDS, only the measurement of the energy shift of the heavy hole hh−>cb transition with and without strain would yield this distinction. This small strain-induced energy shift of the heavy hole band, however, cannot be detected by our experimental setup.

The feature observed is rather sharp on an energy (and also on a wave vector) scale compared to the broad structures of the surface or interface anisotropies. It can be argued that the sharpness of such a feature is a type of measure of delocalization, i.e., only a change in periodicity due to loss of long range order is expected to produce broad localized structures on an energy scale.

B. *In situ* measurements

The *in situ* RDS data were taken at 100, 200, 300, 400, 600, 1000, 1400, 2000, 3000, 5100, and 10 400 Å. In Fig. 2 the $\Delta r/r$ and in Fig. 3 the $\Delta\theta$ data (with the exception of the two largest thicknesses, where even derivatives do not show resonance features) at 250 °C under "no flux" conditions with closed Zn as well as Se Knudsen cells are displayed. Although it seems from Figs. 2 and 3 that the stress changes sign by going from 400 to 600 and then to 1000 Å, it has to be noted that in this measurement the signal heights are modulated by interference fringes arising from the interface induced anisotropy and the surface induced anisotropy.

In Fig. 2 one observes in the $\Delta r/r$ data that the resonance between 2.56 and 2.62 eV is superimposed on an increasing or decreasing background due to interference fringes. At 400 Å, for example, there is a positive background because the fringe peaks at 2.8 eV and the structure at 2.56 eV has the same sign as the fringe. At 600 Å the fringe peaks at 2.4 eV and the resonance is superimposed onto a decreasing interference fringe. Also in this case the structure at 2.56 eV has the same sign as the fringe. Therefore the built-in stress (or strain) situation did not change sign in going from 400 to 600 Å, nor does it change sign by increasing the thickness to 1000 Å, where the in-plane strain anisotropy feature shows the same resonance structure as at 400 Å. In Sec. III C2 (Fig. 9) we will discuss simulated $\Delta r/r$ as well as $\Delta\theta$ data, which show this oscillating behavior. In order to get a clear assessment of the sign of the strain (tensile or compressive) only samples at the same thickness can be compared or the interference fringes have to be modeled by simulation. The same is true for the $\Delta\theta$ data, with the exception that here the first fringe starts to develop even at smaller thicknesses with a negative slope.

Another interesting point is that the absolute height of the resonance in $\Delta\theta$ data is quite a bit higher than the height of the resonance in $\Delta r/r$ data. This can best be explained by referring to Ref. 8, where it was shown that birefringence differences in the real part of the dielectric function or the refractive index of, e.g., windows more strongly affect the measured $\Delta\theta$ data than the $\Delta r/r$ data. In Fig. 4 we plot the absolute value of the height of the resonance structure between 2.56 and 2.62 eV as a function of overlayer thickness L. It is observed that starting at a certain thickness (100–200 Å) the resonance structure starts to develop. Comparing this finding with the simulated data of Fig. 8, it becomes clear that a certain layer thickness is needed in order to enable the

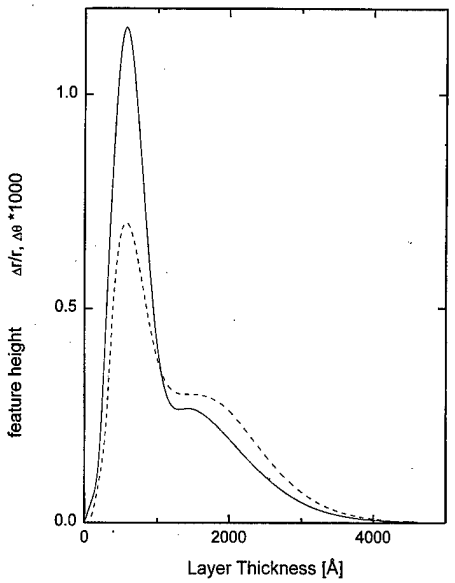


FIG. 4. Thickness dependence of the height of the absolute value of the resonance structure between 2.56 and 2.62 eV. The dashed curve displays the values of $\Delta r/r$, and the solid line displays the values of $\Delta\theta$.

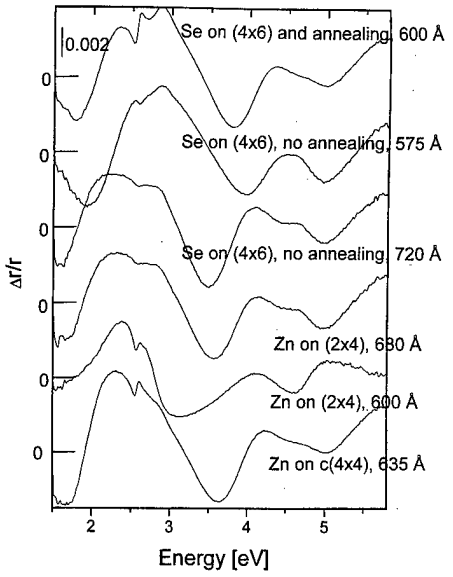


FIG. 5. Dependence of the resonance structure on interface preparation. The spectra are numbered from top to bottom: a, b1, b2, c, d, e. The second and the third curves from the top were measured on sample b at different thicknesses. All data with the exception of d were measured under Zn flux. The sample thicknesses are indicated.

RDS apparatus to pick up the difference in Δn or $\Delta\epsilon$.

A comparison to simulated data (Fig. 8, Sec. III C 2), where the growth temperature dielectric data of ZnSe and GaAs at 2.6 eV and also the surface and interface anisotropies were used,[2] shows that the increase of the absolute height of the RD signal can also be due to the increasing optical path length within this simulation where a constant anisotropy ($\Delta n = 0.01$) in the dielectric function of the ZnSe epilayer was assumed. The decrease seen in Fig. 4 can only be partly explained with the interference oscillations, because at higher layer thicknesses the signal should recover and even increase, especially if the major change in the dielectric function is in its real part.

The maximum height of the resonance is reached at 600 Å, and then the height of the resonance is reduced until it is no longer detectable at 3000 Å and above. These numbers fit quite well to the proposed critical thickness of approximately 1500 Å, although it has to be noted that RDS measures the in-plane strain anisotropy, and the reported numbers for critical thicknesses are coming from *ex situ* x-ray diffraction and photoluminescence data. Because the majority (>90%) of misfit dislocations in ZnSe/GaAs(001) layers have a 60° angle between the Burger's vector and the dislocation line (i.e., $b = a/2$ [110]), it is not unreasonable to assume that the introduction of misfit dislocations relaxes the in-plane as well as the out-of-plane strain.

Figures 5 and 6 display the dependence of the heights of the resonance structure on surface termination and on the interface structure. In Fig. 5 there are six RDS spectra displayed; they were taken on five samples all at layer thicknesses of around 600 Å at growth temperature. The layer numbers from top to bottom are a, b1, b2, c, d, e. Sample b was measured two times, once with a thickness of 575 Å and once with a thickness of 720 Å. The text in Fig. 5 indicates the interface structure of GaAs prior to growth and the II or VI compound which was deposited first at the interface possible annealing steps, as well as the exact thicknesses. All these spectra were taken under Zn flux with the exception of sample d, which was measured under no flux conditions. However, referring to Ref. 2, the differences in RDS data for different surface terminations below 3 eV are small.

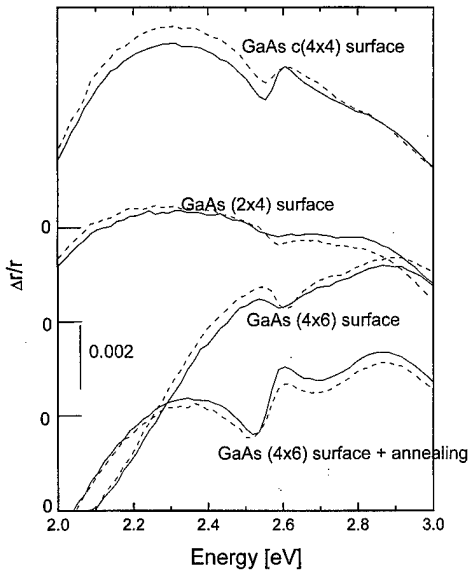


FIG. 6. Dependence of the resonance structure on surface termination. The dashed line represents data taken under Se flux, and the solid line represents data taken under Zn flux. The sample numbers from top to bottom are e, c, b1, and a.

All these samples (a, b1, b2, c, d, e) have thicknesses of around 600 Å, a thickness as was shown before where the signal is largest. There is only one case where the resonance structure is really strong, namely, when Se is deposited on a Ga rich (4×6) GaAs surface and an annealing step is used. We suppose that this annealing step rearranges the location of the Ga and/or the Se atoms and gives rise to an increased GaSe or Ga_2Se_3 layer thickness. When Zn is deposited on one As layer (2×4) surface or when Se is deposited on a (4×6) GaAs surface without postannealing, the height of the resonance structure is small. Zn on a double layer $c(4\times4)$ surface results in intermediate resonance heights.

Figure 6 shows the variation of the resonance structure of samples a, b1, c, and e at the thicknesses indicated in Fig. 5 in the energy range of 2–3 eV when the surface termination is changed. The dotted lines show the data taken under the Se beam, and the solid lines show the data taken under the Zn beam. The absolute as well as relative changes are quite strong and can be detected quite well with current RDS setups. The text in Fig. 6 indicates the interface structure. Although it cannot be argued that either one of the two surface terminations (Zn or Se) is increasing or decreasing the resonance height, it is seen that Zn coverage turns the resonance more to a structure, where the peak is at the high energy side, whereas Se coverage is changing the resonance, such that the peak is at the low energy side of the feature.

These experimental findings can be explained in the following way. Knowing, that the Se atom is larger and that it has four electrons more than the Zn atom, we assume that the Se exerts higher stress along the backbond directions (to the lower Zn layer) and thereby increases the tensile strain or reduces the compressive strain along the projection of the backbond direction on the (001) plane. This would imply, that resonance structures with the peak at the high energy side are a signature for compressive strain, and structures under tensile strain show the peak at the low energy side of the resonance structure. The same kind of behavior was also found at a layer thickness of 1400 Å. However, this argument is at the present stage pure hypothesis and has to be supported by further experiments.

C. Model

Here we will discuss the strain and polarization selection rules giving rise to the anisotropic optical behavior (Sec. III C 1), experimental procedures to determine the difference (anisotropy) in the dielectric functions along the two orthogonal axes (Sec. III C 2), and origins for the occurrence of anisotropic strain feature (Sec. III C 3).

1. Strain and polarization selection rules

In Fig. 7, on the left side, the band structure of ZnSe with the conduction band and three valence bands (lh: $J=3/2$, $m_j=\pm3/2$, hh: $J=3/2$, $m_j=\pm1/2$, spin split (ss) off band: $J=1/2$, $m_j=\pm1/2$) is shown for the unstressed case and on the right side it is shown under compressive stress. It has to be noted that not only the energy bands shift, but that the hh–>cb and the ss–>cb transition, become polarization dependent.[10,11] In this article only the results of Refs. 10 and 11 necessary to explain the experimental findings are reviewed.

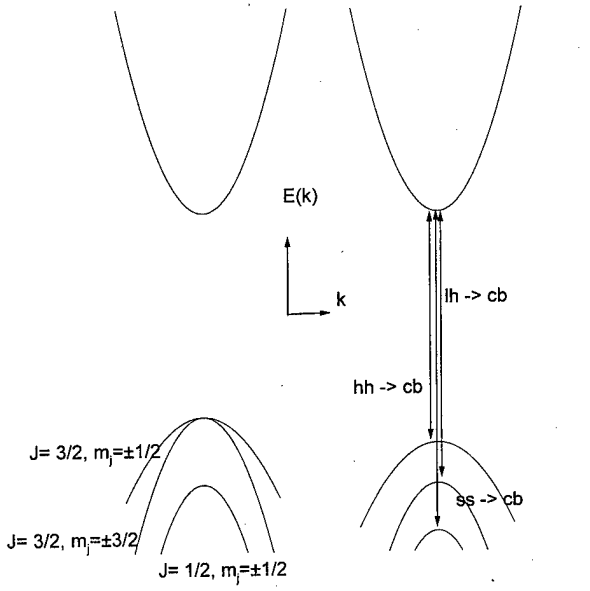


FIG. 7. Left side shows the three valence bands and the conduction band near $k=0$ without external stress. The right side shows the effect of compressive stress on the bands and the allowed transitions. The selection rules of the hh–>cb and the ss–>cb transitions are stress dependent.

From the wave functions of the three valence band states the selection rules and relative intensities for the lh–>cb, hh–>cb, and ss–>cb transitions can be calculated for light polarized parallel (∥) and perpendicular (⊥) to the stress axis in first order of the stress X (compressive stress is positive), respectively, α_0 with $\alpha_0=2b(S_{11}-S_{12})X/\Delta$:

$$\begin{aligned}&\text{lh}\text{->}\text{cb}: I^{\parallel}=0, \quad I^{\perp}=1/2\ P^2;\\ &\text{hh}\text{->}\text{cb}: I^{\parallel}=2/3\ P^2(1+\alpha_0); \quad I^{\perp}=1/6\ P^2(1-2\alpha_0);\\ &\text{ss}\text{->}\text{cb}: I^{\parallel}=1/3\ P^2(1-2\alpha_0), \quad I^{\perp}=1/3\ P^2(1+\alpha_0);\end{aligned} \tag{1}$$

P is the matrix transition element. The S_{ij} are the tabulated stiffness coefficients, calculated in Ref. 12 from the elastic moduli c_{11}, c_{12}, and c_{44},[13] Δ denotes the energy difference in the unstressed case between the lh, hh, and the ss bands, and b is the tetragonal shear deformation potential. (For ZnSe we use the following room temperature values: $b=-1.2$ eV, $S_{11}=2.3\times10^{-11}$ Pa^{-1}, $S_{12}=-8.9\times10^{-12}$ Pa^{-1}, $\Delta=0.4$ eV.)

The measured signal $\Delta r/r$ can be estimated for a stress tensor X by taking the difference of the intensities between the two orthogonal polarizations $[\bar{1}10]$ and $[110]$ of the projections parallel and perpendicular to the stress direction X:

$$\frac{\Delta r}{r}=2\frac{[I^{\parallel}_{\bar{1}10}(X)-I^{\parallel}_{110}(X)]+[I^{\perp}_{\bar{1}10}(X)-I^{\perp}_{110}(X)]}{I^{\parallel}_{\bar{1}10}(X)+I^{\parallel}_{110}(X)+I^{\perp}_{\bar{1}10}(X)+I^{\perp}_{110}(X)}. \tag{2}$$

In first order X the result for the lh–>cb transition is zero. Using the above values for the calculation of the heights of the resonance structure of the hh–>cb transition in ZnSe in

Eq. (2) for a uniaxial stress of 4.3×10^7 Pa along [100] (the main diagonal elements of the strain tensor are 0.001, −0.000 38, −0.000 38) yields

$$\frac{\Delta r}{r}=2\frac{2/3(1+\alpha_0)-2/3+1/6(1-2\alpha_0)-1/6}{2/3(1+\alpha_0)+2/3+1/6(1-2\alpha_0)+1/6}$$
$$\approx 2\alpha_0/5=0.0033. \tag{3}$$

This number seems reasonable as an estimate, but to get exact numbers on the stress or strain, it would be important to build a model, which yields the relative intensities of the polarization dependent absorption per unit length ($\Delta\epsilon$ or Δn). From the above estimation we conclude that the built-in in-plane stress in our epilayers corresponds to a value approximately of the order of 10^7 Pa.

It is not clear to us why the ss−>cb transition is not seen at 3.1 eV. Possible explanations are that the absorption of the epilayer is already so high that amplification by multiple reflections does not occur.

There is another possibility to establish a connection between $\Delta r/r$ and the calculated intensities of Eq. (1). To us it is currently not clear if Eq. (2) or the following equation should be used:

$$\frac{\Delta r}{r}=2\frac{r^{\parallel}(X)-r^{\perp}(X)}{r^{\parallel}(X)+r^{\perp}(X)}$$
$$=2\frac{[I_2^{\parallel}(X)+I_1^{\parallel}(X)]-[I_2^{\perp}(X)+I_1^{\perp}(X)]}{[I_2^{\parallel}(X)+I_1^{\parallel}(X)]+[I_2^{\perp}(X)+I_1^{\perp}(X)]}$$
$$=2\frac{[0+2/3(1+\alpha_0)]-[1/2+1/6(1-2\alpha_0)]}{[0+2/3(1+\alpha_0)]+[1/2+1/6(1-2\alpha_0)]}=1.5\alpha_0. \tag{4}$$

The signs for parallel and perpendicular are in this case equivalent to the [110] and the [$\bar{1}$10] directions. We do not favor Eq. (4) because the lh−>cb transition as well as the hh−>cb transition are at certain distinct energies and we only see one resonance. If the lh−>cb transition as well as the hh−>cb transition would be extremely sharp, we would consequently observe at two distinct energies, separated by an energy difference of $2b(S_{11}-S_{12})X$, two peaks with $\Delta r/r$ equal to ±1. This would imply that the heights of the resonance structure are due to the energy width of the transitions. However, in this case, with RDS one could observe also the energy positions of the transitions. There is one experimental fact against it: Stressing the sample does increase the heights of the resonance (e.g., by a factor of 3), however it does not measure its increase in the width of it. This indicates that the resonance originates from only one transition, which has to be the hh−>cb due to Eq. (1). We currently consider Eq. (2) to be the correct definition.

2. Optical model

We will now discuss the evolution of the resonance feature with increasing thickness and how the anisotropy in the dielectric function along two orthogonal directions can be obtained.

At energies and thicknesses where the epilayer is transparent, the measured $\Delta\tilde{r}/\tilde{r}$ can best be interpreted as a superposition of three terms as shown in expression (2) of Ref. 1:

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{1}{A}\left(B\frac{\Delta\tilde{r}_{oa}}{\tilde{r}_{oa}}+C\frac{\Delta\tilde{r}_{so}}{\tilde{r}_{so}}+\Delta n_0 C\frac{4\pi iL}{\lambda}\right), \tag{5}$$

with

$$A=(1+Zr_{oa}r_{so})(r_{oa}+Zr_{so}),$$
$$B=r_{so}(1-Zr_{so})(1+Zr_{so}), \tag{6}$$
$$C=Zr_{so}(1-r_{oa})(1+r_{oa}),$$

with the subscripts o, a, and s denoting the overlayer, the ambient, and the substrate, respectively. The quantities r_{so} and r_{oa} are the Fresnel reflection coefficients and $Z=\exp(4\pi i n_0 L/\lambda)$ with epilayer thickness L and wavelength of light λ.

The first term describes the surface induced anisotropy, the second the interface induced anisotropy, and the third the dichroism. In Eq. (6) the first and second terms also include the contribution of Δn as it is derived in Ref. 14. The quantity Δn_0 is related to the bulk effects in the overlayer; however, interference oscillations also occur if Δn_0 is zero. As proved in Ref. 1, Δn_0 is negligible for energies which are not at critical points of the epilayer. Knowing the interface $\Delta\tilde{r}_{so}/\tilde{r}_{so}$ as well as surface anisotropy $\Delta\tilde{r}_{oa}/\tilde{r}_{oa}$, it is, in principle, possible to determine Δn_0. However, from a numerical point of view the problem is ill posed, because two large numbers (the first and second terms on the right-hand side) are subtracted from the measured $\Delta\tilde{r}/\tilde{r}$, yielding large error bars for Δn_0. Furthermore, the dielectric functions of the epilayer as well as of the substrate have to be known with high accuracy in order to determine Δn_0.

It has to be noted that a five phase model (ambient-anisotropic surface-anisotropic epilayer-anisotropic interface-isotropic substrate) should be a good model for the ZnSe/GaAs system. However, the complexity, as well as the many parameters not known, only allow one to model such structures numerically. In the remainder of Sec. III C 2 we will present some analytical formulas, with which it is possible to relate the measured $\Delta\tilde{r}/\tilde{r}$ to anisotropies in the interface contribution to the dielectric function.

In the thickness and energy regime, where the epilayer is optically opaque (like an absorbing anisotropic substrate), so that no interference can occur, $\Delta\epsilon=(\epsilon_x-\epsilon_y)$ can be obtained from Eq. (5). The limit is well known and is a special case of equation (4.243) in Ref. 15:

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{2(\sqrt{\epsilon_y}-\sqrt{\epsilon_x})}{-1+\sqrt{\epsilon_x}\sqrt{\epsilon_y}}\approx\frac{\Delta\epsilon}{(-1+\epsilon_x)\sqrt{\epsilon_x}}. \tag{7}$$

With this equation it is possible to obtain from the RDS data the anisotropy in the dielectric functions for ZnSe in the energy region, where the epilayer is optically opaque.

If there is a very thin anisotropic overlayer with thickness d and anisotropy $\Delta\epsilon_0$ with an optically opaque epilayer with

a dielectric function ϵ_x and a small $\Delta\epsilon_e$, a series expansion in these two terms yields for the measured

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{\Delta\epsilon_e}{(-1+\epsilon_x)\sqrt{\epsilon_x}}+\frac{4\pi i d\Delta\epsilon_0}{\lambda(-1+\epsilon_x)}.$$

If a series expansion in the thickness for $\Delta\tilde{r}/\tilde{r}$ of the epilayer is not valid any more, the formulas become rather lengthy. For an anistropic epilayer of arbitrary thickness (index 1) on an isotropic substrate $\epsilon[2]$, a series expansion in the overlayer anisotropy $\Delta\epsilon[1]$ yields

$$\begin{aligned}\frac{\Delta\tilde{r}}{\tilde{r}}=&\left[\Delta\epsilon[1]\left(\frac{E^{8I\pi d[1]\sqrt{\epsilon_x[1]}/\lambda}\lambda(\sqrt{\epsilon_x[1]}-\sqrt{\epsilon_x[2]})^2}{\sqrt{\epsilon_x[1]}}-\frac{\lambda(\sqrt{\epsilon_x[1]}+\sqrt{\epsilon_x[2]})^2}{\sqrt{\epsilon_x[1]}}\right.\right.\\&\left.\left.+4E^{4I\pi d[1]\sqrt{\epsilon_x[1]}/\lambda}(2I\pi d[1](\epsilon_x[1]-\epsilon_x[2])+\lambda\sqrt{\epsilon_x[2]})\right)\right]\\&\times[\lambda(E^{8I\pi d[1]\sqrt{\epsilon_x[1]}/\lambda}(-1+\epsilon_x[1])(\sqrt{\epsilon_x[1]}-\sqrt{\epsilon_x[2]})^2+(-1+\epsilon_x[1])(\sqrt{\epsilon_x[1]}+\sqrt{\epsilon_x[2]})^2\\&-2E^{4I\pi d[1]\sqrt{\epsilon_x[1]}/\lambda}(1+\epsilon_x[1])(\epsilon_x[1]-\epsilon_x[2]))].\end{aligned}\tag{8}$$

Provided that the interface and the surface anisotropy are equal to 0, it would be possible to calculate $\Delta\epsilon[1]$; however, the same arguement as above with the ill-posed numerical problem and insufficiency knowledge of the dielectric functions of the overlayer and the substrate is valid.

In Fig. 8 simulated RDS data at 2.6 eV ($\Delta r/r$ and Δq) are shown as a function of layer thickness from 0 to 2000 Å using the growth temperature dielectric functions and the interface and surface anisotropies published.[2] If no layer anisotropy is assumed, the data slightly decrease with increasing thickness. If the anisotropy Δn_0 is assumed to be 0.01 (only real), an increase in the interference would be observed if there is no other mechanism such as misfit dislocations. In Fig. 9 simulated spectral data of $\Delta r/r$ and $\Delta\theta$ at thicknesses of 400 and 600 Å are shown, also using the growth temperature dielectric functions and the interface and surface anisotropies. The refractive index anisotropy was also assumed here to be $0.01+i0$. As asserted before, the peak sign really changes due to interference.

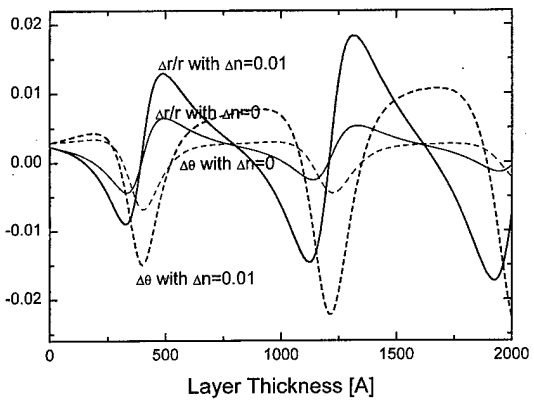


FIG. 8. Simulated evolution of the $\Delta r/r$ and $\Delta\theta$ values at 2.6 eV up to 200 nm using the growth temperature dielectric functions and the interface and surface anisotropies of Ref. 2.

3. Origin of the anisotropy

To us it is not perfectly clear at the moment, why an anisotropic in-plane strain results, however we would like to point out the following arguments.

(1) The linear electrooptic effect (LEO) induces similar sharp feature at E_0, but our analyses using spectroscopic data of the LEO coefficient have shown that the observed intensity and thickness dependency cannot be explained by LEO.[16]
(2) The observation of the peak at growth temperature [and temperature-dependence experiments on (110) samples] excludes the possibility of excitonic transitions being involved since these are highly temperature dependent.
(3) Comparison of the thickness-dependence data (Fig. 2) and the simulation suggests that the observed strain anisotropy is reduced upon lattice relaxation, which leads us to the conclusion that the length of the projection of

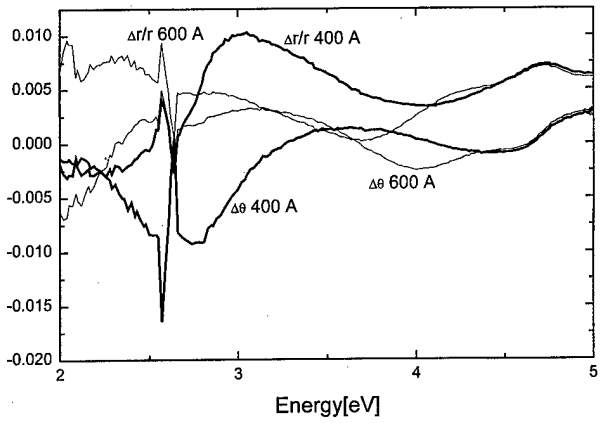


FIG. 9. Simulated spectral data of $\Delta r/r$ and $\Delta\theta$ between 2 and 3 eV at thicknesses of 400 and 600 Å, using the growth temperature dielectric functions and the interface and surface anisotropies of Ref. 2. The refractive index anisotropy was assumed to be $0.01+i0$. Please note that the peak sign changes due to interference.

Zn–Se backbonds onto the (001) plane is different from the length of the projection of Se–Zn backbonds onto the (001) plane if the anisotropy feature is observed. This interpretation is also consistent with the dependence on the interface structure and on the surface termination.

There have been additional reports on in-plane anisotropies in the II–VI heterostructures obtained with different experimental techniques: RHEED observations performed by Eymery *et al.*[17] in the ZnTe/CdTe strained layers showed a different evolution of the $[\bar{1}10]$ and the [110] azimuths in the oscillation intensity up to a critical thickness of 5 Å and in the width of the streaks going above the critical thickness. *Ex situ* SE measurements performed on ZnSe/GaAs heterostructures reported by Dahmani *et al.*[7] showed a resonance type structure at the E_0 transition at 2.7 eV as well as at the $E_0+\Delta_0$ position at 3.1 eV. The latter authors observed the transition from the light valence band to the conduction band, which could be due to an effect in the second order of strain. These authors attribute the resonance structures to a change in the out-of-plane lattice parameter, which is perhaps not consistent with the fact that SE measures the projection of the dielectric tensor along the line defined by the intersection of the plane of incidence and the sample surface.[18] In our opinion, their results could possibly be interpreted also in terms of differences in the in-plane lattice constants, giving rise to an in-plane strain anisotropy.

IV. CONCLUSION

We have shown *in situ* data for a ZnSe epilayer on GaAs at various thicknesses and discussed what amount of the lattice mismatch is accommodated by in-plane strain and and when, according to the diminishing resonance measured by RDS, misfit dislocations start to form. The dichroism due to strain is shown as a function of energy at room temperature, and some dichroism also occurs in the $E_0+\Delta_0$ and in the E_1 region. RDS could therefore also be employed for measuring the difference in the piezooptical constants of strained epilayers. The possible technological importance is evident for self-assembled structures,[19] and also for tuning optical properties in laser structures.

ACKNOWLEDGMENTS

This study was partly supported by NEDO and the Visiting Researchers Program of NAIR. One of the authors (K.H.) is grateful to the Austrian Science Foundation for Grant No. P11672-PHY.

[1]T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, J. Vac. Sci. Technol. B **14**, 3052 (1996).
[2]T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, Phys. Rev. Lett. **77**, 326 (1996), and references therein.
[3]J. Petruzello, B. L. Greenberg, D. A. Cammack, and R. Dalby, J. Appl. Phys. **63**, 2299 (1988).
[4]A. Ishibashi, J. Cryst. Growth **159**, 555 (1996).
[5]Z. Yang, I. K. Sou, Y. H. Yeung, and G. K. L. Wong, Appl. Phys. Lett. **66**, 2235 (1996).
[6]Z. Yang, I. K. Sou, Y. H. Yeung, and G. K. L. Wong, J. Vac. Sci. Technol. B **14**, 2973 (1996).
[7]R. Dahmani, L. Salamanca-Riba, N. V. Ngyuen, D. Chandler-Horowitz, and B. T. Jonker, Appl. Phys. Lett. **64**, 3620 (1994).
[8]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[9]L. H. Kuo, K. Kimura, A. Ohtake, S. Miwa, T. Yasuda, and T. Yao, J. Vac. Sci. Technol. B **15**, 1241 (1997).
[10]F. H. Pollak and M. Cardona, Phys. Rev. **172**, 816 (1968).
[11]D. G. Thomas, J. Appl. Phys. **32**, S2298 (1961).
[12]*Fundamentals of Semiconductors*, edited by P. Y. Yu and M. Cardona, ISBN 3-540-61461-3.
[13]*Landolt Börnstein Series 17b*, edited by Gumlich (Theis, Tschierse), p. 144.
[14]T. Yasuda, Thin Solid Films (in press).
[15]R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Amsterdam, Amsterdam, 1977).
[16]T. Yasuda (unpublished).
[17]J. Eymery, B. Daudin, D. Brun-Le Cunff, and N. Boudet, Appl. Phys. Lett. **66**, 3456 (1996).
[18]D. E. Aspnes, J. Opt. Soc. Am. **70**, 1275 (1980).
[19]J. Tersoff, Phys. Rev. Lett. **76**, 1675 (1996); Phys. Rev. B **53**, 16 334 (1996).

Interpretation of reflectance anisotropy spectroscopy spectra of ZnSe(001) grown on GaAs(001) in terms of bulk, interface, and surface contributions

A. M. Frisch,[a] Ch. Schultz, T. Herrmann, and V. Emiliani
Technische Universität Berlin, Sekretariat PN6-1, Hardenbergstrasse 36, 10623 Berlin, Germany

D. Wolfframm and D. A. Evans
University of Wales, Aberystwyth, Ceredigion S423 3BZ, Wales, United Kingdom

M. Korn
Universität Würzburg, Germany

U. Rossow
TU Ilmenau, Germany

N. Esser and W. Richter
Technische Universität Berlin, Sekretariat PN6-1, Hardenbergstrasse 36, 10623 Berlin, Germany

(Received 21 January 1998; accepted 21 May 1998)

We combine reflectance anisotropy spectroscopy (RAS) with low energy electron diffraction, Auger electron spectroscopy, and soft x-ray photoemission spectra to analyze the surface optical anisotropy of ZnSe(001). Clean surfaces were prepared by thermal desorption of a protective Se cap from ZnSe layers of different thicknesses grown by molecular beam epitaxy on GaAs(001). Two surface reconstructions have been prepared by subsequent annealing, the Se-rich (2×1) reconstruction and the Zn-rich $c(2\times2)$ reconstruction. By modifying the surfaces either through submonolayer deposition of Sb or a short exposure to atmosphere it was possible to distinguish between surface and bulk/interface contributions to the optical anisotropy. Only on disordered, very Se-rich ZnSe(001) surfaces prepared at low annealing temperatures RAS features possibly related to electronic surface states are found. RAS spectra of the (2×1) and the $c(2\times2)$ surfaces are correlated with surface morphology and ordering rather than surface reconstruction and show features near the critical points of the bulk ZnSe band structure. © *1998 American Vacuum Society.* [S0734-211X(98)09204-X]

I. INTRODUCTION

In recent years reflectance anisotropy spectroscopy (RAS) has been established as an optical method for analysis of surface dielectric anisotropies.[1,2] Especially for GaAs(001) the optical response for different reconstructions was studied in detail. RAS features characteristics of different reconstructions were identified and related to the microscopic surface structure.[3,4] For ZnSe (another zincblende material) the optical response of the surface is not yet well understood.[5–9] Several *in situ* RAS studies during growth or growth interruption in molecular beam epitaxy (MBE) or metalorganic vapor phase epitaxy (MOVPE) have been reported. Despite reflection high energetic electron diffraction (RHEED), however, complementary surface analytical methods have not been used in combination therewith. Commonly, two surface reconstructions are found on ZnSe(001), the Se-rich (2×1) and the Zn-rich $c(2\times2)$.[5–10] Moreover, a change from (2×1) to a Se-rich $c(2\times2)$ was observed with increasing Se coverage.[7] Total energy calculations using the local density approximation-density functional theory (LDA-DFT)[11,12] have been performed to study the microscopic structure of the three different reconstructions. However, direct experimental evidence for the atomic structures is still missing. Surface states associated with the dangling bonds of Se dimers on the (2×1) reconstructed surface have been predicted from the LDA-DFT calculations.[11] In a recent angle resolved photoemission spectroscopy (ARPES) analysis a feature possibly correlated with one of these states has been found.[13] On the other hand, RAS studies during growth revealed several different characteristic line shapes, but the surface contribution to the optical anisotropy and its microscopic origin have not been addressed. Thus, in the present study the anisotropic optical response of clean ZnSe surfaces of MBE grown layers on GaAs(001) is studied under ultrahigh vacuum (UHV) conditions in combination with electron-spectroscopic measurements.

II. EXPERIMENT

We performed RAS, low energy electron diffraction (LEED), and Auger electron spectroscopy (AES) measurements on pseudomorphic (50, 90, and 160 nm) and relaxed (500 nm, 4 μm) ZnSe layers grown by MBE on GaAs(001). All samples were Se capped after growth for protection during the transport through ambient atmosphere. Surface analysis was performed in an UHV chamber equipped for LEED, RAS, and AES (base pressure $<10^{-10}$ mbar). Following the

[a]Corresponding author;
electronic mail: FRISCH@GIFT.PHYSIK.TU-BERLIN.DE

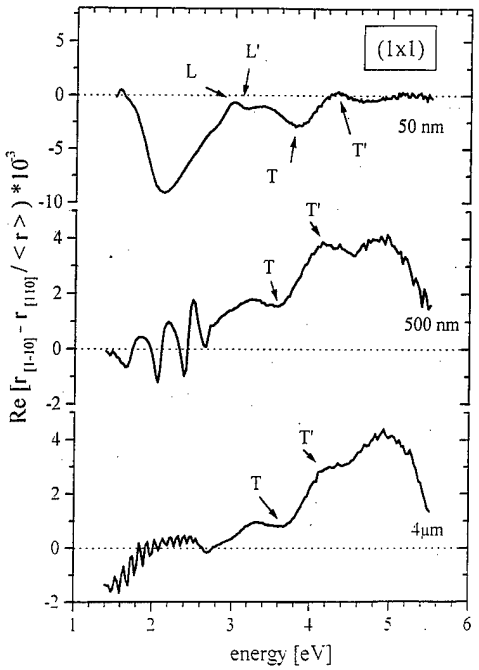


FIG. 1. RAS spectra (real part) of clean ZnSe(001) surfaces after decapping at 180 °C. ZnSe layer thicknesses are indicated in the plot.

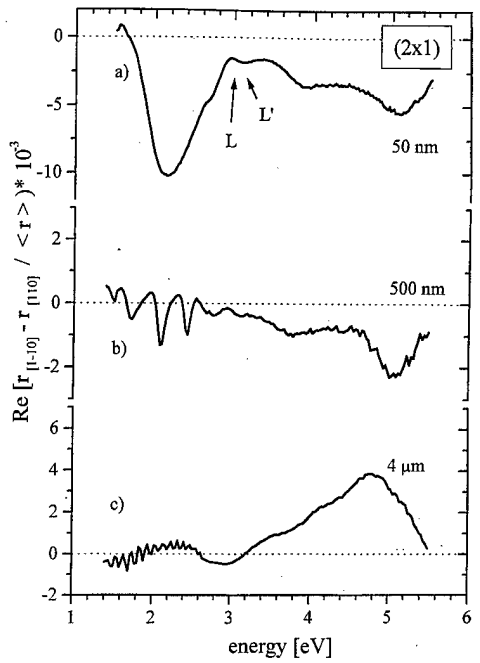


FIG. 2. RAS spectra (real part) of (2×1) reconstructed ZnSe(001) surfaces: thin, pseudomorphic layer with high degree of surface order (a), thick, relaxed ZnSe with high (b) and low (c) degree of surface order.

initial decapping at a temperature of 180 °C, well defined, reconstructed surfaces were prepared by several annealing cycles at temperatures up to 420 °C. Deposition of Sb onto the surface was performed at room temperature by evaporation of Sb from a Knudsen cell. The LEED system consisted of a four grid reverse view optics and a sensitive digital video camera for recording reflection patterns. AES spectra in the range of 30–1400 eV were taken for primary energies of 3 keV using a standard cylindrical mirror analyzer. The RAS spectra were taken *in situ* in the spectral range from 1.5 to 5.5 eV through a strain free quartz window. The RAS unit consisted of a Xe lamp, a single grating monochromator, a photoelastic modulator, Glan Air prisms and a S20 photomultiplier. Soft x-ray photoemission spectra (SXPS) on a 50 nm thick, pseudomorphic ZnSe layer were taken with an angle resolved photoelectron spectrometer (VG ADES 400) at the TGM2 beamline at BESSY (Berliner Elektronen-Speicherring Gesellschaft für Synchrotonstrahlung) storage ring in Berlin. All photoemission spectra were taken in normal emission with an overall resolution <0.3 eV at a photon energy of 100 eV. In order to determine the binding energies and the overall energy resolution the Fermi edge from a gold foil in electrical contact with the sample was recorded. The UHV chamber at BESSY was also equipped with a strain free window for RAS measurements.

III. RESULTS AND DISCUSSION

A. Clean ZnSe(001) surfaces

After decapping at 180 °C a weak (1×1) reconstructed, extremely Se-rich surface [coverage above 1 monolayer (ML)] is found. RAS spectra of this surface for different ZnSe layer thicknesses are shown in Fig. 1. Two features are found exclusively on these extremely Se-rich surfaces: a strong minimum at 3.9 eV (labeled T) and a weak maximum at 4.3 eV (labeled T'). For thick, relaxed layers an additional broad maximum centered around 5 eV appears which shifts the apparent position of T and T' somewhat to lower energies. On thin layers contributions from the interface are also seen, which will be discussed later.

Annealing the samples at temperatures between 200 and 290 °C results in the Se monolayer terminated (2×1) reconstructed surface, and the features T and T' disappear. Depending on layer thickness and surface quality, three different RAS line shapes for (2×1) reconstructed surfaces can be observed: Figure 2(a) shows the RAS spectrum of a thin, pseudomorphic ZnSe layer with a high degree of surface order, Fig. 2(b) shows a spectrum of a relaxed layer with high degree of surface order (sharp, bright half order LEED spots), and Fig. 2(c) show a spectrum of a thick relaxed ZnSe layer with a low degree of surface order (broad, weak half order LEED spots). Despite Fabry–Pérot oscillations between 1.5 and 2.7 eV (below the E_0 gap of ZnSe), two features are seen in Fig. 2(c): a minimum near the $E_0+\Delta_0$ gap of ZnSe at 3 eV and a broad feature with a maximum at 4.8 eV close to the E_1 gap of ZnSe. The RAS spectra in Figs. 2(a) and 2(b) (high surface order) instead show mainly a dielectric function derivative-like shape with main features at 4.6 eV (maximum) and at 5.05 eV (minimum), close to the E_1 and $E_1+\Delta_1$ gaps of ZnSe, respectively.

Further annealing in the range between 300 and 420 °C results in a $c(2\times2)$ reconstructed surface with a termination of half a monolayer of Zn. The corresponding RAS spectra (Fig. 3) show for all layer thicknesses the dielectric function derivative-like line shape, similar to those observed on the

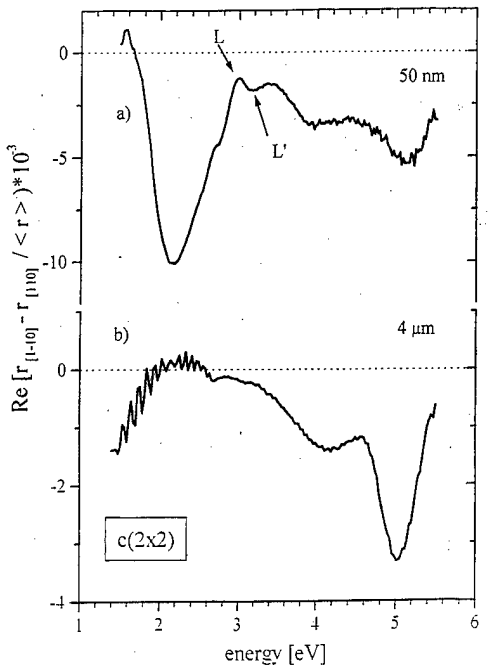


FIG. 3. RAS spectra (real part) of $c(2\times2)$ reconstructed ZnSe(001) surfaces: (a) thin, pseudomorphic layer; (b) thick, relaxed ZnSe layer.

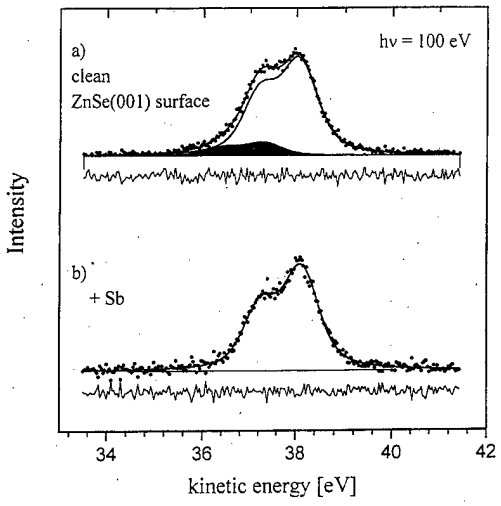


FIG. 4. Se $3d$ core level spectra of a fairly well ordered (2×1) reconstructed ZnSe(001) surface taken at a photon energy of 100 eV before (a) and after (b) Sb deposition.

well ordered (2×1) reconstructed surfaces [see Figs. 2(a) and 2(b)]. So far, we can summarize, that there exist two main RAS line shapes for clean reconstructed ZnSe(001) surfaces, which do not uniquely depend on the reconstruction, but merely on annealing temperature, degree of surface order, and layer thickness. Main features seen in the RAS spectra occur near gaps of ZnSe (despite features T and T') and the overall line shapes of the RAS spectra are either dielectric-function or dielectric-function-derivative-like.

Features T and T' found in the RAS spectra of the surfaces immediately after decapping (Fig. 1) at spectral positions distinct from the ZnSe band gaps may have two possible origins: either they are due to electronic surface transitions or to the optical response of Se crystallites on the surface. In the first case, the surface transitions would be associated with the very Se-rich $c(2\times2)$ reconstruction found in MBE at low substrate temperatures under Se background pressure.[7] The appearance of the weak (1×1) LEED image on the decapped surfaces would reflect a lower degree of long range order than on the MBE prepared surfaces. In fact, RAS spectra taken from this Se-rich $c(2\times2)$ reconstructed surface in MBE showed the same features,[7] however, the surface electronic band structure has not yet been determined and thus the possible attribution to surface states cannot be clarified.

Alternatively, the optical anisotropy of trigonal Se crystallites on the ZnSe surfaces could explain the T and T' features.[14,15] On the decapped surfaces Se crystallites due to cap remainders could still be present after the initial annealing step.[16] Likewise, Se crystallites might have nucleated on the ZnSe surfaces in MBE under extreme Se-rich conditions.[7]

The dielectric-function-derivative-like line shape of the RAS spectra is observed for both (2×1) and for $c(2\times2)$ reconstructions. However, it is dependent of surface order [Figs. 2(a), 2(b), and 3]. This line shape can be explained by surface modified bulk states.[17,18] Anisotropic contributions may be induced by the surface morphology,[8,9,19,20] by microscopic surface steps,[17] by microscopic surface strain in the top few layers (due to the different atomic positions at the surface with respect to the bulk,[11]) or by piezo-electric contributions.

The RAS spectra showing the broad, dielectric-function-like structure will be discussed in Sec. III B.

Interface contributions can be separated by the layer thickness dependence of the RAS line shape as has been shown by Yasuda *et al.* in detail.[8–10] In the transparent spectral region below the E_0 gap of ZnSe interface contributions are seen for all samples, however overlapped by Fabry–Pérot oscillations. Above 3 eV the light penetration depth in ZnSe (below 90 nm) limits interface contributions to thin ZnSe layers. Consistent with Refs. 8–10 the main interface contributions observed in our spectra [Figs. 1(a), 2(a), and 3(a)] are the linear electro-optic effect at the E_1 and $E_1+\Delta_1$ gaps of GaAs (at 2.9 and 3.1 eV, labeled L and L') on doped GaAs substrates ($n>10^{16}$ cm^{-3}), and a weakly structured negative background depending on initial growth conditions.

B. Adsorbate modified ZnSe(001) surfaces

In order to test the surface or bulk origin of the RAS features, the surfaces were modified either by submonolayer deposition of Sb or short-time exposure to atmosphere (below 1 min). Deposition of 0.8 ML Sb on a fairly well ordered (2×1) reconstructed surface leads to the deterioration of the LEED pattern to a weak (1×1) one. Corresponding SXPS spectra taken before and after Sb deposition are shown in Fig. 4. On the clean surface the Se $3d$ core level shows a component shifted to higher binding energies which disap-

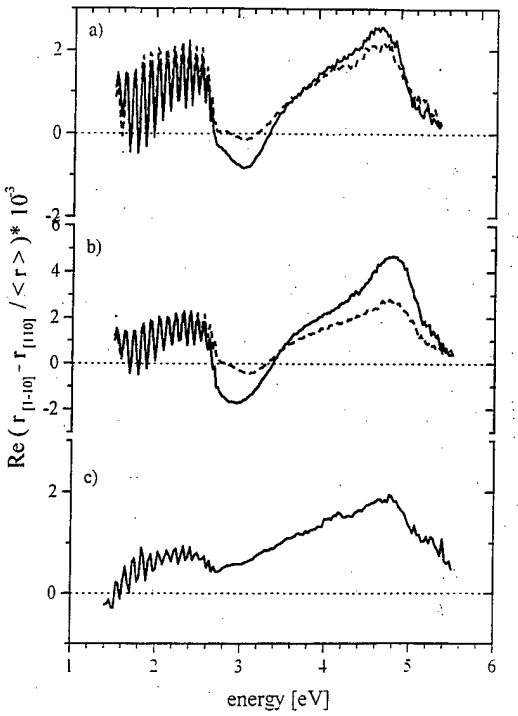


FIG. 5. RAS spectra (real part) for a fairly well ordered (2×1) reconstructed ZnSe(001) surface (layer thickness 4 μm): (a) before (full line) and after Sb deposition (dashed line); (b) before (full line) and after AES (dashed line); (c) initially $c(2\times2)$ reconstructed surface after contamination.

pears after Sb deposition. No new components arise in the Se $3d$ structure after Sb deposition. Also, no significant changes are found in the valence band spectra (not shown here). Therefore we conclude that Sb terminates the Se-rich ZnSe surface without inducing a strong chemical exchange reaction. We interpret the disappearance of the high energy Se $3d$ component as due to a reaction of Sb with the excess Se present on the surface (due to the limited resolution, the Se $3d$ emission of the reacted compound cannot be resolved in the core level spectra). As discussed already in Sec. III A, the excess Se could be due either to cap remainders[16] or to a Se-rich surface structure similar to the Se-rich $c(2\times2)$ reconstruction observed in MBE.[7]

The RAS spectra of the fairly well (2×1) reconstructed surface before and after Sb deposition are shown in Fig. 5(a). The minimum at 3 eV (at the $E_0+\Delta_0$ gap of ZnSe) disappears after Sb termination, while the broad maximum at 4.8 eV (E_1 gap of ZnSe) remains nearly unaffected. Both features are sensitive to electron irradiation: AES measurements induce an overall reduction of the amplitude by a factor of 2 [Fig. 5(b)]. Furthermore, we would like to point out that the broad feature at 4.8 eV is strongly dependent on the ZnSe layer thickness: For thin pseudomorphic layers, well below the critical thickness of around 160 nm, no such feature is seen, whereas for thicker, relaxed layers this feature becomes prominent. Moreover, if an initially $c(2\times2)$ reconstructed surface showing the derivative-dielectric-function-like line shape is exposed to atmosphere, the broad feature appears again with an amplitude similar to that found after AES measurements [Fig. 5(c)]. Therefore we conclude that it is related to bulk properties, most likely induced by an anisotropic relaxation of strain.

The structure at 3 eV, on the other hand, disappear after Sb deposition on the surface. Interestingly this feature, which coincides with the spectral position of the $E_0+\Delta_0$ gap of ZnSe, is not seen on all disordered (2×1) reconstructed surfaces, but only on thick, relaxed ZnSe layers. Therefore we conclude that it originates from surface modified bulk states, similar to the derivative-dielectric-function-like RAS spectra discussed in the previous section.

IV. CONCLUSIONS

The surface optical anisotropy of ZnSe(001) layers grown by MBE on GaAs was studied for different ZnSe layer thicknesses and reconstructions. Surface contributions were derived by comparing RAS with LEED, AES, and SXPS results for clean and adsorbate-modified surfaces. Only a few characteristic RAS line shapes are reproducibly found which, however, are not uniquely correlated with the different surface reconstructions. We conclude that the RAS spectra of ZnSe(001) are mainly related to surface modified bulk states that contribute to the optical anisotropy depending on surface order, strain, and morphology.

ACKNOWLEDGMENTS

The authors would like to acknowledge financial support by the Bundesministerium für Bildung und Forschung (Grant No. 05 622KTA 0). One of the authors V.E. is grateful to the Alexander von Humboldt foundation for financial support. The authors would also like to thank the staff of BESSY for friendly cooperation and experimental help.

[1] *Epioptic: Linear and Nonlinear Optical Spectroscopy of Surfaces and Interfaces*, edited by J. McGilp, D. Weaire, and C. H. Patterson (Springer, Berlin, 1995).
[2] I. Kamiya, D. E. Aspnes, H. Tanaka, L. T. Florez, J. P. Harbison, and R. Bhat, Phys. Rev. Lett. **68**, 627 (1992).
[3] N. Esser, A. I. Shkrebtii, U. Resch-Esser, C. Springer, W. Richter, W. G. Schmidt, F. Bechstedt, and R. Del Sole, Phys. Rev. Lett. **77**, 4402 (1996).
[4] D. Pahlke, M. Arens, N. Esser, D. T. Wang, and W. Richter, Surf. Sci. **352–354**, 66 (1996).
[5] D. Spaltmann, S. J. Morris, C. C. Matthai, and R. H. Williams, J. Phys. D **28**, 2574 (1995).
[6] H. Akinaga and K. Tanaka, Surf. Sci. **82**, 298 (1994).
[7] J. T. Zettler, K. Stahrenberg, W. Richter, H. Wenisch, B. Jobst, and D. Hommel, J. Vac. Sci. Technol. B **14**, 2757 (1996).
[8] T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, J. Vac. Sci. Technol. B **14**, 3052 (1996).
[9] T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, A. Ohtake, C. G. Jin, K. Tanaka, and T. Yao, J. Vac. Sci. Technol. B **15**, 1212 (1997).
[10] T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, Phys. Rev. Lett. **77**, 326 (1996).
[11] C. H. Park and D. J. Chadi, Phys. Rev. B **49**, 16467 (1994).
[12] A. Garcia and J. E. Northrup, Appl. Phys. Lett. **65**, 708 (1994).
[13] Y. Zhang, J. Y. Xue, J. A. Con Foo, A. P. J. Stampfl, D. Wolfframm, D. A. Evans, J. D. Riley, R. C. G. Leckey, A. Ziegler, B. Mattern, R. Graupner, M. Hollering, R. Denecke, and L. Ley, Surf. Sci. **377–379**, 288 (1997).
[14] S. Tutihasi and I. Chen, Phys. Rev. **158**, 623 (1967).

[15]A. G. Leiga, J. Opt. Soc. Am. **58**, 880 (1967).
[16]W. Chen, A. Kahn, P. Soukiassian, P. S. Mangat, J. Gaines, C. Ponzoni, and D. Olego, J. Vac. Sci. Technol. B **12**, 2639 (1994).
[17]L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes, J. Vac. Sci. Technol. B **15**, 1196 (1997).
[18]D. Olguin and R. Baquero, Phys. Rev. B **51**, 16 891 (1995).
[19]S. Tomiya, R. Minatoya, H. Tsukamoto, S. Itoh, K. Nakano, A. Ishibashi, and M. Ikeda, *Proceedings of the 23rd International Conference on the Physics of Semiconductors, Berlin 1996*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore), p. 1079.
[20]S. Miwa, K. Kimura, T. Yasuda, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, *Proceedings 23rd International Conference on the Physics of Semiconductors, Berlin 1996*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore), p. 899.

Reflectance difference spectroscopy spectra of clean (3×2), (2×1), and $c(2\times2)$ 3C-SiC(001) surfaces: New evidence for surface state contributions to optical anisotropy spectra

U. Rossow[a)]
TU Ilmenau, Inst. f. Physik, Pobox 100565, D-98684 Ilmenau, Germany

K. Lindner and M. Lübbe
TU Chemnilz, Professur f. Halbleiterphysik, D-09107 Chemnitz, Germany

D. E. Aspnes
Physics Department North Carolina State University, Raleigh, North Carolina 27695-8202

D. R. T. Zahn
TU Chemnitz, Professur f. Halbleiterphysik, D-09107 Chemnitz, Germany

(Received 21 January 1998; accepted 26 May 1998)

In this study we applied the surface sensitive optical method reflectance difference spectroscopy to investigate the optical response of clean cubic 3C-SiC(001) surfaces. The main reconstructions, the Si-rich (3×2) and (2×1) as well as the C-rich $c(2\times2)$ were prepared by annealing in Si flux. For (3×2) we find a strong spectral feature at 4.2 eV that is not observed for the other two reconstructions. Since the direct gap of 3C-SiC is at higher photon energies the 4.2 eV feature must be related to the surface. Angle resolved ultraviolet photoelectron spectroscopy measurements indicate surface states around the $\bar{X}$ point of the unreconstructed surface Brillouin zone. We conclude that the 4.2 eV feature is associated with this surface state and suggest that it is linked to the E_2 gap of bulk Si at 4.2 eV. © *1998 American Vacuum Society.* [S0734-211X(98)10904-6]

I. INTRODUCTION

Reflectance-difference spectroscopy/reflectance anisotropy spectroscopy (RDS/RAS) is now an established surface diagnostic that is used rather extensively to monitor surfaces during processes such as epitaxial growth and to characterize well defined clean and passivated surfaces. However, the origins of surface-induced optical anisotropy (SIOA) spectra, as determined by RDS, remain unclear.

For many surfaces we find that RDS spectra have dielectric function- and derivative-like line shapes that are dominated by the (surface modified) bulk dielectric function or its energy derivative.[1–8] In these cases the origin of SIOA spectra can be regarded as surface modified bulk optical responses.[9] In general SIOA spectra are more complicated and show both characteristics.

In addition, contributions involving surface states are expected. It is unclear in which situations surface states contribute and whether their contribution can be included in the basic line shapes.

Some candidates for surface state contributions to the optical response are arsenic-rich GaAs(001) surfaces. In previous work it has been shown that RDS signals of GaAs(001) surfaces are extremely sensitive to the chemical composition, i.e., the As/Ga ratio. For two of the corresponding arsenic-rich reconstructions; the (2×4) and the $c(4\times4)$, it has been argued that the signal originates from As dimer related states. However, recent findings that the line shape of the (2×4) RDS spectra resemble that of the bulk dielectric function ϵ_b and that the spectral feature at 2.6 eV shows the same temperature dependence as that of the E_1 interband critical point of GaAs indicate that surface states are not the origin of the (2×4) RDS signal.[5] In contrast, for $c(4\times4)$ ϵ_b describes only the part of the spectrum for photon energies above 3.5 eV and the temperature dependence of the 2.6 eV feature is different. Hence, it is likely that the 2.6 eV feature is related to surface states. Although many attempts have been made to calculate SIOA spectra for $c(4\times4)$ GaAs(001) the differences in line shape between theory and experiment are too large to decide this question from theory.

In order to obtain more data to examine the question of the contribution of surface states we investigated clean cubic 3C-SiC(001) in the visible and near ultraviolet (UV) spectral range below the direct gaps of 3C-SiC. Since 3C-SiC is an indirect semiconductor we do not expect (strong) bulk features below the direct gap of 5.8 eV and therefore, surface state can be more easily separated from bulk contributions compared to III–V and II–VI compound semiconductors.

In this article we investigate the main reconstructions for SiC(001), i.e., Si-rich (3×2) and (2×1) as well as C-rich $c(2\times2)$. A review of the various discussed surface structures can be found in Ref. 10.

II. EXPERIMENT

The samples used were commercially available (Cree Inc.) epitaxially grown cubic 3C-SiC on vicinal Si(001) 3.5° off towards [110]. The 3C-SiC material consists of two layers: First, a 2.5×10^{16} cm^{-3} n-type doped layer of 2 μm

a)Electronic mail: rossow@physik.tu-ilmenau.de

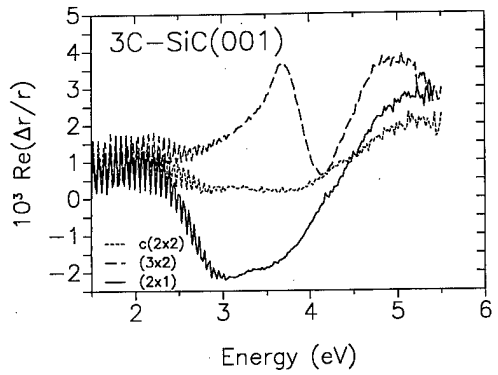


FIG. 1. Re $\Delta r/r$ for various reconstructions of SiC(001).

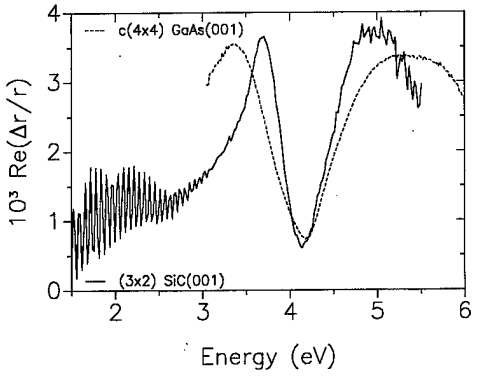


FIG. 2. Re $\Delta r/r$ of (3×2) SiC(001) compared to that of $c(4\times4)$ GaAs(001). The latter spectrum was rigidly shifted to higher energies by 1.55 eV.

thickness was deposited, then a second 6.9×10^{16} cm^{-3} p-type doped layer of thickness 3 μm was grown on top. The samples were prepared in an ultrasonic bath with organic solvents, rinsed in distilled water, and mounted on both sides with Ta clamps for direct heating in the ultrahigh vacuum (UHV) chamber. Cleanliness and surface order were checked by ultraviolet photoelectron spectroscopy (UPS) and low-energy electron diffraction (LEED). The LEED patterns reveal single-domain surfaces. The spots appeared to be sharp with low background with the exception of the (3×2) which showed some streaking of the half-order spots.

The different reconstructions were prepared by heating to temperatures between 870 and 1030 °C. Here, the Si-rich reconstructions of SiC were produced by additional Si flow. Details of the preparation procedure can be found in Ref. 11.

The RDS measurements were performed with an Aspnes-type spectrometer.[12,13] Due to the use of Glan polarizer prisms the spectral range was limited in the UV to photon energies less than 5.5 eV. The samples were measured in one azimuth and therefore no background corrections could be performed. However, from a measurement of an untreated sample we know that the background is small enough to be neglected.

RDS measures the difference in reflectance $\Delta r/r=2(r_\alpha-r_\beta)/(r_\alpha+r_\beta)$ between two directions α and β which are perpendicular to each other and the surface normal. We define α to be in the direction 2× and 3× for the (2×1) and (3×2) reconstructions, i.e., in the direction of the Si dimers (in view of the Dayan–Hara model),[14] respectively.

III. RESULTS AND DISCUSSION

Figure 1 shows RDS spectra for the main reconstructions of SiC(001). Sharp features are visible in the spectra below 3 eV due to layer interferences. The penetration depth of light in cubic SiC is much smaller than the layer thickness above photon energies of 3.7 eV. Therefore, SiC–Si substrate interface contributions vanish for photon energies above approximately 3.5 eV and all features above this energy are due to the SiC layer and surface. The $c(2\times2)$ and (2×1) surfaces do not show any sharp features. Which types of line shapes these spectra have is difficult to decide in this limited spectral range. It seems though that the $c(2\times2)$ has a derivativelike spectrum. The RDS spectrum is hardly different from that of the background spectrum. The (2×1) spectrum could be of dielectric function type. Most interesting is the large broad feature around 4.2 eV for the extremely Si-rich (3×2) reconstruction, which appears neither in (2×1) nor in $c(2\times2)$. The direct gaps of cubic SiC are all above 5.8 eV which is beyond the accessible spectral range and the untreated surface shows essentially no signal. Consequently, the 4.2 eV feature is most likely induced by the surface.

The 4.2 eV feature of (3×2) SiC(001) has a line shape similar to the 2.62 eV feature of $c(4\times4)$ GaAs(001), as demonstrated in Fig. 2. The main difference seems to be that the $c(4\times4)$ feature is much broader. The similarity is much more obvious when the energy axis is scaled, see Fig. 3. For a Lorentzian line shape scaling the energy axis is equivalent to changing the broadening parameter. From a comparison we can conclude that both spectral features have a similar origin. In the case of $c(4\times4)$ GaAs(001) we argue that the feature is due to the extra As dimers on top of the As-rich surface, while in the case of SiC(001) (3×2) the extra Si dimers are responsible.

Since SiC has no direct gaps in the spectral range considered here, the involved states must be localized at the surface. We suggest that the surface states are related to the Si dimers and have a character like that of bulk Si for the fol-

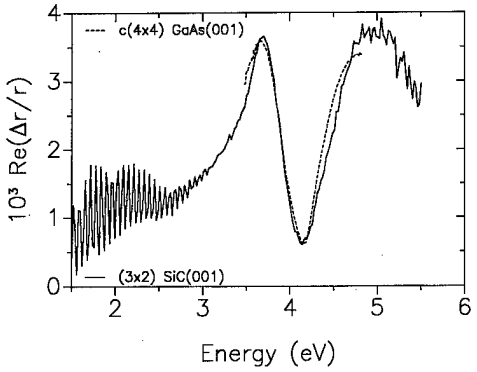


FIG. 3. Same as in Fig. 2 but with the energy axis scaled by a factor of 3/5.

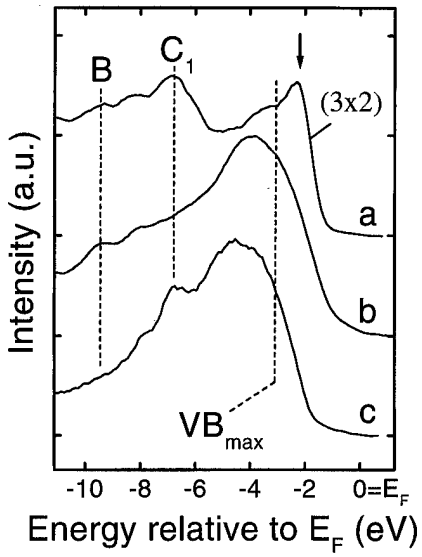


FIG. 4. Examples of angle resolved UPS spectra. The angle has been chosen such that there are strong contributions from the $\bar{X}$ point of the surface Brillouin zone of the unreconstructed surface. The surface state contribution for (3×2) can clearly be seen in the upper spectrum while it is missing in the lower two for (2×1) and *c*(2×2), respectively.

lowing reasons: (i) Bulk Si has the strongest spectral feature—the E_2 interband critical point at 4.2 eV—close to the (3×2) RDS feature. (ii) Angle resolved UPS investigations indicate that only the (3×2) reconstruction of SiC(001) has surface states located near the $\bar{X}$ point of the unreconstructed surface Brillouin zone (see Fig. 4). Now $\bar{X}=2\pi/a(1/2,1/2)$ is the projection of states especially of the *K* point of the bulk Brillouin zone and E_2 has large contributions from *K*. (iii) Finally, bulk As has its E_2 gap at 2.8 eV close to the RDS feature at 2.62 eV observed for *c*(4×4) GaAs(001) (at elevated temperatures).

IV. SUMMARY AND CONCLUSIONS

We found a broad but strong spectral feature at 4.2 eV for (3×2) SiC(001) which has a line shape similar to that of the 2.62 eV feature known for *c*(4×4) GaAs(001). These features are most likely related to the surface and can be linked to surface states. We suggest that the surface states have a character of the corresponding bulk materials which are Si and As in the case of (3×2) SiC(001) and *c*(4×4) GaAs(001), respectively. Consequently, we observe the features at spectral positions close to the E_2 gaps of the bulk material, which have the highest oscillator strength. However, the appearance of surface state features in RDS spectra does not imply that surface states contribute directly to RDS spectra. Still it is possible that surface states contribute only via the surface local field effect and that RDS spectra are of the dielectric function type. Future research will clarify this point.

ACKNOWLEDGMENTS

This work was funded by the Alexander von Humboldt Foundation, by the Deutsche Forschungsgemeinschaft (DFG) through the Innovationskolleg "Methoden und Matcrialien für den Nanometerbereich," and by the Office of Naval Research under Contract No. N-00014-93-1-0255.

[1]T. Yasuda, L. Mantese, U. Rossow, and D. E. Aspnes, Phys. Rev. Lett. **74**, 3431 (1995).
[2]U. Rossow, L. Mantese, T. Yasuda, and D. E. Aspnes, Appl. Surf. Sci. **104/105**, 137 (1996).
[3]L. Mantese, U. Rossow, and D. E. Aspnes, Appl. Surf. Sci. **107**, 35 (1996).
[4]U. Rossow, L. Mantese, and D. E. Aspnes, J. Vac. Sci. Technol. B **14**, 3070 (1996).
[5]U. Rossow, L. Mantese, and D. E. Aspnes, *On the Origin of Surface Induced Optical Anisotropy Spectra, Proceedings of the 23rd International Conference on the Physics of Semiconductors (ICPS), Berlin, Germany, July 21–26, 1996*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), Vol. 2, p. 831; M. Wassermeier, J. Beherend, J.-T. Zettler, K. Stahrenberg, and K. H. Ploog, Appl. Surf. Sci. **107**, 48 (1996).
[6]U. Rossow, N. V. Edwards, M. D. Bremser, R. S. Kern, H. Liu, T. W. Weeks, Jr., R. A. Stall, R. F. Davis, and D. E. Aspnes, Mater. Res. Soc. Symp. Proc. **449**, 781 (1997).
[7]L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes, J. Vac. Sci. Technol. B **15**, 1196 (1997).
[8]L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes, Thin Solid Films **313–314**, 557 (1998).
[9]U. Rossow, L. Mantese, and D. E. Aspnes, Appl. Surf. Sci. **123/124**, 237 (1998).
[10]V. M. Bermudez, Phys. Status Solidi B **202**, 447 (1997).
[11]M. Lübbe, K. Lindner, S. Tautz, S. Sloboshanin, J. A. Schaefer, and D. R. T. Zahn, J. Vac. Sci. Technol. B (submitted).
[12]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[13]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[14]M. Dayan, J. Vac. Sci. Technol. A **3**, 361 (1985); S. Hara, J. Kitamura, H. Okushi, S. Misawa, S. Yoshida, and Y. Tokumaru, Surf. Sci. **357/358**, 436 (1996).

Processes of quantum dot formation in the InAs on GaAs(001) system: A reflectance anisotropy spectroscopy study

D. I. Westwood,[a] Z. Sobiesierski, and C. C. Matthai
Department of Physics and Astronomy, University of Wales, Cardiff, Cardiff CF2 3YB, United Kingdom

E. Steimetz, T. Zettler, and W. Richter
Institut für Festkorperphysik, Technische Universitat Berlin, Hardenbergstr.3b, D10623, Berlin, Germany

(Received 21 January 1998; accepted 28 May 1998)

Reflectance anisotropy spectroscopy (RAS) has been used to monitor the formation and development of InAs islands grown onto GaAs(001) substrates by molecular beam epitaxy. Time resolved measurements show entirely different responses at different photon energies, corresponding to different aspects of the islanding process. At a photon energy of 2.6 eV RAS is sensitive to the onset of islanding, whereas the 4.0 eV signal appears to be sensitive to the continuous interisland wetting layer. Thus, by using the 4.0 eV signal it is possible to follow the real time development of the islanded surface. In particular, it is found that (1) during growth and immediately beyond the point of islanding, a fixed fraction of the incoming flux is directly transferred to the islands, and this fraction increases with increasing growth temperature up to 100%; (2) as a consequence of this, the wetting layer can increase in thickness beyond the 1.6 monolayers islanding thickness, to at least ~2 monolayers; (3) the importance of islands acting as sinks for the acquisition of material from the wetting layer is revealed; (4) the dynamic equilibrium set up between the islands and the wetting layer is such that the thickness of the latter increases significantly with sample temperature. © *1998 American Vacuum Society.*
[S0734-211X(98)12204-7]

I. INTRODUCTION

It is well known that as films of InAs are grown on GaAs(001) they undergo a number of transitions as the amount of deposited material is increased.[1] Initially, the InAs is pseudomorphically strained to the GaAs substrate. However, above ~1.6 monolayers (ML) and due to the large mismatch (~7.2%) it islands coherently, the islands being separated by a continuous two-dimensional (2D) interisland wetting layer (WL). With increased thickness misfit dislocations are generated and the islands coalesce to form a continuous but relaxed film. Both pseudomorphic and coherently islanded InAs films embedded in GaAs have recently generated a large amount of interest as two-dimensional quantum wells[2] and quasizero dimensional quantum dots,[3] respectively. The main stumbling block to application in practical devices has been island uniformity, showing the need for a greater understanding and control of the whole process. Since measurements are required at the monolayer scale, this is not a simple task, and most *in situ* studies to date have been performed using scanning tunneling microscopy (STM). Although these have revealed a wealth of important information,[4,5] they only give a snapshot of the surface and it is entirely possible that significant redistribution occurs during the required postgrowth cooldown and sample transfer to the STM stage. By contrast reflectance anisotropy spectroscopy (RAS), a relatively new optical technique that is highly sensitive to surfaces, subsurfaces, and buried interfaces, is capable of real time monitoring in the growth environment and therefore offers great prospects for improving the understanding of such ultrathin systems.[6,7] In the RAS technique near normal incidence polarized electromagnetic radiation interacts with the surface (and bulk) electronic dipole moments and the difference in reflection between the two orthogonal polarization states is measured.

It has already been shown that the RAS spectra of differently reconstructed semiconductor surfaces, such as GaAs(001), exhibit particular identifiable features for each structure. Aspnes *et al.* demonstrated that some of these features (in this case at ~2.6 eV) could be associated with the dimer bonds on the surface and their orientation with respect to the polarization of the incident light.[6,8] Hence the RAS signal changes sign as the GaAs(001) surface changes from an As-rich $c(4\times4)$ to an As stabilized (2×4).

Using *ab initio* pseudopotential calculations Bass and Matthai obtained excellent agreement with experimental results for the As-rich $c(4\times4)$ GaAs(001) surface.[9] Subsequent analysis of the local density of states for the top layers of the GaAs surface led these authors to deduce that both the main features in the RAS spectrum (the dip at 2.6 eV and the peak at 3.6 eV) have their origins in transitions between very similar *bulklike* valence band states and surface states at two distinct energy levels in the conduction band. The dramatic difference between the (2×4) and $c(4\times4)$ signals at ~2.6 eV can then be understood with the energetically lower conduction band state being associated with the directional As–As dimer bonds. The importance of adatom species on the reflectance anisotropy has also been demonstrated through tight binding calculations on the Sb– and Sn–

[a]Electronic mail:westwood@cf.ac.uk

Si(111) $\sqrt{3}x\sqrt{3}$ reconstructed surface and pseudopotential calculations of Si–GaAs $c(4\times4)$.[10]

In this study the first few monolayers of InAs deposited on GaAs(001) by molecular beam epitaxy (MBE) were studied *in situ* by both RAS and reflection high-energy electron diffraction (RHEED). Particular attention was paid to temperature and temporal effects through time resolved RAS measurements at different photon energies. In this case it appears that the signal at 4.0 eV is the most interesting since it seems to be intimately related to the wetting layer and largely independent of the islands, at least at the early stages of relaxation which are of most interest for use in quantum dot structures. It will be shown that this allows the dynamics of surface reordering to be followed in real time.

II. EXPERIMENT

All experiments were carried out using a VG Semicon V80H MBE reactor and an RAS system, described in detail elsewhere.[11] The RAS system had a working spectral range from 1.5 to 5.5 eV and measured the difference (Δr) between the anisotropic complex reflectance (r) along the [−110] and [110] optical eigenaxes within the (001) surface crystallographic plane, normalized to the mean reflectance (r):

$$\frac{\Delta r}{\bar{r}} = 2\frac{r[\bar{1}10]-r[110]}{r[\bar{1}10]+r[110]}. \quad (1)$$

Only the real part of the RAS signal was investigated to avoid residual strain effects associated with the viewport. Measurements could be acquired spectroscopically with typical scan times of ~3 min or at fixed energy in which case real time measurements are possible.

Thermal desorption of the surface oxides from the GaAs(001) substrate was performed at ~600 °C under an As_4 flux of $F(As_4)\sim5\times10^{14}$ molecules $cm^{-2}\,s^{-1}$. A ~1 μm undoped GaAs buffer layer was then grown at a sample temperature of ~580 °C before growth was stopped and the sample temperature lowered, so that experiments could be performed in the range 350–530 °C. For the time resolved measurements InAs films were deposited at a rate of ~0.025 ML/s to a thickness of up to 4 ML.

III. RESULTS

A. RHEED and spectroscopic RAS measurements

Upon initiation of InAs growth, RHEED observations showed an initial gradual change from a $c(4\times4)$ pattern to an "asymmetric" (1×3) pattern at ~0.3 ML InAs coverage, where the 3× features were slightly offset towards the half-order position. With increasing coverage up to 0.7 ML the 3× features developed, moving towards their correct positions, i.e., the (1×3) pattern became "symmetric." The asymmetric RHEED pattern has been explained by the coexistence of domains of two reconstructions on the same surface and is similar to that previously observed for submonolayer films of Si on GaAs(001).[12,13] Above 1 ML the pattern gradually changed to (1×1) before becoming spotty, indicative of islanding, at ~1.6 ML. This transition in the RHEED pattern, being very distinct and abrupt even at very low growth temperatures, was used to identify the islanding transition and acted as a marker for the RAS studies. Following termination of growth after 2 ML, redistribution on the surface was evident since the RHEED spots were observed to slowly sharpen with time.

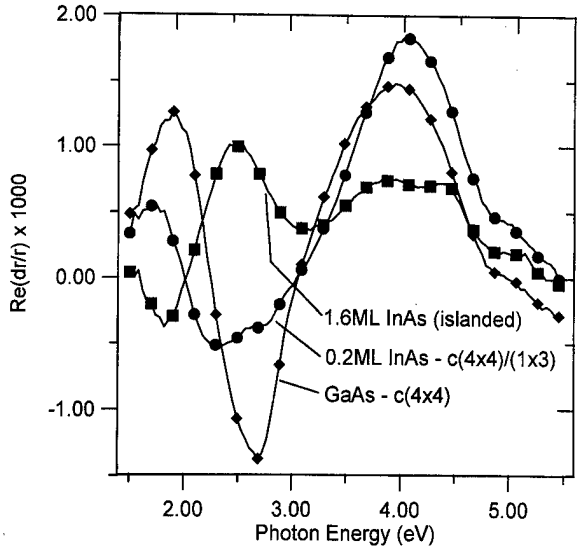


FIG. 1. RAS spectra taken immediately after the growth of different amounts of InAs at a growth temperature of 350 °C: (♦) the original GaAs substrate [$c(4\times4)$]; (●) 0.2 ML InAs [$c(4\times4)/(1\times3)$ mixed reconstruction]; (■) 1.6 ML InAs, i.e., at the point of islanding (spotty RHEED pattern). The symbols are positioned at every 5th data point.

To illustrate the spectroscopic changes in the RAS signal with InAs growth, traces are shown in Fig. 1 for deposition at 350 °C for the clean $c(4\times4)$ GaAs surface and following growth of 0.2 ML [close to the point where the As-rich $c(4\times4)$ reconstruction is finally lost] and 1.6 ML of InAs (i.e., the point of islanding). From Fig. 1 the photon energies of most interest, i.e., where the largest changes in signal occur, can be identified as ~2.6 and ~4.0 eV. Changes in the sign of the signal from negative to positive at 2.6 eV are known to be related to the directional nature of the As–As bonds on the surface. As discussed above, the classic example of this is the change in sign on going from the excess As $c(4\times4)$ to the As stabilized (2×4) reconstructions of GaAs(001), where the reconstruction change involves the loss of the outer layer of As. The initial rise of the 4.0 eV RAS signal with InAs thickness seen in Fig. 1 also mimics that observed for the GaAs(001) $c(4\times4)-(2\times4)$ transition and it may therefore be concluded that, as far as RAS is concerned, the initial changes observed at both energies (i.e., at 0.2 ML InAs coverage) are essentially due to loss of the excess As layer from the surface and the establishment of a surface that is similar to a mixed GaAs(001) $c(4\times4)/(2\times4)$. Such a change could easily have been predicted since the In–As bond strength is lower than that of Ga–As. Hence, although InAs and GaAs(001) display somewhat similar reconstructions, the boundaries between them are shifted to much lower temperature for InAs.[14] It is therefore

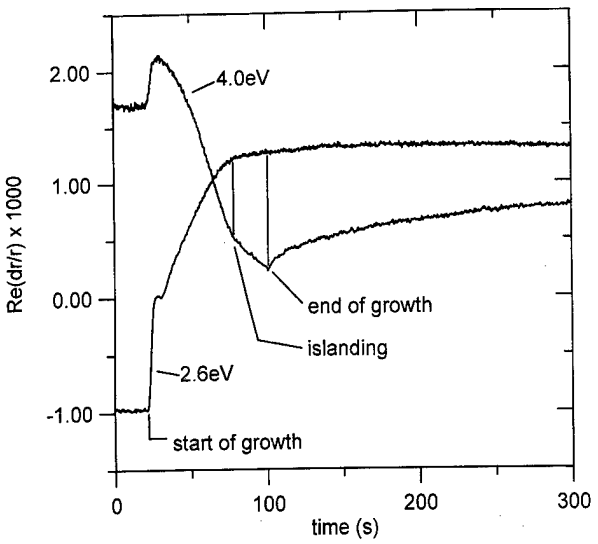


FIG. 2. RAS transients monitored at photon energies of 2.6 and 4.0 eV before, during, and after the deposition of ~2 ML of InAs on GaAs at a rate of ~0.025 ML/s and a growth temperature of 450 °C.

reasonable that the (submonolayer) InAs/GaAs(001) surface is incapable of maintaining an excess As layer and it is the consequence of this that is observed. However, the fact that these surfaces are not identical is obvious; due to the presence of In the InAs/GaAs(001) displays an asymmetric (1 ×3) pattern rather than a (2×4) and differences are also apparent in the RAS, where at 1.6 ML coverage the maxima at ~2.6 and ~4.0 eV are both shifted by ~0.2 eV to lower energy than observed for GaAs surfaces. The most notable distinction is that, after its initial rise at 0.2 ML, at 1.6 ML the 4.0 eV signal decreases significantly below the initial GaAs value. This has been reported previously,[15,16] has no precedent on straightforward GaAs surfaces, and is important to note when it examining the time resolved traces described later.

B. Deposition of 2 ML of InAs at 450 °C

Although the spectroscopic measurements above are extremely useful in setting the scene, it is known that above the islanding thickness the RAS signal changes with time as surface redistribution occurs,[15,16] as RHEED also indicates. At low growth temperatures spectroscopic measurements have indicated time scales of tens of minutes[16] in which case such measurements could reasonably be used to follow the redistribution (it takes ~3 min for a full spectrum to be recorded). However, faster processes occurring at higher sample temperatures or those occurring on shorter time scales at low temperatures could well be missed. In order to properly follow the dynamics of the growth/redistribution process, time resolved measurements (at fixed photon energies) must be performed. Such measurements at 2.6 and 4.0 eV are shown in Fig. 2 for the case of 2 ML deposition at a rate of ~0.025 ML/s and a growth temperature of 450 °C. The total scan time was 400 s with a lead time of ~10 s followed by growth times of 80 s. Figure 2 is useful and important since it displays the characteristic features of virtually all time resolved measurements under different conditions and may be compared with the spectroscopic changes apparent in Fig. 1. Upon initiation of growth both the 2.6 and 4.0 eV RAS signals increased very rapidly before briefly saturating at ~6–12 s (0.15–0.3 ML), corresponding to the loss of the $c(4\times4)$ pattern. The 2.6 eV response then returned to a (less steeply) rising signal once more while the 4.0 eV decreased towards zero. In fact the changes in the 4.0 eV signal are linear with deposited thickness from ~0.8 ML up to the point of islanding (1.6 ML). Thereafter, the main points to note are

(1) the sensitivity of the 2.6 eV signal and relative lack of sensitivity of the 4.0 eV signal to the onset of islanding (the 2.6 eV signal saturates at this point and is largely constant thereafter as has been observed previously[15,16]);
(2) the total absence of any indication of the cessation of growth at 2.6 eV concurrent with a very strong discontinuity at 4.0 eV;
(3) the sensitivity to postgrowth surface changes at 4.0 eV only.

Although detailed and difficult calculations will be necessary to understand these effects properly, it is possible to relate them to the onset of islanding, cessation of growth, and postgrowth development on a phenomenological basis. To do this a number of factors need to be considered. First, the RAS signal in such a system as this may include contributions from the surface order, the thickness of the overlayer, and possibly also the nature of the buried interface.[8,17] Second, interpretation of the data is much more difficult when both a pseudomorphic wetting layer and 3D islands are present. However, just beyond the point of islanding the islands that are present are likely to be small, of low density, and cover only a small fraction of the surface area.[5] These islands are consequently unlikely to contribute significantly to the RAS signal and so can be ignored. Therefore, assuming any interface effects to be already fully established, the constant signal at 2.6 eV immediately following islanding may result either from

(1) a constant (in both surface order and thickness) wetting layer; or
(2) a balance between the effects of a changing surface order and an increasing wetting layer thickness; or
(3) a constant surface contribution and a negligible effect from an increasing wetting layer thickness.

Although none of the above possible explanations can be definitely ruled out (1) requires that all of the incident In flux is directly transferred to the islands and is contradicted by the observation that the 4.0 eV signal does not behave in the same fashion. Since the weight of both experimental and theoretical evidence indicates that ~2.6 eV signals correlate exclusively with the surface As–As bonds (2) is very unlikely, and so (3) is the most likely explanation. With a constant surface order following islanding, the (changing) 4.0 eV signal must be due solely to a still increasing wetting

layer thickness. In addition, the linear relationship between the RAS signal and deposited material between ~0.8 and 1.6 ML suggests that it can be used as a direct measure of the wetting layer thickness. Indeed, a change in slope at the point of islanding is entirely expected since, by definition, some material must then be incorporated into islands. To the best of our knowledge this is the first evidence that the wetting layer can progress and become thicker than the value it has at the point of islanding (i.e., 1.6 ML). Despite the significant amount of study of this subject, it is perhaps not surprising that it has not been previously reported since, as will be shown, it only occurs at temperatures lower than those usually employed for quantum dot growth and because our RAS studies are in effect the first thickness dependent, real time *in situ* measurements.

The possibility of interpreting the 4.0 eV data in a quantitative or semiquantitative way in terms of the wetting layer thickness is the crux of this article. Success here would be very important since knowing the amount of material in the wetting layer can also reveal the amount present in islands and lead to a great step forward in our understanding of the whole islanding process. In the remainder of this article we seek to explore both the validity of our interpretation of the RAS data and the implications it has for our understanding of the quantum dot formation process.

C. Deposition of 2 ML of InAs at different temperatures

Previous work has suggested that the final equilibrium wetting layer thickness has a value of ~1 ML,[5] and our own spectroscopic RAS studies have indicated thicknesses in the range ~0.8–1.2 ML depending on temperature.[16] Therefore the thick wetting layers predicted above must be inherently unstable, leading to the expectation of subsequent redistribution of material to the islands through a kinetically limited process. In addition, it might be expected that the change in slope at the point of islanding might vary. It appears possible to monitor both processes by observation of the 4.0 eV time resolved signal. Experiments to check this were performed in the growth temperature range 350–530 °C using the same growth rate (~0.025 ML/s) and film thickness (~2 ML) as used in Sec. III B. RHEED showed that the reconstruction of the GaAs starting surface was $c(4\times4)$ except for the highest temperature (530 °C) for which it was (2×4). In addition, islanding of the InAs took place after exactly the same growth period with the exception of 530 °C where a longer period was required. It is thought that this delay in islanding was due to the reevaporation of In from the surface at high temperature, a view supported by other measurements at higher temperatures.

Figure 3 shows time resolved scans for the deposition of 2 ML of InAs taken at 4.0 eV. Differences at the onset of growth relate to the initial reconstruction changes [$c(4\times4)$ or $(2\times4)-(1\times3)$] which took place more quickly with increasing growth temperatures. However, it is the point of islanding and the cessation of growth that follows that are of most interest here. At 350 °C there appears to be only a slight inflection at the point of islanding. Thus it appears that

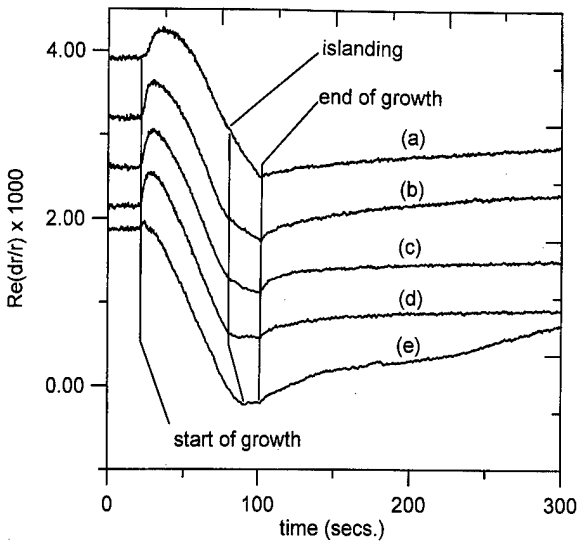


FIG. 3. RAS transients monitored at a photon energy of 4.0 eV before, during, and after the deposition of 2 ML of InAs on GaAs at a rate of ~0.025 ML/s and growth temperatures of (a) 350, (b) 450, (c) 475, (d) 500, and (e) 530 °C.

even beyond the point of islanding most of the incident In flux continues to be incorporated into the continuous interisland film at this growth temperature. At all growth temperatures the RAS response varies linearly from 0.8 ML up to the point of islanding; however at higher growth temperatures its inflection became more pronounced so that at 500 °C and above an essentially flat trace was then obtained until cessation of growth, implying that under these conditions all of the incident flux was incorporated immediately into islands.

It is worth noting that the RAS response varies approximately linearly with time both before and after the point of islanding, the difference is one of slope. This implies that following islanding a certain fixed fraction of the incoming In flux is directly transferred to islands, at least up to 2 ML of deposited material. This fraction depends on the growth temperature so that it appears to be ~30% at 350 °C, ~75% at 450 °C, and 100% at 500 °C and above. The behavior at high growth temperatures provides supporting evidence for the independence of the 4.0 eV signal of the material in the islands. If rather than being independent, some nonlinear thickness averaging was in fact taking place, then such a flat response would be difficult to explain.

Postgrowth redistribution of the surface as a function of growth temperature is slightly more complicated. Although there is clearly a greater rate of recovery of the 4.0 eV signal following growth at 450 as opposed to 350 °C, above this temperature there is very little change. This may be understood by considering that it is reasonable that the rate of redistribution must increase not only with sample temperature but also with strain energy, the latter being related to the wetting layer thickness. Thus, at the lowest temperature since only a small fraction of the In flux initially goes into islands, the wetting layer is thickest. Although the strain energy driving the redistribution is largest at these lower temperatures, the low diffusion rate ensures that the In redistri-

bution rate remains low. As the growth temperature is increased, the enhanced In mobility results in more material being incorporated into islands during the growth process limiting the strain energy available to drive the postgrowth redistribution. Although this argument is rather simplistic, it probably highlights the most important factors operating in this case; a fuller description will be revealed later in the article.

One final comment we want to make is that the postgrowth 4.0 eV, 530 °C trace deviates away from those for lower temperatures after more than ~200 s. This is consistent with a thinning of the InAs wetting layer, which we attribute to desorption of In from the surface: this implies that all the loss of In is initially borne by the InAs islands. This is not a statement that evaporation occurs preferentially at the islands since In transfer between island and wetting layer may occur. However, it does once again provide further evidence of the independence of the 4.0 eV signal of the islands.

D. Deposition of varying thicknesses of InAs at 450 °C

The variable temperature study above provided support for the interpretation of the 4.0 eV RAS signal during the initial stages of growth. Up to 2 ML layer deposition the RAS signal is always approximately linear. However, there must be a limit to this either through the introduction of nonlinearities in the RAS response itself or through, for example, the increasing strain energy driving an enhanced redistribution rate. To address this issue, layers from ~1 to ~4 ML of deposited thickness were grown at a rate of ~0.025 ML/s and a growth temperature of 450 °C.

Figures 4(a) and 4(b) show the resulting 4.0 eV time resolved scans; they are separated into two parts for clarity. It can be seen that the traces coincide up to the point of cessation of growth indicating the reproducibility of the measurements. Hence varying the deposited thickness primarily provides access to the postgrowth redistribution under different conditions. The redistribution during growth is entirely covered by the measurement of the thickest (~4 ML) deposition.

The behavior of the RAS response for the thickest film deposited is remarkable. As the amount of deposited material increases, the 4.0 eV RAS signal, rather than decreasing monotonically, goes through a minimum before eventually beginning to increase. This is not entirely unexpected since the resulting increase in the wetting layer thickness raises the strain energy available to drive redistribution of material to the islands. However, in this simple picture a balance must eventually be reached where the rate of redistribution equals the growth rate resulting in a flat trace or constant wetting layer thickness. Diminution of the wetting layer would only then occur after cessation of growth. Instead, RAS indicates that the wetting layer undergoes thinning during the growth process, implying that the redistribution process is more complex.

Examination of the postgrowth recovery of the various

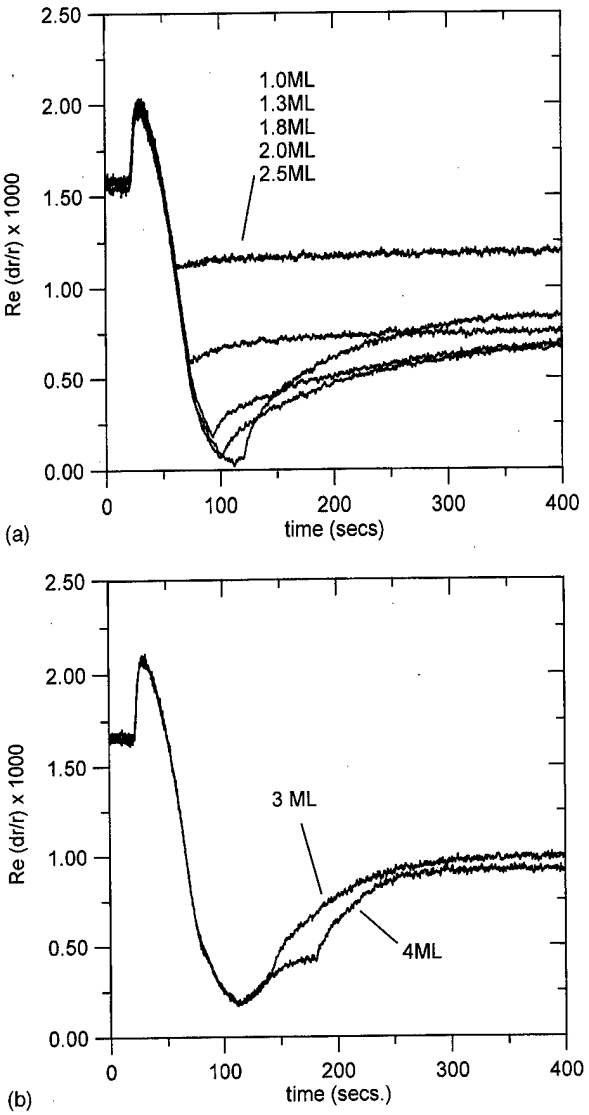


FIG. 4. RAS transients monitored at a photon energy of 4.0 eV before, during, and after the deposition of various thicknesses of InAs on GaAs at a rate of ~0.025 ML/s and a growth temperature of 450 °C. Divergence of the traces occurs at the termination of growth and is indicated. (a) Deposited thicknesses of ~1.0, 1.3, 1.8, 2.0, and 2.5 ML. (b) Deposited thicknesses of 3.0 and 4.0 ML.

amounts of deposited material is also revealing. Although the change for the thinnest (1 ML) film is very small, for the 1.3 ML case there is a clear rise in the signal which appears to saturate after ~100 s; this occurs despite the fact that this film is thinner than the thickness required for islanding. Perhaps the most important observation from this data set is the difference between the 2.0 and 2.5 ML cases. Despite the fact that growth is terminated at almost the same RAS signal and therefore the same wetting layer thickness, the recovery of the latter is much more rapid, quickly surpassing that of the slightly thinner film. This observation confirms the complexity of the redistribution process and leads to the comment that the most obvious change occurring as growth proceeds is that more material is present in the form of islands.

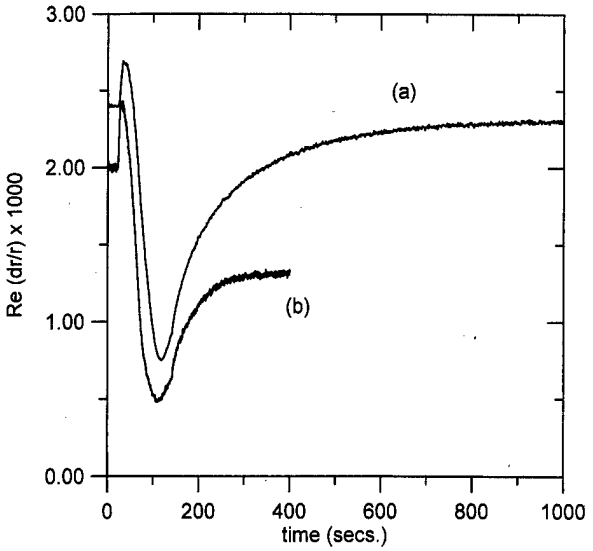


FIG. 5. RAS transients monitored at a photon energy of 4.0 eV before, during, and after the deposition of ~3 ML of InAs on GaAs at a rate of ~0.025 ML/s and growth temperatures of (a) 350 and (b) 450 °C.

It has previously been suggested that these can act as sinks for material from the wetting layer![4,5]

Finally, it is clear that the recovery of the three "thickest" films has reached equilibrium, leading to a saturation wetting layer thickness of ~1.35 ML in each case. Although the 1.8 and 2.0 ML films have lower RAS signal values, it is also clear that these surfaces are still evolving even ~300 s after termination of growth. This clearly suggests that the recovery process proceeds more quickly when more material is present in islands.

E. Saturation in the recovery of 3 ML InAs films at different temperatures

The previous discussions have hinted at aspects of the behavior of the InAs growth on GaAs(001) but would benefit from more clearcut data. For example, in the 2 ML deposition, variable temperature study it is clear, by examining the postgrowth behavior, that the signals do not convincingly saturate at any temperature within the time scales considered. At low temperature this is entirely consistent with our previous spectroscopic study[16] where saturation was observed but it took well over 1 h to achieve. By contrast, at high temperatures Fig. 3 demonstrates that reevaporation of InAs is significant and so saturation can never occur here. Finally, Sec. III D contained evidence that the recovery time is a function of the amount of material deposited.

To demonstrate these effects more clearly 3 ML of material was deposited at a rate of ~0.025 ML/s at temperatures of 350 and 450 °C and the resulting 4.0 eV RAS scans are shown in Fig. 5. (the two traces are offset for clarity). The most striking feature of these traces is the difference in the extent of the recovery of the RAS signals for the two temperatures. Following the thrust of this article it can be estimated that the final equilibrium wetting layer thicknesses are ~0.8 and ~1.35 ML for 350 and 450 °C, respectively, in agreement with both our previously published results[16] and with those given in Sec. III D. It is also clear that following growth at 350 and 450 °C the times required to reach saturation are ~600 and ~100 s, respectively. Thus the time scale is much shorter at the higher temperature and, remembering that more than 1 h was required for 2 ML at 350 °C, it is also much quicker with a greater deposited thickness. A final point to be noted here is that, as has already been shown for 450 °C depositions, the RAS signal at 350 °C also goes through a minimum and has a positive gradient at the termination of 3 ML growth.

IV. SIMULATING THE RAS DATA/WETTING LAYER THICKNESS

We have already indicated that the RAS signal at 4.0 eV can be used, at the very least, as a guide to the thickness of the wetting layer that exists between InAs quantum dots/islands and therefore that it provides a useful tool for real time monitoring of the formation of InAs quantum dots on GaAs(001). Some phenomenological interpretations of the islanding process consistent with the data were proposed and these appear to be quite straightforward, suggesting that it is not a difficult task to model the essential features of the growth/redistribution process mathematically.

The following model is a simple simulation of the thickness of the InAs wetting layer through the deposition and postgrowth rearrangement processes. It is not intended to be an accurate representation of the physical processes that occur and no attempt will be made to "fit" the RAS data at this point, but it does serve to illustrate some of the issues that a more realistic model should address. It should also be remembered that comparisons between data and simulations rest on the assumption that the 4.0 eV RAS data do indeed reflect the wetting layer thickness above InAs coverages of ~0.8 ML. In essence this means that the linear response approaching the point of islanding acts as the calibration by which to measure the postislanding behavior.

Changes in the wetting layer thickness arise from growth due to the incident In and As_4 fluxes as well as to loss of material to islands which are assumed to be effectively invisible to the RAS measurement at this photon energy. The growth rate (g_r) used is the same as the experimental value of ~0.025 ML/s, which for 2 ML deposition corresponds to a growth time of 80 s.

The change in wetting layer thickness in 1 s, $\delta(WL)$, given by

$$\delta(WL) = g_r \times (1-f) - A \times \text{sink} \times (WL - 1.6) - B \times \text{sink} \times [WL - WL(0)], \quad (2)$$

where all terms are thicknesses expressed in monolayers.

In the first term, f is the fraction of incident flux transferred directly into islands when the thickness of the wetting layer is greater than 1.6 ML, the critical thickness for islanding (i.e., f is 0 for $WL < 1.6$ ML, and so nonzero values for f only occur during the islanding process). Estimated values for f as a function of growth temperature have already been given for the case of a growth rate of ~0.025 ML/s.

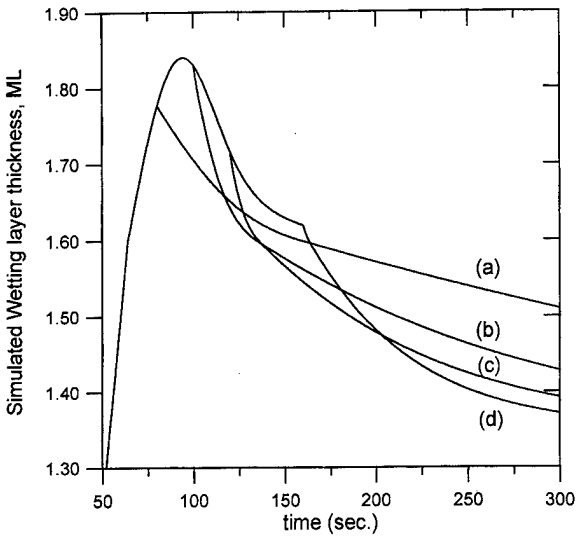


FIG. 6. Simulation, using Eq. (2) of the wetting layer thickness as a function of time, with the parameters: $g_r = 0.025$ ML s^{-1}, $f = 0.5$, $A = 0.08$ ML^{-1}, $B = 0.007$ ML^{-1} and $WL(0) = 1.35$ ML. Here growth is started at $t = 0$ s and terminated at (a) 2, (b) 2.5, (c) 3, (d) 4 ML.

In the second and third terms A and B are scaling factors and the quantity "sink" represents the total cumulative amount of material transferred to islands, i.e., total deposited thickness = sink + WL (in the absence of reevaporation). Its inclusion provides a mechanism for the efficient removal of material from the wetting layer by the islands and such a term is essential in order to account for the positive going RAS signals, for example, above $d = \sim 2$ ML at $Tg = 450$ °C. A precedent does exist for this since both Joyce *et al.*[5] and Ramachandran *et al.*[4] have previously introduced the requirement of a "sinking mechanism" for island formation and explained the mechanism in terms of an anisotropy in the probability of In atoms going up or down steps. Obviously for island formation the greater probability must be for going up a step. Two terms were found to be necessary to account for the fast recovery to the 1.6 ML signal level and the subsequent, relatively much slower, recovery. In the $[WL - WL(0)]$ factor, $WL(0)$ is the final equilibrium wetting layer thickness, which has already been noted to vary as a function of sample temperature.

Bearing in mind the inverse relationship between the RAS signal and wetting layer thickness, Fig. 6 shows simulations that mimic the variable thickness depositions at 350 and 450 °C in Figs. 4 and 5. It can be seen that the simulation reproduces the essential features of the data very well including crossover of the 2.5 and 2.0 ML traces in Fig. 4(a) following cessation of growth and the peculiar shape of the 4 ML trace during growth.

The above simulations make two important points. First, they show that, despite the initially rather speculative explanation of the form of the 4.0 eV RAS signal, it has proved consistent with data. In addition, the simplicity of the model leaves the important mechanisms of relaxation quite clear. At higher temperatures, where the value of the fraction of material directly transferred to the islands is large (i.e., f is ~ 1), it is this that acts as the important initial strain relaxation mechanism. Postgrowth relaxation is also not very significant here since the equilibrium wetting layer thickness is close to 1.6 ML. By contrast at low temperatures where direct transfer is small and the equilibrium wetting layer thickness, $WL(0)$, is ~ 1 ML the second term, scaled to 1.6 ML, becomes important as the amount of material in the islands (or the sink) increases and the third term, scaled to $WL(0)$ provides the final, slower recovery.

V. DISCUSSION

Although the results presented here have revealed the strength of the RAS technique for following in real time the movement of InAs material about the GaAs(001) surface, it is not clear why this should be so. In particular, questions that may be asked include (1) Why is the relationship between the 4.0 eV RAS signal and wetting layer thickness apparently so linear? (2) Why is the RAS signal insensitive to the presence of islands? (3) Why can the wetting layer only be followed at high photon energies (in our case 4.0 eV)? (4) What are the limitations to the application of this method? The answers to these questions are probably inseparable and are certainly largely unknown at this time. However it is certainly worth discussing them within the framework of our work and that of others.

In addressing the first question, it is worth noting that the argument for believing that the islands formed on the surface are invisible to the 4.0 eV signal applies equally well to other photon energies. So, a better way to phrase this question might be to ask, Why it is that only high photon energies are sensitive to InAs wetting layer thickness? And in particular, Why it is that the 4.0 eV signal decreases by the amount that it does? The best place to start is to compare the spectra obtained from the surfaces of bulk GaAs and InAs(001) to assess the possible effects due to a straightforward transition from one to the other. Although there does not seem to be a large data set available for InAs(001), Steimetz *et al.*[18] have shown that the features for InAs are very similar to those for GaAs, the main difference being a shift to lower energies by 0.3–0.4 eV. This is consistent with previous experience with group III arsenides, AlGaAs features having been observed to be shifted to slightly higher energies with respect to GaAs.[19] The InAs spectrum shows a high energy peak at ~ 3.6 eV and a significant drop off in signal at 4.0 eV, seemingly consistent with the changes seen here after ~ 0.5 ML of growth. However, this is not the complete story since the spectroscopic information presented here and a more thorough previous study[16] show that the GaAs ~ 2.65 eV feature shifts only by ~ 0.15–~ 2.5 eV, well short of the final 2.3 eV bulk position. Therefore a similar small shift of the higher energy feature would result in a relatively small decrease in signal at 4.0 eV and so it seems that the data cannot be explained in terms of a simple transition from a GaAs to InAs spectrum. This is perhaps not surprising since unusual effects in high energy (4.0–5.5 eV) RAS features have previously been observed for the case of at least the first 25 ML

deposition of AlAs on GaAs(001) where the very small lattice mismatch results in planar growth throughout.[19]

Considering the RAS signal at ~2.6 eV, it has long been established that this is sensitive to As–As bonds, in particular their direction and number, leading to the view expressed in Sec. III B that the saturation at the point of islanding is related to the surface order of the wetting layer (in terms of these bonds) reaching a constant level, despite the fact that its thickness continues to increase. It is rather curious that saturation occurs at the same thickness at which islanding is initiated and intriguing to consider whether this is coincidental or not. In any case it is probable that it is this high sensitivity to surface bonding that results in lack of sensitivity to wetting layer thickness. Since higher energy signals are less sensitive to surface bonds, they are intrinsically more likely to respond to further thickness changes. However, the initial changes here do indicate some dependence on surface bonding and raise another note of caution concerning the interpretation of the 4.0 eV data. In the regime 0.8–1.6 ML, the 2.6 eV signal changes, implying a change in surface bonding. It is therefore likely that the 4.0 eV signal will also be influenced to some extent and so alter the relationship between the RAS signal and wetting layer thickness before and after islanding, and affect any derived wetting layer thicknesses.

With no explanation for the origin of the changes observed in the 4.0 eV signal it is clear that there is, similarly, no explanation at present for the apparent linearity of the RAS response with wetting layer thickness. This makes it difficult to address the limitations to this linearity and its independence of the islands. However, some information can be inferred from the behavior of the signal and in particular its comparison with simulations, which seem to indicate that even with the deposition of the equivalent of 4 ML (at 450 °C) of InAs the islands have no significant effect. This is also directly evident in Fig. 5 where the signals from different thicknesses of InAs from 2 to 4 ML seem to converge to much the same level. This is a rather unexpected result since at 4 ML the onset of island coalescence is expected[5] and therefore the initial argument that at the onset of islanding the islands cover only a small fraction of the surface certainly does not apply. This subject is quite complex since it is not clear what signal to expect from the islands themselves. If the dimensions of the islands are large enough and are anisotropic in the plane of the surface, then some contribution to the RAS signal might be expected from them.[18,20] However, if they are not, then their stepped or faceted nature is likely to reduce their RAS response compared to the two-dimensional wetting layer. It is this latter case that seems most appropriate to the measurements presented here. In addition, as the islands cover more of the surface, then the response from the wetting layer must necessarily decrease. Consideration of this is complicated by the fact that the equilibrium thickness of the wetting layer varies with temperature so that the amount of material in the islands is a complex function of deposited thickness, temperature, and time.

At this point it is worth discussing the main points arising from this work in terms of what is already known about this system from STM studies. It has already been stated that such studies have introduced the requirement of a "sinking mechanism" for island formation.[4,5] At the same time these also revealed the existence of 2D and quasi-three-dimensional (3D) islands in addition to a uniform wetting layer and large, conventional islands. RAS is obviously not capable of distinguishing these features, although assuming that the linear change in the RAS signal with wetting layer thickness holds, then the RAS signal accurately represents an average InAs thickness. Therefore this analysis has no problems with 2D islands, they simply contribute to the wetting layer signal. What is not so clear is what the effect of the quasi-3D islands might be since no criteria have been established for when islands become invisible to RAS.

By contrast with the behavior of islands as sinks, there has been no previous evidence either that the wetting layer can have a thickness greater than 1.6 ML, or that the equilibrium between the wetting layer and the islands is temperature dependent. It is probably not surprising that STM studies have failed to pick these points up since most are performed at room temperature after the dots have been grown, quenched, and removed from the growth environment. At 450 °C and with thin deposited films it takes only of the order of 30 s for the wetting layer to decrease below 1.6 ML and, following growth at low temperatures, where the redistribution processes are slowed significantly, STM images unfortunately tend to be quite poor.[21] The situation is made more difficult because of the fact that as the sample is cooled the equilibrium wetting layer thickness decreases. Hence when films are deposited at high temperatures, they will tend towards the lower temperature equilibrium and many observations may be artifacts of the quenching process. A variable temperature study of quantum dot size suggests itself as a good test of this result provided desorption of material is kept under control. However, room temperature STM has been successful in revealing the reduction of the wetting layer below 1.6 ML,[5] in agreement with and providing support for this and our previous spectroscopic study.[16]

The obvious alternative test of the observations made here is comparison with information from studies of "real" quantum dots, i.e., those overgrown, for example, with GaAs, and examined by photoluminescence, transmission electron microscopy, etc. However the overgrowth process itself is fraught with complications since it is well known that In segregates on the surface of GaAs during growth.[22] The complete removal of quantum dots has even been found to result from such processes.[5]

One last point to consider is the reason why the equilibrium wetting layer thickness should increase with sample temperature. As already stated, it has been shown that the transfer of material from the wetting layer is controlled by an anisotropy in energetic barriers up and down the steps associated with the islands. This anisotropy will be more important at low temperatures since the surface atoms have lower energy. At higher temperatures the surface atoms have higher energy and so are less likely to be trapped. Hence it is

entirely reasonable that a thicker wetting layer exists at higher temperatures.

Apart from the very new predictions of the InAs/GaAs(001) growth mechanism, the ability to follow the redistribution of material in real time will be invaluable in determining the important temperature and strain dependent diffusion rates and activation energies important to a real understanding of the physics of the islanding process and ultimately its control. Such parameters have previously been derived from room temperature STM studies which have been shown to be ideally suited to this purpose. However, STM and RAS should be viewed as complementary since STM reveals information on the detailed morphology of the wetting layer in addition to island sizes and densities, which are inaccessible to RAS.

VI. CONCLUSIONS

RAS has been applied to the *in situ* study of islanded films of the prototypical quantum dot system, InAs on GaAs(001), and has revealed information not available by use of the more conventional techniques of STM and RHEED. The radically different RAS responses at different photon energies are remarkable. It appears that the 4.0 eV response above a coverage of ~0.8 ML can apparently be used as a measure of the thickness of the interisland two-dimensional wetting layer since the material present in islands becomes entirely or largely invisible to the technique. In the absence of desorption this allows the distribution of InAs on the GaAs surface and the interplay between the continuous wetting layer and the islands to be followed. Based on this the following conclusions are drawn: (1) during growth and immediately beyond the point of islanding, a fixed fraction of the incoming flux is directly transferred to the islands; this fraction increases with growth temperature up to 100%; (2) as a consequence of this, the wetting layer can increase in thickness beyond 1.6 ML and the islanding thickness to at least ~2 ML; (3) the importance of islands acting as sinks for the acquisition of material from the wetting layer is revealed; (4) the dynamic equilibrium set up between the islands and the wetting layer is such that the latter's thickness increases significantly with sample temperature. In addition, it should be possible to model the simulations to extract some of the parameters necessary in order to understand the physics of the islanding process.

ACKNOWLEDGMENTS

This work was supported by the U.K. EPSRC Low Dimensional Structures and Devices Initiative and by the DAAD and the European H.C.M. GOODS (Growth of One Dimensional Structures) programme. The authors would like to thank T. S. Jones and J. G. Belk (Imperial College, London) and J. E. Macdonald (University of Wales, Cardiff) for useful discussions.

[1]F. Houzay, C. Guille, J. M. Moison, P. Henoc, and F. Barthe, J. Cryst. Growth **81**, 67 (1987).
[2]O. Brandt, K. Ploog, L. Tapfer, M. Hohenstein, R. Bierwolf, and F. Phillipp, Phys. Rev. B **45**, 8443 (1992).
[3]J.-Y. Marzin, J.-M. Gérard, A. Izraël, D. Barrier, and G. Bastard, Phys. Rev. Lett. **73**, 716 (1994).
[4]T. R. Ramachandran, R. Heitz, P. Chen, and A. Madhukar, Appl. Phys. Lett. **70**, 640 (1997).
[5]B. A. Joyce, J. L. Sudijono, J. G. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedensky, and T. S. Jones, Jpn. J. Appl. Phys., Part 1 **36**, 4111 (1997).
[6]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. B **6**, 1127 (1988).
[7]Z. Sobiesierski, D. I. Westwood, and C. C. Matthai, J. Phys.: Condens. Matter **10**, 1 (1998).
[8]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. B. Harbison, Phys. Rev. B **46**, 15 894 (1992).
[9]J. M. Bass and C. C. Matthai, J. Vac. Sci. Technol. B **14**, 3075 (1996).
[10]H. Anyele and C. C. Matthai, J. Phys.: Condens. Matter **8**, 6585 (1996).
[11]S. M. Scholz, A. B. Müller, W. Richter, D. R. T. Zahn, D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Vac. Sci. Technol. B **10**, 1710 (1992).
[12]M. R. Fahy, M. J. Ashwin, J. J. Harris, R. C. Newman, and B. A. Joyce, Appl. Phys. Lett. **61**, 1805 (1992).
[13]A. H. Levermann, D. A. Woolf, D. I. Westwood, and J. E. Macdonald, Surf. Sci. **352–354**, 352 (1996).
[14]S. M. Newstead, R. A. A. Kubiak, and E. H. C. Parker, J. Cryst. Growth **81**, 49 (1987).
[15]D. I. Westwood, Z. Sobiesierski, E. Steimetz, TR. Zettler, and W. Richter, Appl. Surf. Sci. **123/124**, 347 (1998).
[16]E. Steimetz, J.-T. Zettler, W. Richter, D. I. Westwood, D. A. Woolf, and Z. Sobiesierski, J. Vac. Sci. Technol. B **14**, 3058 (1996).
[17]T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, J. Vac. Sci. Technol. B **14**, 3052 (1996).
[18]E. Steimetz, J.-T. Zettler, F. Schienle, T. Trepk, T. Wethkamp, W. Richter, and I. Sieber, Appl. Surf. Sci. **107**, 203 (1996).
[19]S. J. Morris, J. T. Zettler, K. C. Rose, D. I. Westwood, D. A. Woolf, R. H. Williams, and W. Richter, J. Appl. Phys. **77**, 3115 (1995).
[20]E. Steimetz, F. Schienle, J.-T. Zettler, and W. Richter, J. Cryst. Growth **170**, 208 (1997).
[21]T. S. Jones (private communication).
[22]J.-M. Gérard and C. d'Anterroches, J. Cryst. Growth **150**, 467 (1995).

Photon-induced localization and final-state correlation effects in optically absorbing materials

D. E. Aspnes,[a] L. Mantese, and K. A. Bell
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

U. Rossow
Institut für Physik, TU-Ilmenau, D-98484 Ilmenau, Germany

(Received 21 January 1998; accepted 20 May 1998)

Two consequences of the absorption of light in optically absorbing materials that appear not to have been recognized previously are: (1) localization of the final electron and hole states involved in the absorption process into wave packets and (2) propagation of these wave packets with their respective group velocities. We demonstrate the existence of these phenomena by applying first-order time-dependent perturbation theory to a simple model that can be solved analytically even when correlations that are ordinarily discarded in the random phase approximation are retained. This approach provides a natural explanation of components in surface- and interface-optical spectra that are related to energy derivatives of the bulk dielectric function ϵ_b and apparent differences in nominally bulk critical point energies E_g and broadening parameters Γ depending on surface conditions. © *1998 American Vacuum Society.* [S0734-211X(98)08404-2]

I. INTRODUCTION

The accurate calculation of surface-optical spectra remains a difficult theoretical challenge. General treatments of the optical responses of materials applicable to the surface-optical problem have been formulated,[1–4] but these are arcane and typically require extensive numerical evaluation, which is not conducive toward gaining a fundamental understanding of basic mechanisms. Here, we discuss a simple model that provides new insight into the optical absorption process in general and the interpretation of surface-optical spectra in particular because it allows us to account explicitly for the exponential decay of the field amplitude of the incident photon. In our model, the resulting final-state correlations, which are discarded in standard random-phase-approximation (RPA) calculations, can be evaluated analytically to first order in time-dependent perturbation theory. This leads to a significantly different picture of optical absorption with respect to that obtained conventionally, where uniform permeation of the material by the photon is assumed.[4]

Specifically, we show that the removal of translational invariance in the propagation direction of the incident photon gives rise to the following: (1) localization of the excited electron and hole, which are wave packets whose spatial profiles initially map exactly onto that of the photon field, as expected from energy-transfer arguments; (2) nonlocal polarization, since these packets are only stable dynamically and must propagate with a group velocity $v_g=\partial\omega(k)/\partial k$ normal to the surface; (3) an unavoidable influence of surface and/or interface effects on apparent values of bulk critical point energies E_g and broadening parameters Γ deduced from optical data, which vary according to surface conditions; (4) as a result, a natural explanation for the appearance of components related to energy derivatives of the bulk dielectric function $\epsilon_b=\epsilon_{b1}+i\epsilon_{b2}$ in surface- and interface-optical spectra; and (5) broadening, which in this formulation is related directly to the finite lifetime of the excited state instead of being described as a phenomenological parameter Γ. Our results also suggest that slab calculations, where limited spatial extent is an unwelcome consequence of limited computing power, may actually provide a more accurate representation of surface-related phenomena than calculations that take the entire bulk into account.

We have previously noted that energy-derivative features observed in surface- and interface-optical spectra[5–8] as well as variations of apparent values of E_g in ellipsometric data[9] are inconsistent with standard RPA calculations of the dielectric response of solids,[10] which are based on semi-infinite Bloch waves. An example[7] of such energy-derivative-type line shapes is shown in Fig. 1. Since the energies of the wave functions from which the E_g are determined are established by their semi-infinite tails into the bulk, these energies cannot be influenced by surface effects, which at most can only change the amplitudes of these functions near the surface as in the case of a surface resonance. Therefore, these features were shown to give direct evidence of localization. Furthermore, since spatial attenuation of a Bloch function is not possible in the static case, this must be a dynamic effect caused by the finite penetration depth of light.[5–7]

The present work provides an analytic validation of these ideas. We use a simple model Hamiltonian consisting of the usual kinetic energy term, a periodic crystal potential $V(x,y)$ in the two dimensions parallel to the surface, and an infinitely deep quantum well extending from $-a\leq z\leq a$ normal to the surface. The eigenfunctions parallel to the surface are assumed to be normalized Bloch functions $\phi_m(x,y)$ with eigenenergies $E_m=\hbar\omega_m$, where m is assumed to represent both band index and wave vector. The eigenfunctions perpendicular to the surface are $\cos(k_n z)/\sqrt{a}$ and $\sin(k_n z)/\sqrt{a}$ for odd and even n, respectively, where $k_n=\pi n/(2a)$. The ei-

[a]Author to whom correspondence should be addressed; electronic mail: aspnes@unity.ncsu.edu

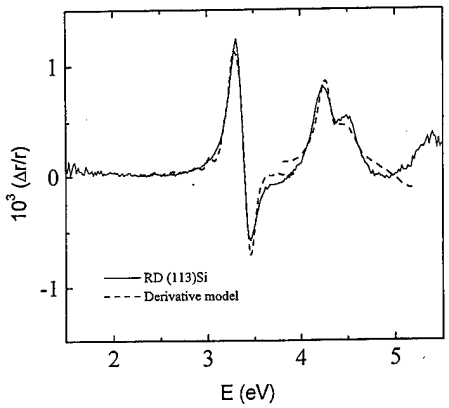


FIG. 1. Solid line: normal incidence RD spectrum of a suboxide-terminated (113)Si surface. (Dashed line) best-fit representations using first- and second-energy derivative spectra of the bulk dielectric function as described in Ref. 7. The (E_0',E_1) complex and E_2 regions were fit separately.

genvalues are $\pi^2n^2\hbar^2/(8m_ea^2)$, where m_e is the free electron mass. The photon is assumed to be turned on at $t=0$ and to decay exponentially in both directions from $z=0$, thus simulating two back-to-back crystals contacted at $z=0$. The empty-lattice nature of the potential normal to the surface allows us to use a harmonic-function representation of the wave functions in this direction and therefore to obtain analytic first-order solutions for the final state without sacrificing any of the essential physics.

Final-state correlation effects have been considered previously,[1–3] however until now only those associated with the outermost atomic layer, e.g., $\mathbf{p}\cdot\mathbf{A}$ effects that appear for p-polarized light[1,2] and boundary-condition effects on the final-state wave functions.[3] In particular, Ref. 2 makes the point explicitly that the classical model for bulk optical properties is microscopically correct. Also, Ref. 2 does not account for Γ, leaving its consideration as a challenge for the reader. For absorbing materials we show that this classical model is not complete, with ramifications including not only apparent values of threshold energies, but also the proper interpretation of the phenomenological broadening parameter. As the analytic solution is somewhat complicated, containing nonvertical terms in momentum and nonconservative terms in energy, we concentrate here on the special case of energy-conserving transitions, deferring a more complete discussion to a later paper.

II. THEORY

We first demonstrate that a surface potential cannot modify the energies of extended (Bloch) states under static conditions. We add a surface potential $V_sd\,\delta(z)$ to the model Hamiltonian, thereby simulating two back-to-back crystals each of thickness a and surface potential $V_sd\,\delta(z)/2$. Since a is of the order of laboratory dimensions and d of Å, representation of the surface potential by a δ function is reasonable. The sine wave functions are unaffected by this addition, while the cosine wave functions develop a first-derivative discontinuity (cusp) at $z=0$, a general characteristic of one-dimensional Green function problems of this type. This change of slope is equivalent to a change of the effective width of the well, and to lowest order in d/a the eigenenergies become

$$E_{mn}=E_m+\frac{n^2\pi^2\hbar^2}{8m_ea^2}+\frac{V_sd}{a}. \tag{1}$$

Thus as $a\to\infty$ the effect of V_s on the energies of these wave functions vanishes as well. The observation of *different* values of E_g for *different* surface conditions for the *same* critical point in optical spectroscopy shows therefore that a must be finite, i.e., the actual wave functions of the crystal must be localized near the surface even though the surface potential may be too weak to create a bound state.[7]

We now ask how this localization occurs. The unperturbed Hamiltonian contains no mechanism for localization, so by analogy to standard problems in electromagnetic theory we consider the photon source, which introduces a characteristic length or penetration depth $L=1/\kappa=1/(2\alpha)$ due to attenuation by optical absorption, where κ and α are the extinction coefficient and absorption coefficient, respectively. We replace the surface-potential term in the Hamiltonian with

$$H'(\mathbf{r},t)=-qEx\,e^{-\kappa|z|}\cos(\omega t)u(t), \tag{2}$$

where $u(t)$ is the unit step function. This represents a photon of linear polarization $\hat{x}$ and extinction coefficient κ arriving at the surface at $t=0$. Although in principle the photon field must be obtained self-consistently,[1,2] for our purposes the present approximation is adequate. Since the magnetic field $\mathbf{H}(\mathbf{r},t)$ and therefore the normal-incidence reflectance R follow from $\mathbf{E}(\mathbf{r},t)$, in this approximation R is not affected by the evolution of the excited state. We write the wave function $\Psi_{mn}(\mathbf{r},t)$ that evolves from the initial state $\phi_m[\cos(k_nz)/\sqrt{a}]e^{-i(\omega_m+\omega_n)t}$ under the action of the photon field as

$$\Psi_{mn}(\mathbf{r},t)=\phi_m(x,y)\frac{\cos(k_nz)}{\sqrt{a}}e^{-i(\omega_m+\omega_n)t}+\sum_{m',n'}a_{mn}^{m'n'}\phi_{m'}(x,y)\frac{\cos(k_{n'}z)}{\sqrt{a}}e^{-i(\omega_{m'}+\omega_{n'})t}. \tag{3}$$

Its evolution with time is given by solving Schrödinger's equation $(H_0+H')\Psi=i\hbar\partial\Psi/\partial t$. To first order in H' the $a_{mn}^{m'n'}$ are given by[11]

$$\frac{\partial a_{mn}^{m'n'}}{\partial t}=\frac{iqEx_{m'm}}{a\hbar}\left(\int_{-a}^{a}dz\,\cos(k_{n'}z)\cos(k_nz)e^{-\kappa|z|}\right)\times e^{i(\omega_{m'n'}-\omega_{mn})}\cos(\omega t)u(t), \tag{4}$$

where $\omega_{mn}=\omega_m+\omega_n$ and $\omega_{m'n'}=\omega_{m'}+\omega_{n'}$ are the eigenenergies of the initial and final states, respectively, in frequency units, and $x_{m'm}=\langle\phi_{m'}|x|\phi_m\rangle$. We assume momentum conservation parallel to the surface, in which case the sum over m' in Eq. (3) reduces to a single term.

Assuming that $L \ll a$ as $a \to \infty$ the integral over z is easily evaluated and we find

$$\int_{-\infty}^{\infty} dz\, \cos(k_{n'}z)\cos(k_n z)e^{-\kappa|z|} = \frac{L}{1+(k_{n'}+k_n)^2L^2} + \frac{L}{1+(k_{n'}-k_n)^2L^2}. \tag{5a}$$

For typical values of L only the resonant term contributes. In a true empty-lattice model this would require $k_{n'}=k_n$, i.e., the initial and final energies would be the same. Accordingly, we suppose a reduced-zone scheme and set $k_{n'}=k_n+\Delta k$, where Δk represents the deviation of $k_{n'}$ from a vertical transition in the conventional band picture. We suppose further that the nonresonant term can be ignored. In principle this assumption is not necessary, but it simplifies the mathematical result. Then

$$\int_{-a}^{a} dz\, \cos(k_{n'}z)\cos(k_n z)e^{-|z|/L} \approx \frac{L}{1+\Delta k^2L^2}. \tag{5b}$$

More generally, the result of the procedure leading to Eq. (5b) is the reciprocal-space representation of the real-space attenuation profile of the photon, even if it is more complicated than the simple exponential assumed here.

We now perform the time integration to obtain

$$\Delta\Psi_{mn} = -\frac{qEx_{m'm}L}{2a\hbar}\,\phi_{m'}(x,y)\sum_{n'}\frac{\cos(k_{n'}z)}{\sqrt{a}(1+\Delta k^2L^2)} \times\left(\frac{e^{-i\omega_{m'n'}t}-e^{-i(\omega_{mn}+\omega)t}}{\omega_{m'n'}-\omega_{mn}-\omega} + \frac{e^{-i\omega_{m'n'}t}-e^{-i(\omega_{mn}-\omega)t}}{\omega_{m'n'}-\omega_{mn}+\omega}\right). \tag{6}$$

We convert the sum to an integral using the density of states π/a, which follows by noting that the allowed values of k_n are $n\pi/a$, where n is odd. Also, we linearize the k dependence of the final-state energy to $\omega_{n'}(k)=\omega_{n'o}+v_g\Delta k$, where $\omega_{n'o}$ is the value of the conduction band energy for $k_{n'}=k_n$. Since $(\omega_{m'n'}-\omega_{mn})$ is not necessarily equal to the photon energy ω, we also define an energy difference $\omega_{m'n'}-\omega_{mn}-\omega=\Delta\omega+v_g\Delta k$. Finally, we discard the second (nonresonant) term in brackets, although again this is not an essential requirement but done to simplify the analytic result. With these considerations $\Delta\Psi_{mn}$ becomes

$$\Delta\Psi_{mn} = -\frac{qEx_{m'm}L}{2\pi\hbar\sqrt{a}}\,\phi_{m'}(x,y)e^{-i(\omega_{m'}+\omega_{n'o}-\Delta\omega)t} \times\int\frac{d\Delta k\,\cos(k_n z+\Delta k z)}{(v_g\Delta k+\Delta\omega)(1+\Delta k^2L^2)}\,(e^{-i(v_g\Delta k+\Delta\omega)t}-1). \tag{7}$$

The integrand contains three simple poles and converges at infinity, so it can be evaluated analytically without approximation by standard contour-integration procedures. The pole involving $\Delta\omega$ expresses energy conservation, and the two involving L express momentum conservation in the direction normal to the surface.

Thus Eq. (7) directly exhibits our first result, that when correlations ordinarily discarded in the RPA are taken into account and the penetration depth of the photon is finite, an excited-state wave packet results. For arbitrarily small times the exponential can be expanded, thereby cancelling the factors $(\Delta\omega+v_g\Delta k)$ and leaving the spatial Fourier transform of the photon field multiplied by $(-it)$. Thus initially the amplitude of the excited state maps exactly onto the photon field independent of the actual spatial distribution of the photon. If $v_g=0$ the mapping remains exact for arbitrarily long times. This is exactly what one would expect from energy transfer arguments: unless the excited state can propagate, the energy that the light beam deposits in the material, which is proportional to the square of the photon field, must simply accumulate as an increase of the final-state probability distribution function $|\Delta\Psi_{mn}|^2$.

As is well known, a wave packet is only stable dynamically and must move with the group velocity v_g. We investigate the consequences of this motion by considering a somewhat simplified version of Eq. (7), that for which $\Delta\omega=0$, i.e., transitions where energy is conserved in the excitation process in the limit of infinite time. Although the full expression is needed to discuss broadening, this simplification allows propagation to be highlighted. Setting $\Delta\omega=0$ and considering only the imaginary (absorptive) part of Eq. (7) yields

$$\Delta\Psi_{mn} = \frac{iqEx_{m'm}L}{2\hbar v_g} \times\left(\phi_{m'}(x,y)\,\frac{\cos k_n z}{\sqrt{a}}\,e^{-i(\omega_{m'}+\omega_{n'o})t}\right) f(z,t), \tag{8a}$$

where

$$f(z,t) = 1-e^{-\kappa|z_g|}\cosh(\kappa z) \quad \text{for } |z|<|z_g|, \tag{8b}$$

$$= e^{-\kappa|z|}\sinh(\kappa|z_g|) \quad \text{for } |z|>|z_g|, \tag{8c}$$

where $z_g=v_g t$ is the distance that a point on the wave front starting at $z=0$ propagates in the time t. This expression has the form of the resonant final-state wave function multipled by a position- and time-dependent envelope function $f(z,t)$ that describes how the wave function is localized and how it evolves with time. Thus Eqs. 8(a)–8(c) illustrate our second result, that the wave packets created are not stable but propagate normal to the surface with a velocity v_g.

Special cases of $f(z,t)$ are:

(a) $|v_g t|\to 0$. We find from Eq. (8c) that

$$\Delta\Psi_{mn} = \frac{iqEx_{m'm}t}{2\hbar}\,e^{-\kappa|z|}\left(\phi_{m'}\,\frac{\cos k_n z}{\sqrt{a}}\,e^{-i(\omega_{m'}+\omega_{n'o})t}\right). \tag{9a}$$

Here, $\Delta\Psi_{mn}$ maps exactly onto the spatial dependence of the photon field, as discussed above and as expected

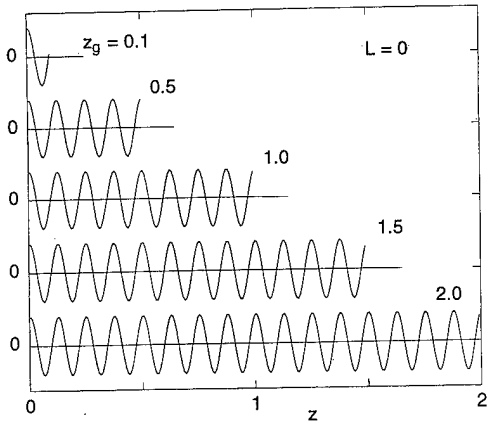


FIG. 2. Wave functions after times where the front is at the values of z_g indicated in the figure. All units are dimensionless. The penetration depth of the photon is $L=0$.

from energy considerations. In this case the absorbed energy simply accumulates without propagation.

(b) $L\rightarrow\infty$. This is the conventional RPA picture. The result can be obtained from either Eq. (8b) or (8c), which is not surprising because in the conventional RPA calculation there is no distinction between the regions defined by $z<z_g$ and $z>z_g$. Energy simply accumulates, since propagation is not relevant.

(c) $L\rightarrow 0$. Here

$$f(z,t)=u(|z_g|-|z|), \tag{9b}$$

which describes a wave that is launched at $t=0$ from $z=0$, is continually renewed at $z=0$, and propagates in both directions with wave fronts at $|z|=|z_g|$. Propagation dominates and energy does not accumulate in the near-surface region.

The above is illustrated in Figs. 2–4. Figures 2 and 3 show the final-state wave packet for various distances z (in dimensionless units) at various times $z_g=v_g t$ for two different photon penetration depths $L=0$ and 0.1, respectively. These figures provide insight into the relative importance of the competition between generation and propagation. As discussed above, propagation dominates completely for $L=0$,

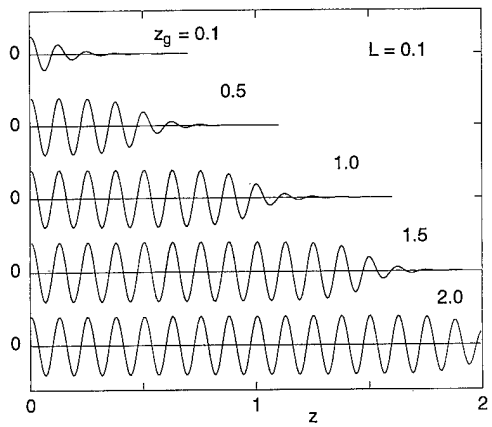


FIG. 3. As Fig. 2, but for $L=0.1$.

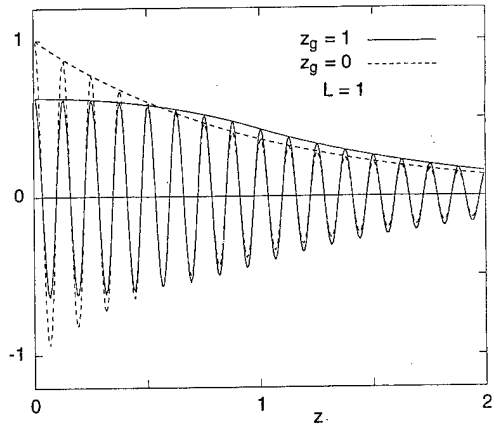


FIG. 4. Comparison of wave functions generated with a photon of absorption length $L=1.0$ with propagation ($v_g\neq 0$, solid lines) and without propagation ($v_g=0$, dashed lines). For $v_g\neq 0$ the front is at $z=v_g t=z_g=1.0$. For $v_g=0$ the amplitude follows the exponential decay of the photon and its scaling simply increases until the excitation scatters.

with the envelope being a unit step function $u(|z_g|-|z|)$. This represents a situation where the amplitude of the wave is determined by the value of the photon field at the surface, no buildup of energy is possible, and the wave front cannot extend past the point given by the product of the group velocity and the time that the photon has been active. For $L=0.1$ the situation is similar, although the rounding of the leading edge due to the finite penetration depth of the photon is evident. In Fig. 4 we compare envelopes of wave functions for $L=1$ with propagation ($v_g t=z_g=1$, solid curves) and without propagation ($v_g=0$, dashed curves). The parameters here are approximately those for optical absorption at the E_2 peak at 4.24 eV in crystalline Si (see below). Without propagation the line shape remains a scaled version of that of the photon field, with an amplitude that in the absence of scattering would increase indefinitely. Figure 4 also illustrates the saturation of the amplitude of the wave function that exists at the surface in the presence of propagation as a result of energy being transported out of the surface region.

Consequently, one result of retaining correlations that are ordinarily discarded in the RPA is to bring the absorption process as described by photon-localized Bloch wave packets into agreement with that expected for hypothetical highly localized (point) excitations. By elementary energy considerations, regions with a higher light intensity will have a higher concentration of such excitations, and if the excitations propagate they will carry energy out of the excitation region into the bulk or toward the surface of the material. In particular, for such excitations the final-state and probability creation rates $d\Delta\Psi_{mn}/dt$ and $d|\Delta\Psi_{mn}|^2/dt$ are proportional to the local photon field and light intensity, respectively, as obtained here. The calculation also shows that the existence of propagation implies that the induced polarization is strictly local only for vanishingly small times or zero group velocity normal to the surface, when the amplitude of $\Delta\Psi_{mn}$ follows that of the photon field. Whether the inherent nonlocality is significant for a particular problem depends on the distance $v_g\tau$ that the excitation travels relative to L before it

scatters, where τ is the lifetime of the excitation.

We now qualitatively consider broadening. In practice the final-state wave function does not evolve indefinitely but terminates at the point where it scatters into another state. In this picture the finite number of cycles over which the excited state exists implies that it can couple to a range of photon frequencies, or alternatively, that an incoming photon of a given frequency can excite final states over a range of frequencies $\Delta\omega \neq 0$. The quantitative description of broadening follows from the calculation of the expectation value of energy absorbed by the material, which to first order involves the matrix element of the polarization operator. This is where energy conservation enters the calculation, but a quantitative treatment is beyond the scope of the present work and will be discussed elsewhere. However, it can be noted that in this picture broadening is described naturally in terms of the finite lifetime of the excitation rather than a phenomenological broadening parameter, and in the usual case of filter theory for individual transitions has the form of the sinc function squared rather than a Lorentzian line shape.

The above discussion is given entirely in terms of the final conduction band state generated from an initial unperturbed valence state $\Psi_{mn}(r,t)$. In fact, conservation of charge implies that the creation of a localized excited electronic state results in the simultaneous creation of a correspondingly localized excited hole state that propagates with the appropriate group velocity of the valence band. Thus expressions equivalent to Eqs. (6) and (7) exist for the excited hole state as well. Also, the use of actual Bloch waves of the material rather than the standing cosine waves of the simple model will result in packets propagating toward as well as away from the surface, thereby also describing photoemission. In the present formalism excitation and propagation, two of the three steps in the three-step model of photoemission, are part of the same process.

III. APPLICATION

We consider estimates appropriate to the (E_0',E_1) complex of c-Si. Here, the extinction coefficient is about 200 Å. From band structure calculations we estimate for an (001) surface and excitation along the Λ axis in reciprocal space that $v_g = dE_{cv}/dk_z \approx 1\ \text{eV}/(\pi\hbar\sqrt{3}/a_0)$, where $a_0 = 5.43$ Å is the lattice constant of c-Si and E_{cv} is the energy difference between conduction and valence bands. Given a broadening parameter Γ of about 40 meV for the complex, we find that $v_g\tau = v_g\hbar/\Gamma = z_g \approx 20$ Å. Therefore, in this case propagation can be neglected and a local description of the dielectric response is probably a good approximation. On the other hand, for the E_2 transition, where v_g and κ are about three times larger and Γ is about twice as large, L and z_g are comparable. Here, the assumption of a local dielectric response may not be justified.

We next consider the effect on the apparent energies of the critical points. The energy E_{mn} of the final-state wave packet $\Delta\Psi_{mn}$ is not that of the resonant component, $\omega_{m'} + \omega_{n'o}$, but is rather given by the expectation value

$$\langle E_{mn}\rangle = \frac{\langle \Delta\Psi_{mn}|H_o + V_s|\Delta\Psi_{mn}\rangle}{\langle \Delta\Psi_{mn}|\Delta\Psi_{mn}\rangle}. \tag{10}$$

It is thus a weighted average of bulk and surface potentials that depends on the overlap between the potentials and $\Delta\Psi_{mn}$. Consequently, the apparent energy of the final state will depend on various factors including, through selection rules, the polarization state of light inducing the transitions. Estimating surface potentials of the order of 1 eV and a relative surface-to-bulk penetration depth on the order of 50–100 Å leads to surface-generated differences on the order of 10–20 meV, consistent with the results of Fig. 1 and other measurements where derivative components are observed. The result differs from a previous interpretation[12] of derivative line shapes in reflectance difference (RD) spectra in that our treatment is general and not specifically restricted to E_1 transitions.

IV. CONCLUSION

In summary, we have shown that absorption of light in optically absorbing materials involves two distinct physical phenomena: the creation of a wave packet that initially maps exactly onto the field of the photon that created it, and propagation of the packet from the region of excitation. The situation is completely analogous to the generation of electromagnetic waves propagating in space. These effects will clearly become more enhanced as the penetration depth of light decreases, consistent with the larger derivative signals obtained for the E_2 transition of crystalline Si near 4.2 eV.[6,7] The results also show that slab calculations, which incorporate spatial confinement as an unwelcome consequence of limited computing power, may actually provide a more accurate representation of surface-related phenomena than calculations that take the entire bulk into account. We have also shown that it may be necessary to treat optical absorption in strongly absorbing materials as an inherently nonlocal phenomenon, certainly in situations where the polarization can propagate an appreciable fraction of the penetration depth of light. Finally, we have discussed only critical point-energy consequences and have ignored surface-amplitude effects, which have recently received considerable attention.[13,14] The theory is easily generalized to incorporate these effects as well, and will be discussed elsewhere.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge support of this work by the Office of Naval Research under Contract No. N00014-93-1-0255, and of further support of UR by the Alexander von Humbolt Foundation.

[1]P. J. Feibelman, Phys. Rev. B **12**, 1319 (1975).
[2]P. J. Feibelman, Phys. Rev. B **14**, 762 (1976).
[3]R. Del Sole, J. Phys. C **8**, 2971 (1975); A. D'Andrea and R. Del Sole, Phys. Rev. **25**, 3714 (1982).
[4]F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).
[5]L. Mantese, U. Rossow, and D. E. Aspnes, Appl. Surf. Sci. **107**, 35 (1996).

[6]U. Rossow, L. Mantese, and D. E. Aspnes, in *Proceedings of the 23rd International Conference on the Physics of Semiconductors*, edited by M. Scheffler and R. Zimmerman (World Scientific, Singapore, 1996), p. 831.
[7]L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes, J. Vac. Sci. Technol. B **15**, 1196 (1997).
[8]T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12**, 1152 (1994).
[9]K. A. Bell, L. Mantese, U. Rossow, and D. E. Aspnes, J. Vac. Sci. Technol. B **15**, 1205 (1997).
[10]E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959).
[11]See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955), p. 195 ff.
[12]D. E. Aspnes and A. A. Studna, J. Vac. Sci. Technol. A **5**, 546 (1987).
[13]K. Uwai and N. Kobayashi, Phys. Rev. Lett. **78**, 959 (1997).
[14]R. Eryigit and I. P. Herman, Phys. Rev. B **56**, 9263 (1997).

Reflection high-energy electron diffraction/scanning tunneling microscopy study of InAs growth on the three low index orientations of GaAs: Two-dimensional versus three-dimensional growth and strain relaxation

B. A. Joyce, T. S. Jones,[a] and J. G. Belk
Semiconductor Materials Interdisciplinary Research Centre, The Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom

(Received 21 January 1998; accepted 11 May 1998)

It is generally believed that strain relaxation in semiconductor heterostructures having a significant misfit (>2%) occurs by the formation of coherent three-dimensional islands, following the growth of one or two continuous two-dimensional monolayers in a manifestation of the Stranski–Krastanov (SK) growth mode. For the InAs–GaAs system, for which the misfit is ≈7.2%, we have shown that this is a very special case, at least during growth by molecular beam epitaxy, as it occurs on only one of the low index orientations, the (001), and then only under As-rich growth conditions. On (110) and (111) surfaces, growth is always two-dimensional layer by layer and strain is relieved by the formation of misfit dislocations. Even when three-dimensional islands are formed on (001) substrates, the process is much more complex than the conventional SK mechanism would imply. We have used a combination of *in situ* reflection high-energy electron diffraction (RHEED) and scanning tunneling microscopy (STM) and *ex situ* transmission electron microscopy to follow the initial growth processes and strain relaxation mechanisms of InAs on GaAs (001), (110), and (111)*A* surfaces. RHEED enables us to establish the symmetry of the wetting layer, whether growth is two- or three-dimensional and the external crystallography of any three-dimensional islands. STM images obtained by rapid quenching from the growth temperature show how growth is initiated, provide quantitative data on island formation (number density and volume), and indicate dislocation formation via strain field morphology effects. Detailed results for all three low index surfaces are presented, together with possible reasons for the major differences between them. © *1998 American Vacuum Society.* [S0734-211X(98)05204-4]

I. INTRODUCTION

The growth modes and strain relaxation mechanisms of heteroepitaxial III–V systems in which there is a moderate degree of mismatch (typically, >2%) have assumed particular importance in the fabrication technology of low-dimensional structures and devices. It is generally assumed that the usual behavior is a Stranski–Krastanov (SK) growth mode, with strain relaxation in the early stages occurring by the formation of three-dimensional (3D) coherent islands, frequently referred to as quantum dots.[1–6] Beyond the stage at which there is significant island coalescence, misfit dislocations are introduced, but these later stages are not strictly relevant to the formation of heterojunctions and will not be discussed here.

In this article we will show, with reference to the InAs–GaAs system, that this so-called usual behavior is in fact limited to one, or possibly two special cases and that the normal situation is represented by two-dimensional (2D) layer-by-layer growth, with misfit dislocation formation after a few monolayers (MLs) providing the means of strain relaxation. All of the films have been grown by solid-source molecular beam epitaxy (MBE), using the *in situ* diagnostic probes of reflection high-energy electron diffraction (RHEED) to obtain dynamic real-time information and scanning tunneling microscopy (STM) to give a series of "snapshot" real-space images, with atomic resolution where appropriate. *Ex situ* transmission electron microscopy (TEM) has also been used to provide confirmation of various dislocation structures. We are not able to offer evidence on whether our results would also apply to other growth techniques such as metal–organic chemical vapor deposition (MOCVD), but we suspect that is likely to be so.

An important aspect of our work is that we have investigated not only different surface reconstructions on (001) substrates, but also other low index orientations, specifically (110) and (111)*A*. This enables us to state unequivocally that lattice mismatch is a necessary, but not sufficient criterion for SK growth. The composition of the 3D islands when they do form is also different from that which is conventionally expected, in that they incorporate substrate material and the nature of the intermediate, or "wetting" layer is also important for a complete description of the overall growth process. We will address each of these points, dealing first with growth modes and strain relaxation mechanisms via a comparison of the growth of InAs on the three low index orientations of GaAs, then with wetting layer formation, and finally the volume, composition, and shape of the islands which form under special circumstances.

II. GROWTH MODES AND STRAIN RELAXATION MECHANISMS

We have used RHEED and STM to follow the growth mode of InAs on three low index orientations of GaAs;

[a]Also at Department of Chemistry, Imperial College, Frankland Road, London SW7 2AZ.

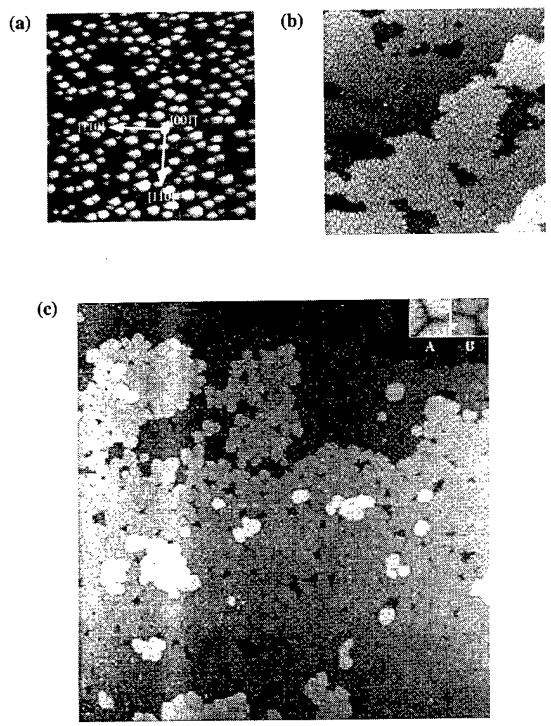


FIG. 1. Filled states STM images of 2 ML of InAs grown on (a) GaAs(001), (b) GaAs(110), and (c) GaAs(111)*A* at 420 °C. Image dimensions 2000 Å ×2000 Å.

(001), (110), and (111)*A*. There is, of course, a constant lattice mismatch of ≈7.2%, so if strain were the only criterion by which the growth mode (and implicitly the strain relaxation mechanism) were determined the results should be independent of orientation. That this is clearly not the case can be seen first from the RHEED response,[7] where for growth on (110)-(1×1) and (111)*A*-(2×2) surfaces, the patterns remain streaked for all thicknesses (up to ≈1000 Å, the maximum we investigated) and intensity oscillations could be readily observed. By contrast, for growth on (001)-2×4 and $c(4\times4)$ surfaces, transmission diffraction spots developed after ≈2 ML deposition, indicating the onset of 3D growth, which was substantiated by there being at most only one measurable oscillation of the specular beam intensity.

This reciprocal space information was confirmed by real-space STM images, obtained by rapid quenching from the growth temperature and *in situ* transfer. The results are illustrated in Fig. 1, which shows images of ≈2 ML InAs deposition at 420 °C on the three low index orientation GaAs substrates. The occurrence of 3D islands is quite apparent on the (001) surface, whereas on the other two orientations the mode is 2D, which is maintained throughout growth of thicker films (see below).

On GaAs(001) substrates, 3D growth is further restricted to As-stabilized reconstructions, specifically 2×4 and $c(4\times4)$. Previous work by several other groups[8–11] has established that if growth is carried out under cation-rich conditions, then even on (001) surfaces growth is 2D, an effect which we have also observed. The remaining low index surface is the (111)*B*, which in principle is also As terminated and has both $\sqrt{19}\times\sqrt{19}$ and 2×2 reconstructions, but with surfaces misoriented by 3° towards $[2\bar{1}\bar{1}]$, Schujman *et al.*[12] observed, by *ex situ* atomic force microscopy (AFM), 3D islands after 1–2 ML deposition, which decorated the steps formed by the vicinal planes. These islands were, however, very different from those shown in Fig. 1(a), being much more crystallographic, and an order of magnitude larger both laterally and vertically; it is not clear whether the two cases of 3D growth can be attributed to a common mechanism, but superficially it appears that an anion-stable surface is a necessary condition. This point will be considered further in the discussion section, but first we will consider the mechanisms of strain relaxation in rather more detail for each orientation separately.

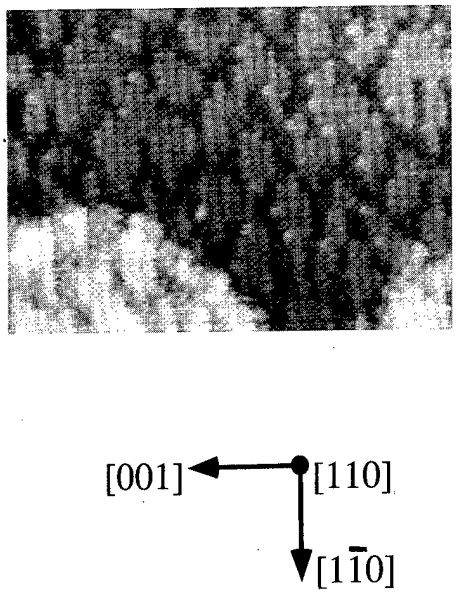


FIG. 2. Filled states STM image of 2 ML of InAs grown on GaAs(110) at 420 °C. Image dimensions 200 Å×150 Å.

A. InAs/GaAs(110)

The general mechanism of strain relaxation in this system has previously been studied by TEM and x-ray diffractometry (XRD). Two distinct classes of dislocations are required to accommodate the crystallographic anisotropy of the (110) plane. Only two low-energy $\{111\}\langle110\rangle$ slip systems are available, based on (111) and $(11\bar{1})$, both of which produce maximum strain relief along [001] within the (110) plane. Conventional slip cannot relieve strain in the orthogonal $[1\bar{1}0]$ direction, so pure edge dislocations are incorporated directly during the initial growth stages. These edge dislocations have line vectors $\mathbf{u}=[001]$ and Burger vectors $\mathbf{b}=a_o/2[1\bar{1}0]$. We have previously shown[13,14] how STM imaging can be used to analyze in detail the edge dislocations and slip steps and we summarize those findings here.

The growth of the first ML is closely similar to the homoepitaxial growth of GaAs(110), with the formation of a high number density ($\approx1.0\times10^{11}$ cm^{-2}) of 2D nuclei, either

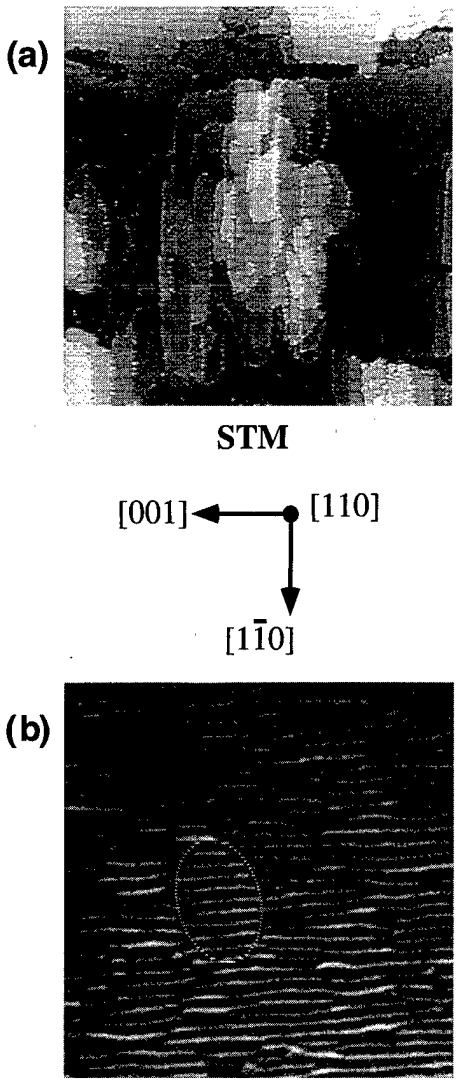


FIG. 3. Dislocations imaged by two different microscopies for different areas of a 5 ML InAs film grown on GaAs(110) at 420 °C: (a) STM and (b) TEM. Image dimensions 2000 Å × 2000 Å.

single ML or double (bilayer) in height, which is also characteristic of homoepitaxy.[15] Beyond 1 ML, however, the morphological similarity between homo- and heteroepitaxy ceases and a surface structure develops which is a precursor to the formation of misfit dislocations at 3 ML. The new structure takes the form of a mosaic, in which a high density of 1 ML deep cracks is generated in the islands. They have a width of 6 Å in the [001] direction, i.e., one missing row of unit cells and a spacing of ≈30 Å. This is illustrated in the STM image shown in Fig. 2. The cracks appear first at the peripheries of the islands and propagate inwards, until by 2 ML the mosaic is a dense array of miniature InAs islands with a number density in excess of 1.0×10^{12} cm^{-2}, although the exact number is somewhat dependent on growth temperature. The onset of plastic deformation occurs at ≈3 ML, and is made manifest in the STM image by a series of rather broad dark lines along [001] in regions where the InAs surface is no longer fractured into small islands. The lines are

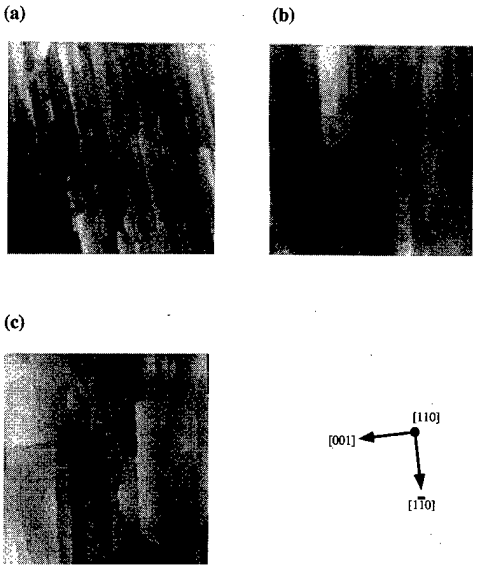


FIG. 4. Filled states STM images of InAs grown on GaAs(110) at 420 °C: (a) 10 ML, (b) 20 ML, and (c) 100 ML. Image dimensions 2000 Å × 2000 Å.

≈0.5 Å deep, with no significant disturbance of the surface atoms around them and the effect can be described as a localized downward displacement of the surface atomic plane associated with misfit dislocations. To obtain confirmation of this explanation, a 5 ML InAs/GaAs(110) film was imaged both by *in situ* STM and *ex situ* TEM, as illustrated in Fig. 3. The dislocation network, imaged via its influence on the surface morphology by STM can be seen to be identical in orientation, periodicity, and form to that resolved by plan-view TEM. This shows that the dislocations are members of the set of edge dislocations having a Burgers vector of $\mathbf{b}=a_o/2[1\bar{1}0]$ and a spacing along $[\bar{1}10]$ of ≈60 Å, which very largely accommodates the +7.2% lattice mismatch. The STM imaging mechanism is then largely reliant on the geometric influence of the dislocation strain field at the surface. These dislocations are completely confined within the (110) plane and cannot be channeled down to the interface by any kind of slip system. From careful examination of the STM images, the nucleation mechanism appears to involve the coalescence of small 2D islands to form a length of misfit dislocation line directly beneath the boundary where the islands came together.

At this stage (<5 ML) growth), the layer is still fully strained in the orthogonal [001] direction. Relaxation here involves slip of surface nucleated half-loop dislocations on inclined {111} planes and these dislocations can also be observed by STM. This is possible because a half-loop dislocation operates by displacing the whole layer by a small amount along an inclined {111} plane, thereby producing a discontinuity at the heterointerface, which is the misfit dislocation. A corresponding displacement occurs at the surface where it is intersected by the {111} slip plane and this is usually referred to as a slip step. The detailed atomic structure can be resolved by STM and we have described this in detail for InAs/GaAs(110) elsewhere.[14] Briefly, the change

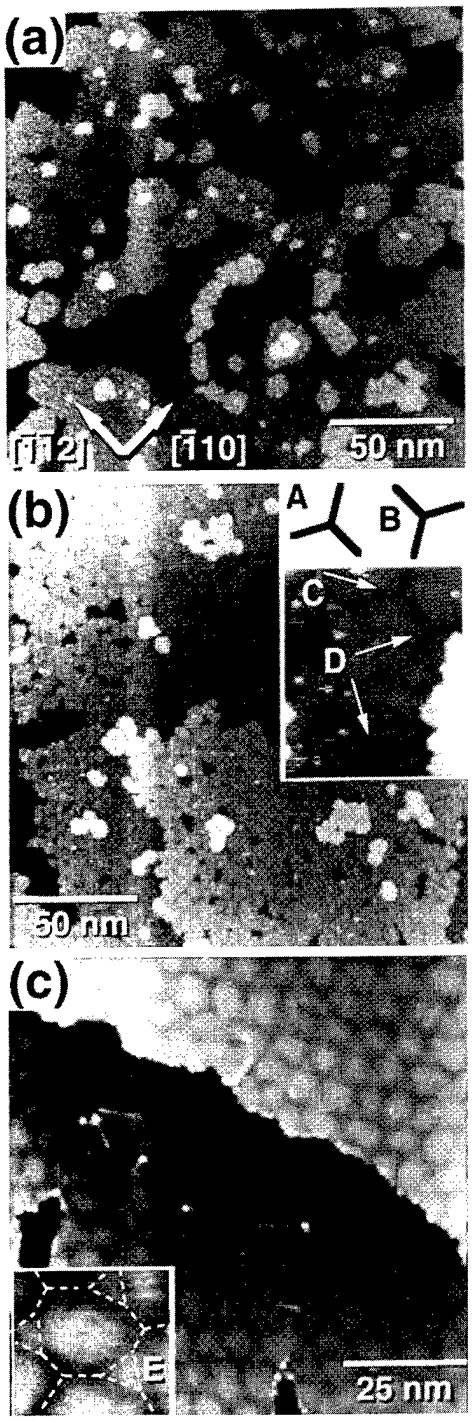


FIG. 5. Filled states STM images of (a) 0.5 ML, (b) 2 ML, and (c) 5 ML of InAs grown on GaAs(111)*A* at 420 °C: Image dimensions 2000 Å ×2000 Å.

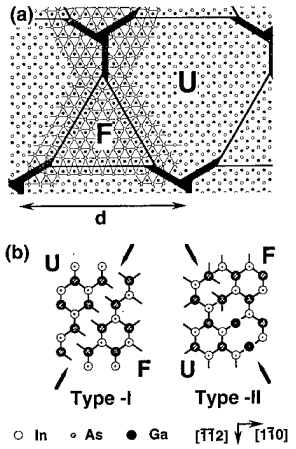


FIG. 6. Model of atomic structure of InAs/GaAs(111)*A* interface. Open (closed) circles show the position of Ga (In) atoms just below (above) the interface. F and U indicate faulted and unfaulted interfaces, respectively. Grey lines and bands display the positions of partial and ideal edge dislocations, while the thin black line is a guide to the eye in tracing the trigonal In lattice. (b) The nominal atomic structure of two possible partial dislocations, with the dislocation line direction indicated by the arrows. Shaded circle shows an As atom at the interface. Two variants of the type II structure are possible.

in surface height is entirely localized to the slipped region of the layer, there are no wider scale effects to perturb the morphology. The sequence of STM images shown in Fig. 4 indicates how the slip step structure develops for film thicknesses between 10 and 100 ML. The complete absence of 2D nuclei indicates that growth occurs by a process of step propagation and for a singular surface almost all of the surface steps will have been formed by dislocation slip. Clearly, the process of strain relaxation along [001] takes place much more slowly than in the orthogonal [1$\bar{1}$0] direction, which has been substantiated by RHEED measurements, where there is an abrupt change in the spacing of the integral order rods after about 3 ML, but no such rapid change is observed for [001].

B. InAs/GaAs(111)*A*

For (111)*A* surfaces, strain relaxation also occurs by the formation of misfit dislocations,[16] but with a different mechanism from that observed on (110) oriented material. It is illustrated in Fig. 5, which shows STM images after the growth of 0.5, 2, and 5 ML, respectively, with a 2×2 reconstruction being maintained throughout (including the initial substrate). The basic mechanism is that as the 2D, approximately hexagonal-shaped islands coalesce in all three equivalent ⟨110⟩ directions at a thickness of ≈2 ML, dislocations are introduced into the first InAs ML. The edge dislocations, having Burgers vectors $\mathbf{b}=a_o/2\langle 110\rangle$ and line directions $\mathbf{u}=\langle 112\rangle$, produce a downward displacement in the surface of ≈0.5 Å due to the effect of their strain fields. There are two possible dislocation junctions in a hexagonal network which are mirror images [labeled A and B in the inset to Fig. 5(b)]. There is a barrier to the complete closure of type B, since small 1 ML deep holes remain at the intersection, but a perfect dislocation triple node is formed with type A. During deposition of the subsequent ML, the B junction closes by forming an open triangle from three partial dislocations ($\mathbf{b}=a_o/6\langle 112\rangle$, $\mathbf{u}=[110]$), so that the final network is still hexagonal, with alternate open nodes, but enclosed within the partial dislocation triangle is a stacking fault in the (111) growth plane [Fig. 5(c) and inset]. The

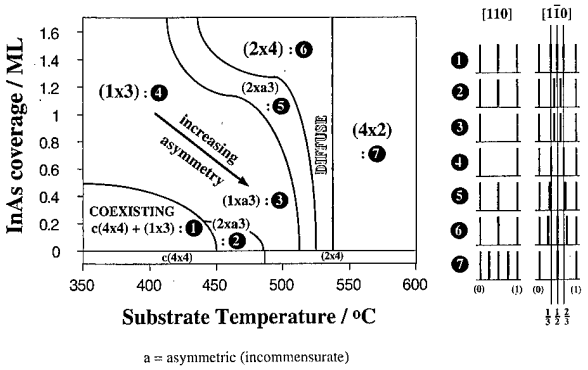


FIG. 7. Schematic diagram indicating the RHEED patterns observed during the 2D growth of InAs on GaAs(001) as a function of nominal InAs coverage and substrate temperature. Patterns themselves seems along the [110] and [1$\bar{1}$0] azimuths are also reproduced for each of the seven characteristic regions: black bars are integral order rods, and grey bars are fractional order rods, with a height proportional to their intensity.

idealized structure is illustrated in Fig. 6. In this system strain relaxation is comparatively slow, being less than 100% at 300 ML (as measured by transmission electron diffraction).

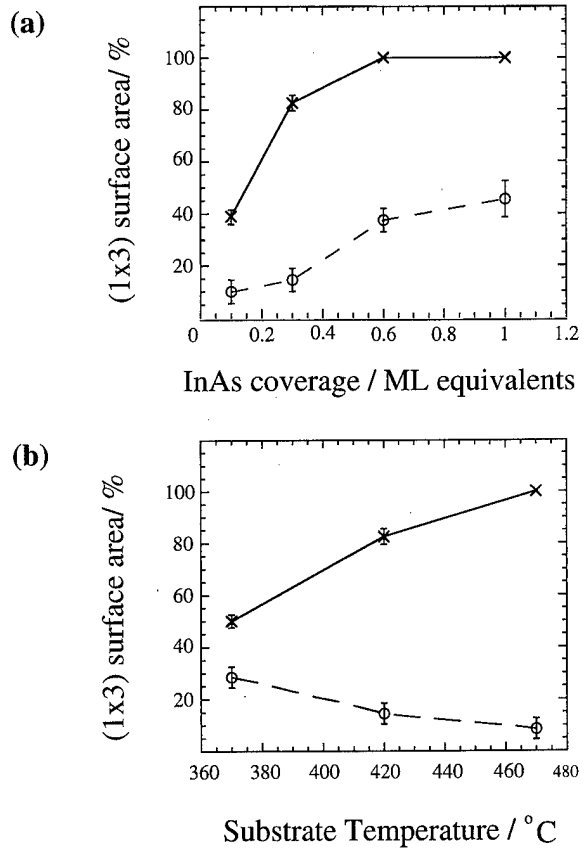


FIG. 8. Plots based on the STM data describing the alloy growth mode for InAs/GaAs(001), in terms of the percentage surface area of the (1×3) reconstruction as a function of (a) InAs coverage at a fixed substrate temperature of 420 °C, and (b) as a function of substrate temperature at a fixed InAs coverage of 0.3 ML; (-×-) represents the total area of the (1×3), which includes a component from (1×3) islands, which is also plotted separately (-O-).

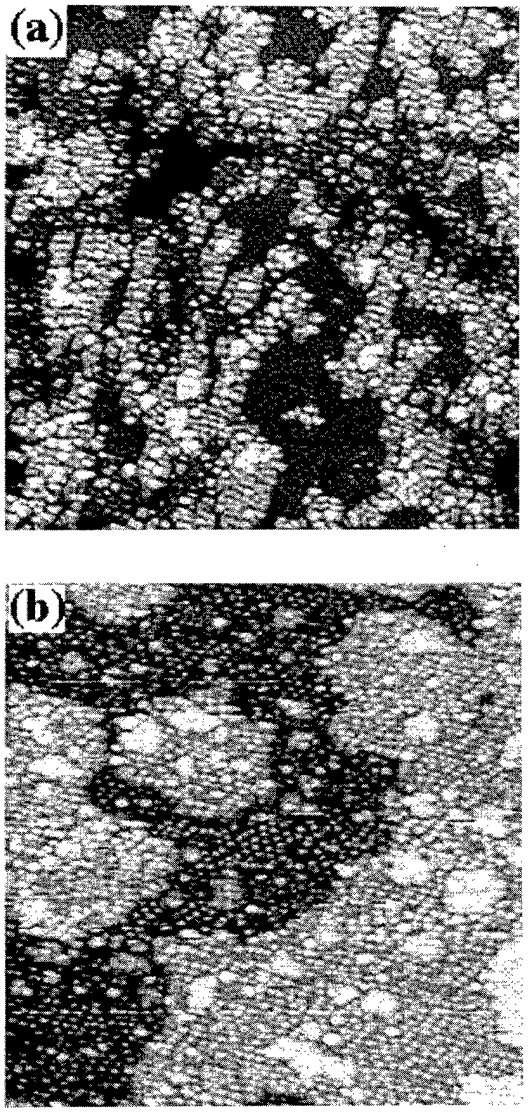


FIG. 9. STM images demonstrating the effect of substrate temperature on the alloy growth mode of InAs/GaAs(001); (a) 0.6 ML InAs at 370 °C (400 Å×400 Å), (b) 0.3 ML InAs at 470 °C (400 Å×400 Å).

To the best of our knowledge, there are no reports on strain relaxation mechanisms for InAs/GaAs(111)B, although as stated earlier, a form of 3D growth occurs.[12]

C. InAs/GaAs(001)-2×4 and c(4×4)

It is abundantly clear in this extremely well-documented system that in the initial growth stages (<2 ML), strain relaxation occurs by the formation of coherent 3D islands.[1–6] Prior to the coalescence of these islands, no dislocations are introduced. The effect has also been treated theoretically[17–19] and we will not comment further at this stage, although in Sec. IV we will provide more details of the size, shape, and composition of the islands.

Two additional points need to be made, however, the first to consider dislocation formation beyond the 3D island coalescence stage and the second in relation to growth on cation-rich surfaces. For relaxation beyond coalescence, the

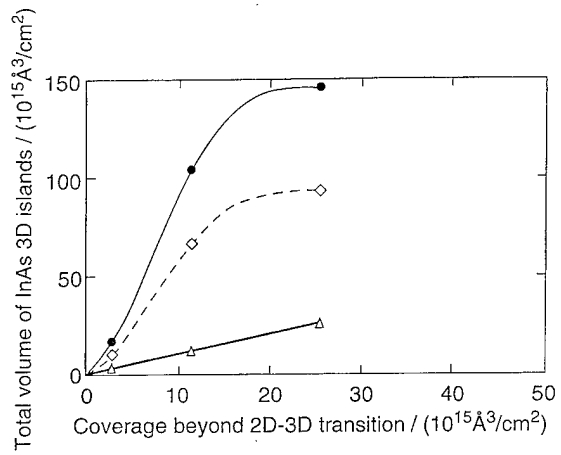


FIG. 10. Plot showing the integrated volume of material in InAs quantum dots (QDs) on GaAs(001) at 420 °C for three coverages beyond the 2D to 3D transition at 1.6 ML (corresponding to nominal InAs thicknesses of 1.7, 2.0, and 2.5 ML). Symbols: triangles, nominal 1:1 volume relationship; circles, measured volume; and diamonds, experimental data scaled by a factor of 0.6 to compensate for tip convolution.

behavior on (001) surfaces involves slip of surface-nucleated half-loop dislocations on inclined {111} planes, analogous to that described above for (110) films. There is a quantitative difference, however, in that misfit segments and surface slip steps are ≈1000 times longer for (001) films. The fact that dislocations are introduced very soon after the islands have coalesced suggests, however, that the island formation process is not very efficient at relaxing strain.

When growth is carried out under strongly In-rich conditions, a 2D layer-by-layer growth mode is observed on (001) substrates, with no indication of coherent island formation.[8–11] The mechanism of strain relaxation is not clear, however, the only evidence coming from the work of Xue *et al.*,[11] who produced high-resolution STM images showing regular "domain-wall" features which are not defined. It is possible that they are in fact surface slip steps, but their correct assignment requires more detailed measurements.

III. INTERMEDIATE (WETTING LAYER) FORMATION

This effect relates predominantly to InAs growth on GaAs(001) surfaces prior to the appearance of 3D islands, but we will also make brief comparisons with (110) and (111)*A* substrates, where it is much less marked.

For (001 surfaces we have followed the formation of this intermediate layer as a function of the amount of InAs deposited, using RHEED and STM).[20,21] The RHEED results are summarized in Fig. 7, from which it is clear that the composition and structure of the wetting layer are not fixed, but are strongly dependent on growth conditions. The most important consideration in relation to strain relaxation models is the stage at which the wetting layer forms a continuous ML, as demonstrated by a homogeneous STM image and a RHEED pattern displaying a single symmetry. This is a strong function of substrate temperature, as shown in Fig. 8, so, for example, at 470 °C, 0.3 ML of InAs produces a complete surface layer, which therefore, has the composition $In_{0.3}Ga_{0.7}As$, while at 420 °C the first ML has the composition $In_{0.67}Ga_{0.33}As$. The effects of substrate temperature on the alloy growth mode are demonstrated in the two STM images shown in Fig. 9 for InAs deposition onto GaAs(001)-$c(4\times4)$. At low temperature, the (1×3) structure is largely confined to 2D islands separated by mostly clean $c(4\times4)$ terraces, while at higher temperatures, the whole surface is already a set of (1×3) terraces (and fewer islands) with no remaining $c(4\times4)$ despite the lower InAs coverage. As further InAs is deposited, the outermost layer becomes more In rich, but clearly the In concentration gradient through this strain mediating layer is temperature dependent. This leads to a temperature-dependent critical layer thickness at which 3D islands form (see Sec. IV).

It is quite evident from these results that alloying of InAs with GaAs takes place very readily on GaAs(001)-2×4 and $c(4\times4)$ surfaces, which means that the overall process of 3D island formation is much more complex than that of the classical SK mechanism. Extensive alloying appears to be restricted to these As-stable (001) surfaces, however, since on (110) and (111)*A* surfaces, the surface coverage corresponds to within 5% of the total amount of InAs deposited.[22] It is possible, therefore, that the intermediate layer and relative ease of alloy formation play an important part in determining the strain relaxation mechanism, but at the moment the evidence is only circumstantial.

IV. 3D ISLAND FORMATION ON As-STABLE GaAs(001) SURFACES

In common with other investigations[1–6] we find from STM images that the islands form spontaneously following the incremental deposition of <0.1 ML of InAs, initially forming predominantly at step edges present in the wetting layer, and reaching saturation number density N_s with a narrow size distribution, within a further 0.5 ML deposition. N_s is temperature dependent, but is typically in the range 5×10^{10} cm^{-2} at 500 °C to 1.0×10^{12} cm^{-2} at 420 °C and there is very little increase in size prior to coalescence. It is frequently stated that the critical thickness of InAs at which 3D growth is initiated is 1.7 ML, but it is in fact temperature dependent, varying from 1.4 ML at 350 °C to 1.8 ML at 500 °C. Furthermore, these thicknesses refer to the amount of InAs deposited, but because of the alloying behavior in the wetting layer (Sec. III), they cannot be considered to be real thicknesses, which in fact would be somewhat greater and of varying composition. These temperature-dependent effects are almost certainly related to the detailed structure and composition of the wetting layer, since it is this which is responsible both for communicating the extent of the misfit strain with the substrate and for adatom migration energetics and kinetics across its surface.

The greatest deviation from conventional SK growth is revealed by estimating the volume of the 3D islands directly from STM measurements and comparing this quantity with the total amount of material deposited beyond the point at which islands first formed. The results are shown in Fig. 10,

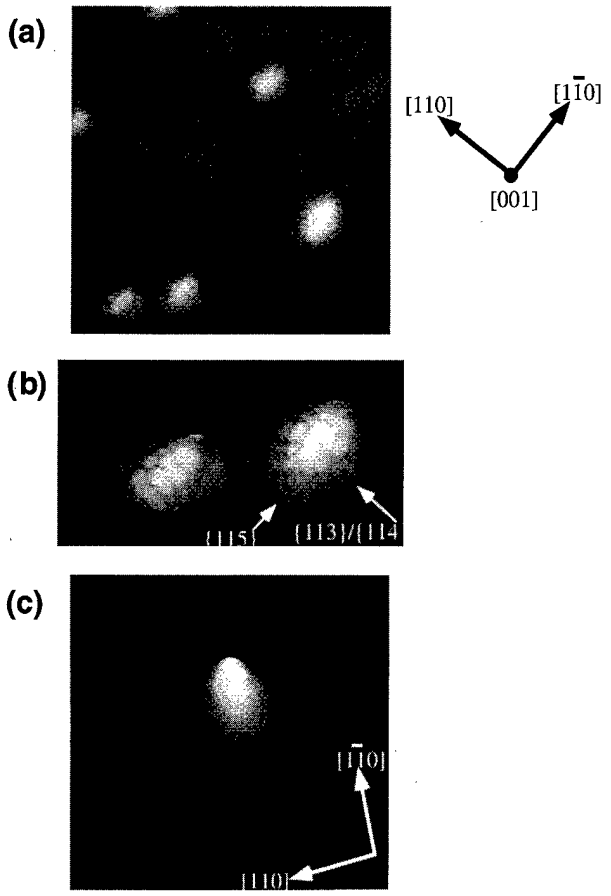


FIG. 11. High-resolution STM images of QDs for 1.7 ML InAs/GaAs at 420 °C: (a) low magnification (500 Å×500 Å) image showing 4 QDs and the wetting layer, (b) higher magnification (200 Å×100 Å) image showing side facets on 2 QDs, and (c) perspective view image (200 Å×200 Å, grey scale: 0–20 Å) showing QD shape anisotropy.

from which it is immediately apparent that the island volume is far greater than the volume of InAs actually deposited. It must be concluded that the additional material is transported from the wetting layer and the islands, as well as the wetting layer, are therefore, (In,Ga)As, not InAs, again illustrating the active role of the wetting layer in this version of the SK growth mode.

Finally, information on the shape of the 3D islands can be obtained from both RHEED patterns and STM images. RHEED shows that facets are formed along only one direction in the surface, so the islands are not symmetrical pyramids.[7] More detail can, however, be extracted from the STM images shown in Figs. 11(a)–(c). Figure 11(a) is a comparatively low magnification image of four islands, from which the RHEED pattern in the [110] azimuth shows the typical chevron features obtained from facets, but in the orthogonal [110] azimuth transmission spots are observed. The angles measured from both techniques show the facets to be close to {113}, also shown in Fig. 11(b). The perspective view in Figure 11(c) demonstrates the anisotropy in island dimensions, which produces a conspicuous ridge along their summit, so that their true shape is significantly different from the idealized square based pyramid which is often used in calculations of their electronic structure.

V. DISCUSSION AND CONCLUSIONS

It is evident from a consideration of the growth of InAs on different orientations and differently reconstructed GaAs surfaces that a 2D layer-by-layer growth mode with strain relaxation by the formation of misfit dislocations is the normal situation. Uniquely, on As-stable (001) surfaces, a highly modified form of Stranski–Krastanov growth occurs and coherent 3D islands relieve some of the initial strain.

Based on this unequivocal evidence from one mismatched III–V system, but which we believe may well be typical, we propose an alternative viewpoint from which to consider strain relaxation in this class of heterostructures. Stated most simply, we suggest that their normal strain relaxation mechanism is by the formation of misfit dislocations, whose details are determined by the slip systems available, while growth follows a 2D layer-by-layer mode over the entire film thickness. Exceptionally, some (undefined) process leads to 3D growth, which coincidentally provides an alternative, if inefficient, means of strain relaxation before the onset of plastic deformation. In other words, it is not strain relaxation which provides the driving force for 3D growth, but rather where there is a separate driving force for this growth mode, strain relaxation occurs as a result.

The open question is then what provides this alternative driving force? If, for example, it were easier for a migrating atom on an anion stable surface to hop up from a lower terrace to a higher one than on a cation stable surface, a basis would exist. At this stage, however, there are virtually no quantitative data on step-edge barriers, but clearly the ability to hop up a step is a prerequisite for 3D growth.

A second alternative could relate to the structure and composition of the wetting layer, which we have shown to be strongly alloyed on As-stable (001) surfaces, but much less so on other orientations. If this intermediate layer is $(In_{1-x}Ga_x)As$, (where x varies with thickness) and not InAs, it is in principle able to provide strain mediation such that the "critical thickness" for dislocation formation is greater than on surfaces where alloying is less pronounced. 3D growth would then be initiated before the need for plastic deformation occurred.

Finally, the process by which 3D islands form is clearly complex and any model must at least account for the formation of a variable composition wetting layer, a narrow size distribution, and the incorporation of substrate material into the islands. No complete model is yet available, but we have made some progress in developing a self-consistent rate equation analysis based on the general rate equation theory of epitaxial growth discussed by Venables.[23,24] Details have been presented elsewhere.[25]

ACKNOWLEDGMENTS

The authors gratefully acknowledge many valuable discussions with Professor Dimitri Vvedensky. Financial support was provided by EPSRC under Grant No. GR/J97540.

[1]A. Madhukar and K. C. Rajkumar, Appl. Phys. Lett. **57**, 2110 (1990).
[2]P. M. Petroff and S. P. DenBaars, Superlattices Microstruct. **15**, 15 (1994).
[3]D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. DenBaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).
[4]J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. André, and O. Vatel, Appl. Phys. Lett. **64**, 196 (1994).
[5]A. Ponchet, A. Le Corre, H. L'Haridon, B. Lambert, and S. Salaun, Appl. Phys. Lett. **67**, 1850 (1995).
[6]W. Seifert, N. Carlsson, M. Miller, M.-E. Pistol, L. Samuelson, and L. R. Wallenberg, Prog. Cryst. Growth Charact. Mater. **33**, 423 (1996).
[7]H. Yamaguchi, M. R. Fahy, and B. A. Joyce, Appl. Phys. Lett. **69**, 505 (1996).
[8]E. Tournie, A. Trampert, and K. H. Ploog, Europhys. Lett. **25**, 663 (1994).
[9]C. W. Snyder, B. G. Orr, and H. Munakata, Appl. Phys. Lett. **62**, 46 (1992).
[10]N. Grandjean, J. Massies, and V. H. Etgens, Phys. Rev. Lett. **69**, 796 (1992).
[11]Q. K. Xue, Y. Hasegawa, T. Ogino, H. Kiyama, and T. Sakurai, J. Vac. Sci. Technol. B **15**, 1270 (1997).
[12]S. B. Schujman, S. R. Soss, K. Stokes, and L. J. Schowalter, Surf. Sci. **385**, L965 (1997).
[13]J. G. Belk, J. L. Sudijono, X. M. Zhang, J. H. Neave, T. S. Jones, and B. A. Joyce, Phys. Rev. Lett. **78**, 475 (1997).
[14]J. G. Belk, D. W. Pashley, C. F. McConville, J. L. Sudijono, B. A. Joyce, and T. S. Jones, Phys. Rev. B **56**, 10 289 (1997).
[15]D. M. Holmes, J. G. Belk, J. L. Sudijono, J. H. Neave, T. S. Jones, and B. A. Joyce, J. Vac. Sci. Technol. A **14**, 849 (1996).
[16]H. Yamaguchi, J. G. Belk, X. M. Zhang, J. L. Sudijono, M. R. Fahy, T. S. Jones, D. W. Pashley, and B. A. Joyce, Phys. Rev. B **55**, 1337 (1997).
[17]J. Tersoff and F. K. LeGoues, Phys. Rev. Lett. **72**, 3570 (1994).
[18]H. T. Dobbs, D. D. Vvedensky, A. Zangwill, J. Johansson, N. Carlsson, and W. Seifert, Phys. Rev. Lett. **79**, 897 (1997).
[19]C. Priester and M. Lannoo, Phys. Rev. Lett. **75**, 93 (1995).
[20]J. G. Belk, J. L. Sudijono, D. M. Holmes, C. F. McConville, T. S. Jones, and B. A. Joyce, Surf. Sci. **365**, 735 (1996).
[21]J. G. Belk, C. F. McConville, J. L. Sudijono, T. S. Jones, and B. A. Joyce, Surf. Sci. **387**, 213 (1997).
[22]J. G. Belk, D. W. Pashley, C. F. McConville, B. A. Joyce, and T. S. Jones, Surf. Sci. (submitted).
[23]D. R. Frankl and J. A. Venables, Adv. Phys. **19**, 409 (1970).
[24]J. A. Venables, Philos. Mag. **27**, 697 (1973).
[25]B. A. Joyce, J. L. Sudijono, J. G. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedensky, and T. S. Jones, Jpn. J. Appl. Phys., Part 1 **36**, 4111 (1997).

Structure of InAs/AlSb/InAs resonant tunneling diode interfaces

B. Z. Nosho[a] and W. H. Weinberg
Center for Quantized Electronic Structures and Department of Chemical Engineering, University of California, Santa Barbara, California 93106

J. J. Zinck
HRL Laboratories, Malibu, California 90265

B. V. Shanabrook, B. R. Bennett, and L. J. Whitman[b]
Naval Research Laboratory, Washington, DC 20375

(Received 21 January 1998; accepted 20 May 1998)

We have used *in situ* plan-view scanning tunneling microscopy to study the surfaces and interfaces within an InAs/AlSb/InAs resonant tunneling diodelike structure grown by molecular beam epitaxy. The nanometer and atomic-scale morphologies of the surfaces have been characterized following a number of different growth procedures. When InAs(001)-(2×4) is exposed to Sb_2 a bilayer surface is produced, with 1 monolayer (ML) deep (3 Å) vacancy islands covering approximately 25% of the surface. Both layers exhibit a (1×3)-like reconstruction characteristic of an InSb-like surface terminated with >1 ML Sb, indicating that there is a significant amount of Sb on the surface. When 5 ML of AlSb is deposited on an Sb-terminated InAs surface, the number of layers observed on each terrace increases to three. Growth of an additional 22 ML of InAs onto the AlSb layer, followed by a 30 s interrupt under Sb_2, further increases the number of surface layers observed. The root-mean-square roughness is found to increase at each subsequent interface; however, on all the surfaces the roughness is ≤2 Å. The surface roughness is attributed to a combination of factors, including reconstruction-related stoichiometry differences, kinetically limited diffusion during growth, and lattice-mismatch strain. Possible methods to reduce the roughness are discussed.
© *1998 American Vacuum Society.* [S0734-211X(98)08304-8]

I. INTRODUCTION

A wide range of band alignments are possible between the nearly lattice matched "6.1 Å" family of III–V semiconductors, InAs, GaSb, and AlSb, allowing a variety of novel quantum well and superlattice-based electronic devices to be fabricated from these materials, including infrared detectors, infrared lasers, and high-speed oscillators. For example, resonant tunneling diodes (RTDs) with switching speeds approaching terahertz frequencies can be constructed using thin AlSb layers [5–10 monolayers (ML)] as tunneling barriers between InAs- and GaSb-based layers.[1–3] The properties of these devices are highly dependent on the barrier thickness,[4] as expected given that transport across the barrier depends exponentially on thickness. RTD performance is also expected to be sensitive to atomic-scale variations in the morphology of the interfaces,[5] including fluctuations in film thickness and composition caused by diffusion and/or intermixing during growth. Characterizing, understanding, and ultimately controlling the atomic-scale structure of the interfaces within RTDs will be essential to reproducing device characteristics within the tight tolerances required for integration into high-speed circuits.

Many recent studies have focused on determining the best method for making interfaces between different 6.1 Å layers using molecular beam epitaxy (MBE).[6–10] For example, two different types of interfacial bonds can be formed during the deposition of AlSb on InAs, either AlAs like or InSb like, depending on the III–V deposition sequence. It has been found both experimentally[11–14] and theoretically[15] that InSb-like interfaces are generally more abrupt than those with AlAs-like bonds. For this reason, among others, AlSb-based tunneling devices are usually fabricated with InSb-like interfaces. One approach for preparing an AlSb-on-InAs interface with InSb bonds using MBE would be to: terminate the In and As_2 fluxes; then start the Sb_2 flux, exposing the surface to just Sb_2 for a brief time (typically 5–30 s, i.e., a "growth interrupt"); and then start the deposition of Al to begin AlSb growth. It has been reported that during the Sb_2 exposure the As on the surface is largely replaced by Sb.[16] One possible way to reduce such an anion exchange might be to use migration enhanced epitaxy (MEE) to form the interface, depositing the individual III and V elements sequentially. For example, to deposit AlSb on InAs with an InSb-like interface one can first deposit an additional ML of In (with no As_2 flux), then briefly expose the surface to Sb_2 (with no In or Al flux). After the desired amount of Sb is deposited, AlSb growth is begun. In principle, this procedure should be less conducive to anion exchange, enabling more abrupt, consistent interfaces to be formed.

In this article we describe our scanning tunneling microscopy (STM) study of the surfaces and interfaces formed during the growth of InAs/AlSb/InAs RTD-like structures, focusing on InSb-like interfaces prepared by MEE. InAs substrates were used, rather than InAs buffer layers grown on GaAs substrates, in order to eliminate any artifacts that

[a]Electronic mail: brettn@engineering.ucsb.edu
[b]Electronic mail: Lloyd.Whitman@nrl.navy.mil

might arise as a consequence of the greater roughness of InAs surfaces grown on GaAs. We have examined the effects of various growth procedures on the apparent anion exchange reaction, and the resulting surface morphology of subsequently deposited films. Based on the analysis of our results, we discuss possible routes for improving the quality and consistency of the interfaces in these structures.

II. EXPERIMENT

The experiments were performed at the Naval Research Laboratory in an interconnected, multichamber ultrahigh vacuum facility that includes a III–V MBE chamber equipped with reflection high-energy electron diffraction (RHEED) and a surface analysis chamber equipped with a STM.[17] All films were grown not intentionally doped using "cracked" As_2 and Sb_2 sources on undoped InAs(001) wafers. The growth rates were monitored by RHEED oscillations. After oxide removal, all growths were begun with a ~0.5 μm thick buffer layer of InAs grown with a 5:1 beam equivalent pressure (BEP) ratio of As-to-In at 1 ML/s with 30 s interrupts under As_2 every 90 s. The growth temperature was kept 10 °C below the temperature at which the (2×4)–to–(4×2) transition occurs during InAs growth (estimated to be 490 °C).[18] At the end of the buffer layer growth, the In was shuttered and a 10 min growth interrupt was performed (i.e., the samples were held at the growth temperature while reducing the As_2 flux), during which time the RHEED patterns progressed from a streaky (2×4) to sharp diffraction spots along each streak. The samples were cooled approximately 100 °C after completion of the InAs buffer layer and prior to the growth of AlSb. In order to minimize unwanted deposition leaking from the shuttered sources, the sample surfaces were rotated away from the sources during this time. Once the temperature stabilized, the samples were rotated back to face the sources and all further growths were performed at the lower temperature.

Each image in our study can be considered to be a "snapshot" of the surface at some point during the growth of InAs/AlSb/InAs. A diagram of the structure grown and the surfaces and interfaces examined is displayed in Fig. 1. Note that although the interfaces were prepared sequentially as indicated, *a new sample was prepared from scratch for each interface studied* to eliminate any history effects associated with vacuum contamination, etc. Starting with the InAs buffer layer described above, the first interface was prepared using MEE, with 1 ML of In (1 s) followed by 2 s of Sb_2 at a BEP of $\sim2.5\times10^{-6}$ Torr. This surface will henceforth be referred to as "the MEE-prepared surface." On two different MEE-prepared surfaces, 5 ML of AlSb was then grown at 0.5 ML/s. One sample was immediately cooled after the AlSb deposition (with no Sb_2 flux), and the other was given a 5 min growth interrupt under Sb_2 for comparison. The final layer studied, 22 ML of InAs, was deposited on the interrupt-terminated AlSb layer, and given a 30 s interrupt under Sb_2 after growth. This surface would be the starting point for the second AlSb layer in a double-barrier resonant tunneling structure.

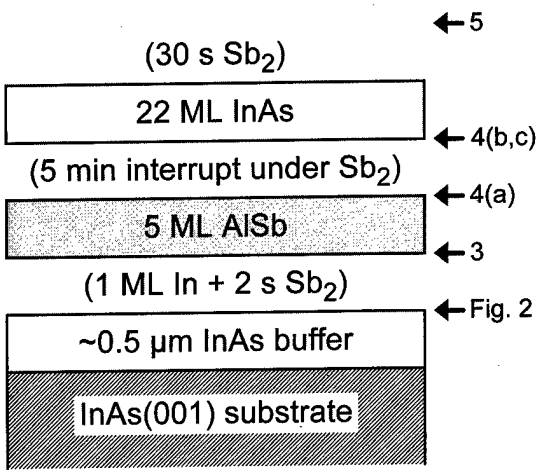


FIG. 1. Schematic diagram of the InAs/AlSb/InAs structure grown by MBE and studied with STM. Surfaces of interest begin with the InAs buffer layer and end at the surface on which a second AlSb barrier layer would be deposited. Notations in parentheses refer to surface treatments applied between the layers. Surfaces that are shown in Figs. 2–5 are indicated by the arrows.

After the completion of each growth, the sample was immediately removed from the MBE chamber and transferred to the STM chamber. The pressure in the transfer section was $<5\times10^{-10}$ Torr, and the entire transfer procedure from the end of the growth until the sample was in the STM chamber typically took <10 min. The samples were allowed to cool further for ~1 h in the STM chamber ($<1\times10^{-10}$ Torr) before imaging. All STM images shown were acquired in constant-current mode with sample biases ranging from −1.5 to −3 V and tunneling currents between 50 pA and 0.5 nA.

III. RESULTS AND DISCUSSION

At the end of the 10 min interrupt the InAs(001) buffer layer is nearly ideal, composed of large (0.1–0.5 μm wide), atomically flat terraces separated by monolayer height (3 Å) steps, as shown in Fig. 2. There are essentially no islands (<1 per 10 μm^2), so the step density is simply determined by the surface orientation—approximately 0.05° from (001). Higher magnification images (not shown) reveal a well-ordered (2×4) surface reconstruction similar in appearance to that observed on As-terminated GaAs(001), consistent with the sharp RHEED spots observed at the end of the growth procedure.

Terminating the InAs(001)-(2×4) surface with InSb bonds via the MEE procedure described above (1 ML In+2 s Sb_2) dramatically changes the nanometer-scale surface structure, as shown in Fig. 3(a). The originally flat surface bifurcates into two levels, with 1 ML deep anisotropic vacancy islands (the lower level) covering 23% of the surface area. An analysis of the two-dimensional height–height autocorrelation function indicates that features on the surface are generally twice as long along $[\bar{1}10]$ than along [110]. A radial distribution analysis of the vacancy islands shows that they are randomly distributed across the surface

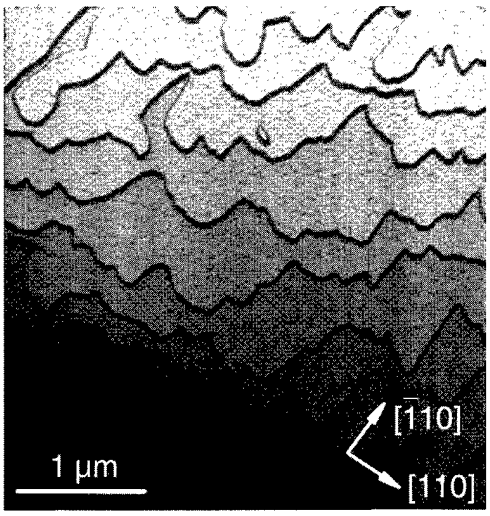


FIG. 2. 3.2 μm×3.2 μm STM image of the InAs(001)−(2×4) buffer layer. Image is displayed as a three-dimensional (3D)-rendered gray scale with a very slight perspective. (This view enhances the 3.0 Å high step edges, which appear dark.)

at a density of 3.7×10^{11} cm^{-2}, with an average area of 42 nm^2 and separation of 17 nm. (These statistics are collated in Table I.) At higher magnification (not shown), a disordered (1×3)-like reconstruction is observed, consistent with the corresponding streaky (1×3) RHEED pattern. The reconstruction has a similar appearance to that observed on InSb,[19] AlSb,[20] and GaSb.[21,22] We believe the atomic-scale structure is similar to that proposed for GaSb(001)-c(2×6),[22,23] which consists of a full 1 ML plane of Sb atoms terminated by $\frac{2}{3}$ ML of Sb in dimer rows [see Fig. 4(d)].

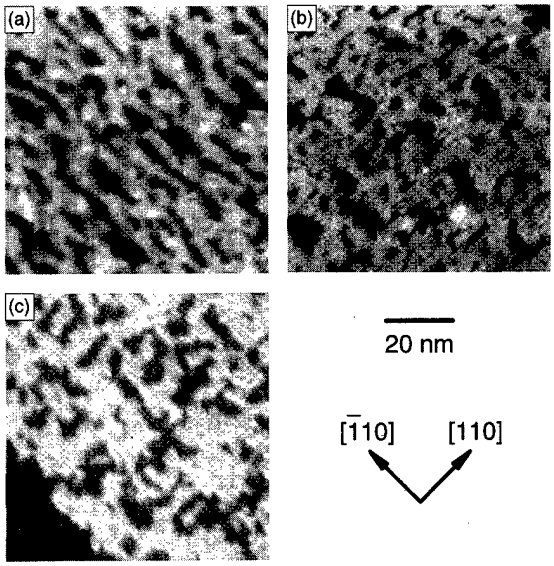


FIG. 3. Filled-state images, 80 nm×80 nm, of an InAs(001)−(2×4) surface exposed to: (a) 1 ML of In followed by 2 s of Sb_2; (b) 2 s of Sb_2; (c) 30 s of Sb_2. In all three images the 3 Å deep vacancy islands cover approximately 25% of the surface.

TABLE I. Statistical comparison of the vacancy islands formed when InSb-like bonds are created on InAs(001) by exposing a (2×4)-reconstructed surface to Sb_2. Island density, average island area, and average island separation are indicated (along with the standard errors) for three different exposures.

Sample	Fraction on lower level	Density (10^{11} cm^{-2})	Average area (nm^2)	Average separation (nm)
1MLIn+2 s Sb_2	23%	3.7±0.5	42±5	17
2 s Sb_2	24%	14.1±0.8	19±4	10
30 s Sb_2	25%	2.9±0.6	58±6	19

As discussed earlier, one approach to changing from InAs to AlSb is to perform a brief growth interrupt under Sb prior to starting Al deposition. For comparison with the MEE procedure, we prepared two InAs surfaces in this way, exposing one to Sb_2 for 2 s and the other for 30 s. Images of these two surfaces are shown in Figs. 3(b) and 3(c), respectively. Although the effects of the Sb exposure are qualitatively similar in all three cases, creating vacancy islands covering about a quarter of the surface, the detailed morphology of the islands varies from case to case (Table I). Whereas the vacancy island density, area, and separation following the 30 s exposure are similar to that following MEE, the island statistics for the 2 s exposure are noticeably different. After the

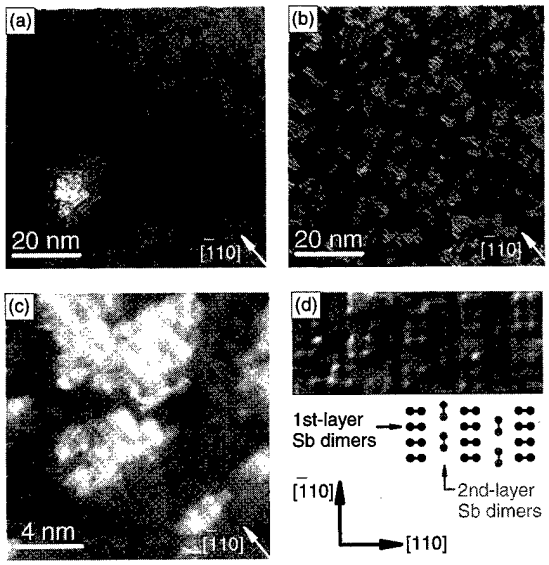


FIG. 4. Filled-state images of 5 ML of AlSb deposited on an InAs surface previously terminated with Sb via MEE [i.e., see Fig. 3(a)]. (a) Surface immediately cooled after growth (shown as a 3D-rendered gray scale, 80 nm×80 nm). The quasi-distinct height levels occur in increments of 1–2 Å (about half the monolayer height). (b) Surface after a 5 min growth interrupt under Sb_2 (also 3D-rendered, 80 nm×80 nm). Islands are of monolayer height, and the wavy corrugation seen on the smaller scale corresponds to the (1×3)-like surface reconstruction. (c) Higher magnification view of (b), 16 nm×16 nm, with atomic-scale features of the surface reconstruction visible. (d) Closeup of the atomically ordered region within the box in (c), gradient enhanced so that the individual atoms within the surface dimers are more easily resolved. Dimer rows separated by 13 Å are seen along the [$\bar{1}$10] direction, as are rotated dimers between the rows in the second layer. Model of the corresponding c(2×6) structure is shown.

shorter dose the island density is nearly an order of magnitude larger, with a much smaller distance between islands (in accord with the total island area which is constant). Dosing with Sb without the In "predose" used in MEE has another interesting effect: the resulting morphology becomes approximately isotropic between the [$\bar{1}$10] and [110] directions. This observation suggests that the Sb adsorption/reaction is relatively isotropic, and that the anisotropy observed after MEE must therefore arise from structural anisotropy in the In prelayer. This would not be surprising given the anisotropy of the As-terminated (2×4) reconstruction.

Even these brief Sb exposures convert the InAs(001)-(2×4) reconstruction to a disordered (1×3)-like structure—characteristic of InSb—implying that Sb is rapidly incorporated into the surface. Although we expect that the MEE-prepared surface is fully Sb-terminated, at this time we have no measure of the actual Sb surface coverages. Under similar experimental conditions in a different MBE facility, the *in situ* threshold photoemission signal from InAs(001)-(2×4) changes dramatically during the first few seconds of Sb_2 exposure, and then recovers to approximately steady state after about 10 s.[24] Collins *et al.*[25] have made similar observations with RHEED. These results imply that after our 2 s exposure the surface may contain a mixture of both Sb and As. The initial signal changes observed in the photoemission and RHEED can be attributed to the disorder rapidly induced by the exchange reaction, with the more gradual recovery associated with the evolution of an Sb-terminated bilayer structure with relative atomic-scale order.

There are a number of possible causes for the bifurcation of the surface that occurs during the formation of an InSb-like interface, including strain and III/V stoichiometry changes. The 7% lattice mismatch between InSb and InAs will result in a compressively strained surface. As seen in the epitaxial growth of InSb/GaAs[26,27] and other III–V heteroepitaxial systems, including In(Ga)As/GaAs[28,29] and GaSb/GaAs,[26,27,30] surface roughening acts as a strain-relief mechanism that helps lower the surface free energy. Thus, the vacancy island formation might be a similar strain-relief phenomenon. One difficulty with this mechanism is the problem with mass balance: the formation of the bilayer in this way requires long-range diffusion of atoms and/or vacancies across the large terraces, of which there is no evidence (e.g., there is no difference in the morphology near the step edges). The second possibility, a change in III/V stoichiometry, is more consistent with our results. Assuming that InAs(001)-(2×4) has the same structure as the analogous GaAs reconstruction, the top two layers of the surface are $\frac{3}{4}$ ML of In covered by $\frac{1}{2}$ ML of As.[31–34] Converting this surface to a $c(2\times6)$ structure composed of 1 ML of In $+\frac{2}{3}$ ML of Sb (or Sb+As) would lead to a surface deficient in In by $\frac{1}{4}$ ML, i.e., a surface with 25% vacancies. Note that such multilayer formation occurs for apparently similar reasons when Ge is deposited on GaAs(001).[35] This stoichiometric effect would also explain why the vacancy island coverage is independent of the way the surface is terminated with Sb. Preliminary results suggest, however, that there are growth procedures that minimize this effect, as will be discussed in detail elsewhere.[36]

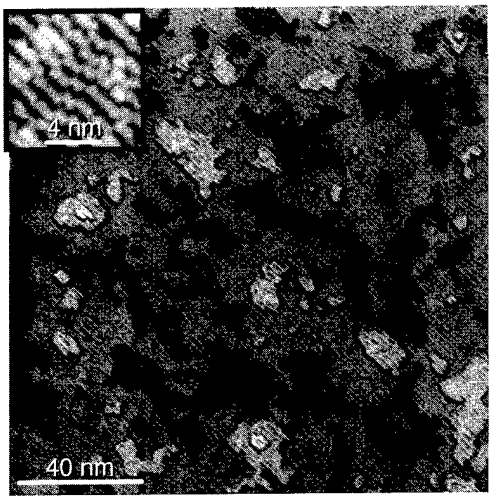


FIG. 5. 160 nm×160 nm 3D-rendered filled-state image of 22 ML of InAs grown on top of an interrupt-terminated AlSb layer [i.e., Fig. 4(b)]. Surface was exposed to Sb_2 for 30 s after growth. (Inset) Atomic-scale view of the (1×3)-like reconstruction observed on the surface.

Moving beyond the formation of the InSb-like interface, we now turn our attention to the morphology of the AlSb barrier layer. Two AlSb films, each 5 ML thick, were grown on an MEE-prepared surface, with one sample immediately cooled after growth and the other given a 5 min growth interrupt under Sb_2. As observed in Fig. 4, the addition of the AlSb increases the number of layers present on the surface. Following the immediate cool-down the surface is relatively rough, with a disordered array of ~1 nm sized features [Fig. 4(a)]. Although the heights of the features occur mostly in increments of 1–2 Å (about half the monolayer height), the levels are not distinct enough for a histogram analysis. In contrast, after the growth interrupt the film exhibits three distinct terrace levels with (1×3)-like atomic-scale order [Figs. 4(b)–4(d)]. Statistical analysis of a number of different images of this surface shows the bottom, middle, and top layers make up 6%, 67%, and 27% of the surface, respectively. Close inspection of regions of this surface with relative atomic-scale order reveal details of the surface reconstruction. In Fig. 4(d) a gradient-enhanced image of such a region is displayed, where the individual atoms within the surface dimers are just resolved. Sb dimer rows running along the [$\bar{1}$10] direction are seen, as are rotated Sb dimers between the rows in the second layer, consistent with the model proposed for III-Sb(001)$-c(2\times6)$ surfaces.[22,23]

The final surface studied was 22 ML of InAs deposited on top of the interrupt-terminated AlSb film. To make this InAs film representative of the next interface in a double–AlSb–barrier structure, its growth was terminated by a 30 s interrupt under Sb_2. As shown in Fig. 5, this Sb-terminated InAs surface has a multilevel morphology, similar to the underlying AlSb surface. Like all the Sb-terminated surfaces, a (1×3)-like reconstruction is observed. Note that there are

TABLE II. Root-mean-square roughness determined for each of the Sb-terminated surfaces. A denotes the surface produced by MEE (alternate In and Sb_2 deposition), and B denotes the surface formed by the addition of 5 ML AlSb followed by a growth interrupt.

Sample	Figure	rms roughness (Å)
1 ML In+2 s Sb_2(A)	3(a)	1.1
2 s Sb_2	3(b)	1.1
30 s Sb_2	3(c)	1.0
A+5 ML AlSb+cool(B)	4(a)	1.5
A+5 ML AlSb+interrupt	4(b)–4(d)	1.6
B+22 ML InAs+30 s Sb_2	5	2.0

many more kinks in the dimer rows on this surface as compared with AlSb, possibly a consequence of the shorter interrupt terminating the growth. Although the number of distinct layers increases to five, the lowest and highest levels together constitute only about 1% of the surface area. The three other levels, from lowest to highest, cover 20%, 67%, and 12% of the surface, respectively.

It is interesting that the morphology of the last InAs surface is noticeably different than that of Fig. 3(c), with island sizes approximately twice as large, even though both were exposed to Sb_2 for 30 s. A variety of factors could contribute to this difference. Because it is more practical to grow an entire structure at the same temperature, the second film was grown at a temperature ~100 °C lower than the buffer layer. Moreover, the second film did not benefit from the long interrupt at the higher temperature given to the buffer layer prior to Sb exposure. At the time Sb exposure begins, these two factors combined are expected to give the second surface a higher density of islands (i.e., more layers) than the nearly ideal buffer layer. Therefore, vacancies created during the short Sb dose will be able to diffuse to and incorporate into existing step edges much more readily than on the very large buffer layer terraces, suppressing the formation of additional vacancy islands.

The root-mean-square (rms) roughness determined for each of the Sb-terminated surfaces is summarized in Table II. The roughness is smallest for the initial bilayer surfaces, and increase as additional layers appear on the surface. The increase in the number of layers present on the growth front may be an indication that the temperature is too low or the growth rates are too high for step-flow. However, given the presence of strain, perfect step-flow may not be possible. Whereas a higher growth temperature and/or growth interrupts may lead to smoother interfaces, this benefit might come at the expense of additional intermixing—an effect known to be detrimental to device properties.[37]

IV. CONCLUSIONS

We have used *in situ* STM to study the evolution of the interfaces formed during the fabrication of an InAs/AlSb/InAs RTD-like structure. InAs(001)-(2×4) buffer layer surfaces were exposed to Sb_2 by several different methods to produce InSb-like bonds at the interface; all methods resulted in a bilayer surface, with vacancy islands covering approximately 25% of the surface. In addition, all the procedures produced a (1×3)-like reconstruction characteristic of an InSb-like surface terminated with >1 ML Sb, indicating that there is a significant amount of Sb on the surface. Starting with an Sb-terminated InAs surface produced by MEE (1 ML In followed by 2 s Sb_2), 5 ML AlSb and then 22 ML InAs/AlSb films were grown. The rms roughness at each subsequent interface was found to increase due to an increase in the number of distinct atomic layers present on the growth surface; however, on all the surfaces the roughness is ≤2 Å. The surface roughness observed is attributed to a combination of factors, including reconstruction-related stoichiometry differences, kinetically limited diffusion during growth, and lattice-mismatch strain.

Finding the optimal point in growth parameter space is a difficult task given the many variables and trade-offs involved. While we have made an initial effort to address some of these issues, it is clear that additional studies are required to isolate and ultimately identify the different factors contributing to the overall surface roughness. It is noteworthy that the interfaces in the RTD structure get rougher with each additional component layer. Thus, improving the initial Sb–InAs interface could lead to a significant reduction in the roughness of the subsequent layers. Further studies that delve into the anion exchange reactions and elucidate how the initial bilayer roughness propagates during the growth, including the effects of kinetic limitations, are required for the development of growth procedures that allow the fabrication of consistently smooth and abrupt interfaces.

ACKNOWLEDGMENTS

The authors would like to thank Dr. D. H. Chow and Dr. J. H. G. Owen for useful discussions throughout this work. This work was supported by the Office of Naval Research; QUEST, a NSF Science and Technology Center for Quantized Electronic Structures (Grant No. DMR 91-20007); DARPA/NSF, Virtual Integrated Prototyping for Epitaxial Growth, Phase I (Grant No. DMS-9615854); and the W. M. Keck Foundation (WHW).

[1]J. S. Scott, J. P. Kaminski, S. J. Allen, D. H. Chow, M. Lui, and T. Y. Liu, Surf. Sci. **305**, 389 (1994).
[2]E. R. Brown, E. R. Söderström, J. R. Parker, L. J. Mahoney, K. M. Molvar, and T. C. McGill, Appl. Phys. Lett. **58**, 2291 (1991).
[3]J. R. Söderström, E. R. Brown, C. D. Parker, L. J. Mahoney, J. Y. Yao, T. G. Andersson, and T. C. McGill, Appl. Phys. Lett. **58**, 275 (1991).
[4]H. Kitabayashi, T. Waho, and M. Yamamoto, Appl. Phys. Lett. **71**, 512 (1997).
[5]P. Roblin, R. C. Potter, and A. Fathimulla, J. Appl. Phys. **79**, 2502 (1996).
[6]M. E. Twigg, B. R. Bennett, and B. V. Shanabrook, Appl. Phys. Lett. **67**, 1609 (1995).
[7]P. M. Thibado, B. R. Bennett, M. E. Twigg, B. V. Shanabrook, and L. J. Whitman, Appl. Phys. Lett. **67**, 3578 (1995).
[8]M. W. Wang, D. A. Collins, T. C. McGill, R. W. Grant, and R. M. Feenstra, Appl. Phys. Lett. **66**, 2981 (1995).
[9]M. W. Wang, D. A. Collins, T. C. McGill, R. W. Grant, and R. M. Feenstra, J. Vac. Sci. Technol. B **13**, 1689 (1995).
[10]B. R. Bennett, B. V. Shanabrook, and E. R. Glaser, Appl. Phys. Lett. **65**, 598 (1994).
[11]K. C. Wong, C. Yang, M. Thomas, and H.-R. Blank, J. Appl. Phys. **82**, 4904 (1997).

[12]J. Schmitz, J. Wagner, F. Fuchs, N. Herres, P. Koidl, and J. D. Ralston, J. Cryst. Growth **150**, 858 (1994).
[13]B. Jenichen, H. Neuroth, B. Brar, and H. Kroemer, Mater. Res. Soc. Symp. Proc. **379**, 493 (1995).
[14]B. Brar, J. Ibbetson, H. Kroemer, and J. H. English, Appl. Phys. Lett. **64**, 3392 (1994).
[15]S.-F. Ren and J. Shen, J. Appl. Phys. **81**, 1169 (1997).
[16]M. W. Wang, D. A. Collins, T. C. McGill, and R. W. Grant, J. Vac. Sci. Technol. B **11**, 1418 (1993).
[17]L. J. Whitman, P. M. Thibado, F. Linker, and J. Patrin, J. Vac. Sci. Technol. B **14**, 1870 (1996).
[18]This temperature was estimated by infrared transmission measurements on an InAs film grown on a GaAs substrate. [See, for instance, B. V. Shanabrook, J. R. Waterman, J. L. Davis, and R. J. Wagner, Appl. Phys. Lett. **61**, 2338 (1992).] These measurements can be sensitive to the growth conditions, surface morphology, etc., and results may vary at different laboratories. For example, a similar experiment in a different lab determined the transition temperature to be ~40 °C higher than we observed.
[19]C. F. McConville, T. S. Jones, F. M. Leibsle, S. M. Driver, T. C. Q. Noakes, M. O. Schweitzer, and N. V. Richardson, Phys. Rev. B **50**, 14965 (1994).
[20]P. M. Thibado, B. R. Bennett, B. V. Shanabrook, and L. J. Whitman, J. Cryst. Growth **175/176**, 317 (1997).
[21]G. E. Franklin, D. H. Rich, A. Samsavar, E. S. Hirschorn, F. M. Leibsle, T. Miller, and T.-C. Chiang, Phys. Rev. B **41**, 12619 (1990).
[22]U. Resch-Esser, N. Esser, B. Brar, and H. Kroemer, Phys. Rev. B **55**, 15401 (1997).
[23]M. T. Sieger, T. Miller, and T.-C. Chiang, Phys. Rev. B **52**, 8256 (1995).
[24]J. J. Zinck (to be published).
[25]D. A. Collins, M. W. Wang, R. W. Grant, and T. C. McGill, J. Vac. Sci. Technol. B **12**, 1125 (1994).
[26]B. R. Bennett, P. M. Thibado, M. E. Twigg, E. R. Glaser, R. Magno, B. V. Shanabrook, and L. J. Whitman, J. Vac. Sci. Technol. B **14**, 2195 (1996).
[27]B. R. Bennett, B. V. Shanabrook, P. M. Thibado, L. J. Whitman, and R. Magno, J. Cryst. Growth **175/176**, 888 (1997).
[28]D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. DenBaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).
[29]V. Bressler-Hill, S. Varma, A. Lorke, B. Z. Nosho, P. M. Petroff, and W. H. Weinberg, Phys. Rev. Lett. **74**, 3209 (1995).
[30]P. M. Thibado, B. R. Bennett, M. E. Twigg, B. V. Shanabrook, and L. J. Whitman, J. Vac. Sci. Technol. A **14**, 885 (1996).
[31]D. J. Chadi, J. Vac. Sci. Technol. A **5**, 834 (1987).
[32]Q. K. Xue, T. Hashizume, T. Sakata, and Y. Hasegawa, Thin Solid Films **281–282**, 556 (1996).
[33]A. R. Avery, C. M. Goringe, D. M. Holmes, J. L. Sudijono, and T. S. Jones, Phys. Rev. Lett. **76**, 3344 (1996).
[34]C. B. Duke, Surf. Sci. **65**, 543 (1993).
[35]X. S. Wang, K. Self, D. Leonard, V. Bressler-Hill, R. Maboudian, W. H. Petroff, and W. H. Weinberg, J. Vac. Sci. Technol. B **11**, 1477 (1993).
[36]B. Z. Nosho, B. V. Shanabrook, B. R. Bennett, W. H. Weinberg, J. J. Zinck, and L. J. Whitman (to be published).
[37]M. J. Yang, W. J. Moore, B. R. Bennett, and B. V. Shanabrook, Electron. Lett. **34**, 270 (1998).

Scanning tunneling microscopy and x-ray photoelectron spectroscopy studies of atomic level structure and Fermi level pinning on GaAs(110) surfaces grown by molecular beam epitaxy

Yasuhiko Ishikawa,[a] Naohiro Tsurumi, Takashi Fukui, and Hideki Hasegawa
Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, N13 W8, Sapporo 060-8628, Japan

(Received 21 January 1998; accepted 26 May 1998)

The relationship between atomic level structure and Fermi level pinning on *n*-type GaAs(110) surface grown by molecular beam epitaxy (MBE) is studied *in situ* by scanning tunneling microscopy (STM)/scanning tunnel spectroscopy (STS) and x-ray photoelectron spectroscopy (XPS). Fermi level by XPS on our MBE-grown *n*-GaAs(110) surfaces is found to lie at around $E_C - 0.5$ eV, indicating the presence of band bending. Among a few characteristic features observed by STM in addition to the basic (1×1) relaxation structure, the dominant holelike structure in the empty state image is correlated with band bending. It shows a gradual decrease of the apparent height over the several atomic distances toward the center in the empty state STM image and gradual increase of the apparent height in the filled state image, indicating the presence of negative charge. It accompanies random site deviations of surrounding atoms. The density of the holelike feature agrees roughly with the required number density of surface pinning centers to cause the observed band bending. This holelike structure originates most likely from a Ga vacancy. However, in STS spectroscopy measurements, a U-shaped continuous gap state distribution extending from the band edges, rather than a peaked discrete deep acceptor state distribution, is observed with a local downward shift of the surface Fermi level. A picture is presented where generation of Ga vacancy introduces local bond disorder and acts like a pinning center having a gap state continuum in accordance with a disorder-induced gap state model. © *1998 American Vacuum Society.* [S0734-211X(98)11004-1]

I. INTRODUCTION

In order to control electronic properties of III-V compound semiconductor surfaces and interfaces, it is important to understand the origin of surface or interface states which generally tend to pin the position of the Fermi level within a certain narrow energy range. Over the past few decades, a number of studies on this so-called Fermi level pinning phenomenon have been carried out, and several models have been proposed concerning the origin and the mechanism of the Fermi level pinning. They include unified defect model (UDM),[1,2] metal-induced gap state (MIGS) model,[3,4] disorder-induced gap state (DIGS) model,[5] and effective work function (EWF) model.[6] However, these models lack direct justification from the microscopic surface/interface structure in atomic scale, and therefore none of them has been persuasive enough over the others.

Based on the microscopic characterization by scanning tunneling microscopy (STM) and scanning tunneling spectroscopy (STS), Pashley *et al.*[7–9] have proposed a model in which kinks in dimer vacancy rows act as single discrete deep surface acceptor sites and cause mid-gap Fermi level pinning on *n*-type GaAs(001)-(2×4) surface grown by molecular beam epitaxy (MBE). This model is hereafter referred to as the kink-acceptor model. It can be regarded as a special case of the UDM, but seems to be the first model based on microscopic observation. However, we have shown recently by detailed STM and x-ray photoelectron spectroscopy (XPS) studies[10] that the kink-acceptor model cannot quantitatively explain the Fermi level pinning on the GaAs (2×4) surface, if one assumes that each kink acts as a single discrete deep acceptor, and have suggested formation of U-shaped gap state continuum from the viewpoint of the DIGS model. It was also shown that the kink-acceptor model is not applicable to InP (2×4) surfaces with kinks.

On the other hand, the (110) surface has been widely used as surfaces for surface science study although they do not have technological importance as possessed by (001) surfaces. This is because clean (110) surfaces can be relatively easily obtained by cleavage in ultrahigh vacuum (UHV) and the surfaces free from surface gap states and free from the Fermi level pinning can be experimentally realized if the cleavage is carefully done. Previous STM works include observation of the (1×1) relaxation structure[11] and study of surface point defects such as vacancies[12,13] and antisite As on a cleaved surface of low temperature (LT) grown GaAs[14] as well as the effect of near surface dopant atoms.[15,16] As for the Fermi level pinning on (110) surfaces, effects of a small amount of metal deposition[17,18] or oxygen exposure[19] to cleaved GaAs(110) surfaces were also investigated previously based on STM and STS measurements. However, unified understanding of the mechanism of the Fermi level pinning does not seem to exist yet.

[a]Present address: Research Institute of Electronics, Shizuoka University, 3-5-1 Johoku, Hamamatsu 432-8011, Japan; electronic mial: royishi@rie.shizuoka.ac.jp

Recently, it has become possible to obtain clean and smooth GaAs(110) surfaces by MBE growth with suitable growth conditions.[20,21] Such growth has been successfully utilized for the fabrication of quantum wire structure on the *in situ* cleaved surface of the quantum well wafer.[20] A structural study using STM was also performed to clarify the kinetics in the MBE growth on the (110) surface.[21]

With these backgrounds, the purpose of this article is to study the relationship between atomic level structure and Fermi level pinning on the MBE-grown n-type GaAs(110) surface in detail by using STM, STS, and XPS. It is an attempt to obtain useful information on the origin of surface states and Fermi level pinning using MBE-grown clean (110) surfaces with band bending.

The Fermi level on our MBE-grown n-GaAs(110) surfaces was found to lie at around $E_C-0.5$ eV, indicating the presence of band bending. Among a few characteristic features observed by STM, in addition to the basic (1×1) relaxation structure, the dominant holelike structure in the empty state STM image was correlated with band bending. It shows the gradual decrease of the apparent height over several atomic distances toward the center in the empty state STM image and the gradual increase of the apparent height in the filled state image, indicating the presence of a negative charge. It accompanies the random site deviation of surrounding atoms. The density of the holelike feature agrees roughly with the required number density of surface pinning centers to cause the observed band bending. This holelike structure most likely originates from a Ga vacancy. However, in tunneling spectroscopy measurements, a U-shaped continuous gap state distribution extending from the band edges rather than a peaked discrete deep acceptor state distribution, was observed with a local downward shift of the surface Fermi level. A picture is presented where the generation of Ga vacancy introduces local bond disorder and acts like a pinning center having a gap state continuum in accordance with the DIGS model.

II. EXPERIMENT

A. Experimental system and MBE growth

A UHV-based multi-chamber system was used. In this system, a solid source MBE chamber, an XPS (Perkin–Elmer PHI 1600) chamber, and a UHV-STM (JEOL JSTM-4600) chamber are connected by a UHV transfer chamber together with other chambers. Details on this system has been described elsewhere.[22,23]

Three sample chips were mounted with In on the sample holder used in this system. Two of these sample chips were cut from an $n^+(>1\times10^{18}\ \mathrm{cm}^{-3})$-GaAs(110) substrate for STM observation and XPS measurement, respectively, and the third chip was cut from an $n^+(>1\times10^{18}\ \mathrm{cm}^{-3})$-GaAs(001) substrate to provide reference for XPS measurement. In the MBE chamber, they were placed next to each other on the holder and underwent exactly the same growth and surface preparation procedure for room temperature measurements. The chip for the STM observation and the other chips for XPS measurements were

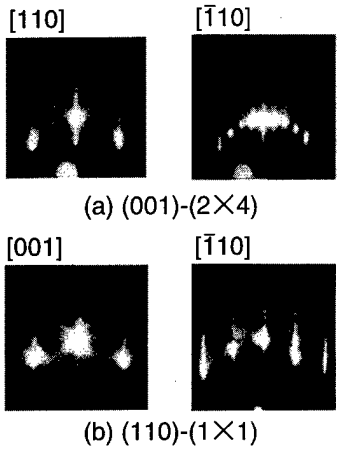


FIG. 1. RHEED patterns for (a) GaAs(001)-(2×4) surface and (b) GaAs(110)-(1×1) surface.

then separated into the two different chambers for STM and XPS. For this purpose, a special sample holder, which was separable in the UHV environment, was used.[10]

For preparation of smooth GaAs(110) surfaces, surface oxides were desorbed from n^+-GaAs substrates in an As_4 flux at 620 °C first, and then thick Si-doped GaAs epitaxial layers (>2000 Å) were grown by the conventional solid source MBE technique. The growth was done at 460–480 °C with a growth rate of 5000 Å/h and the As_4 beam equivalent pressure of $(2\pm1)\times10^{-5}$ Torr. The use of this growth condition led to successful formation of smooth (110) surfaces as described in Ref. 20. Reconstruction patterns were monitored by reflection high-energy electron diffraction (RHEED) with an electron energy of 15 keV. RHEED patterns were (1×1) for (110) surfaces during and after the growth. For (001) surfaces, RHEED patterns appeared diffuse (2×1) during the growth, but changed to $c(4\times4)$ when the supply of Ga and Si was stopped.

After the growth was completed, the samples were heated up to 580 °C and annealed in an As_4 flux for about 10 min before cooling for smoothing the surface on the atomic scale. By this annealing, the (001) surface showed (2×4) β reconstruction[24] with a clear Laue circle, as shown in Fig. 1(a), and (110) surfaces showed a (1×1) pattern with sharp streaks, as shown in Fig. 1(b) where the difference of the spacing of the streaks between [001] and [$\bar{1}$10] incidences is due to the intrinsic difference of the lattice constants along these directions.

Surfaces for room temperature characterizations were prepared by the following procedure in order to prevent adsorption of excess As on the surface. Namely, the sample was cooled gradually reducing the intensity of the As_4 flux in such a way that the (2×4) β reconstruction for the (001) surface was always maintained down to room temperature.[10]

Carrier concentration of the grown layer was determined from capacitance–voltage $(C-V)$ characteristics of Au/GaAs Schottky diodes fabricated by the conventional vacuum evaporation method after STM or XPS characterization of the samples. All the layers grown in the present study

showed n-type conduction. For the (001) layer, the electron concentration n was found to be almost equal to that for the conventional growth at around 580 °C, where n is equal to Si doping concentration N_{Si} up to $5-7\times10^{18}$ cm^{-3}.[10] The (110) layer showed slightly lower electron concentration than that for the (001) layer grown in the same run.

After the samples were cooled down to room temperature, they were rapidly transferred to the STM chamber or to the XPS chamber through the UHV transfer chamber.

B. STM and STS characterizations

STM observations were made at room temperature using electrochemically polished tungsten tips with the constant current mode (0.10–0.30 nA). Both empty and filled state images were taken at various sample voltages.

In addition to the imaging, tunneling spectroscopy measurements were also made in the following way. Namely, after the tip was fixed at the desired position on the surface, tunneling current–voltage ($I-V$) curves were obtained by breaking the STM feedback loop and ramping the sample bias voltage. During $I-V$ measurements, no intentional movement of the tip toward the surface was made as was done previously to increase the current levels at voltages near 0 V.[8,17,18] Instead, ramping of the sample bias voltage was made 64 times at the same fixed position and an average of $I-V$ curves was taken which resulted in the current measurement sensitivity extending over 4 orders of magnitude. The conductivity (dI/dV) was calculated from the $I-V$ characteristic. In order to obtain the state density, the quantity of $(dI/dV)/(I/V)$ was calculated using a one-pole Fourier low-pass filter technique[18] with the pole frequency of $\Delta V^{-1}=1/3$ V^{-1} to prevent the divergence in the normalization when V approaches band edges.

C. XPS characterization

XPS measurements were done at room temperature using monochromatic Al $K\alpha$ radiation ($h\nu=1486.6$ eV) in order to measure the surface Fermi level position. Valence band spectrum as well as several core level spectra such as As 3d and Ga 3d were taken. The binding energies of the spectra were carefully calibrated by separate measurements of the Au $4f_{7/2}$ peak position.

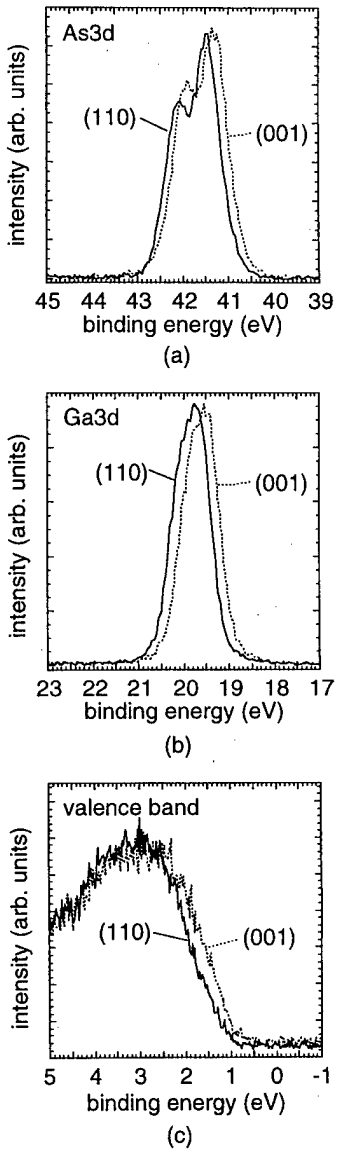


FIG. 2. XPS spectra from (a) As 3d core level, (b) Ga 3d core level, and (c) valence band taken on MBE-grown Si($N_{Si}=5\times10^{17}$ cm^{-3})-doped GaAs(110) and (001) surfaces. Electron concentrations were 2×10^{17} cm^{-3} for (110) surface and 5×10^{17} cm^{-3} for (001) surface, respectively. Solid lines are for (110) surface and dashed lines are for (001) surface.

III. EXPERIMENTAL RESULTS

A. Fermi level position at the MBE-grown n-type GaAs(110) surface

Typical XPS spectra from As 3d core level, Ga 3d core level, and the valence band taken on Si($N_{Si}=5\times10^{17}$ cm^{-3})-doped MBE-grown n-GaAs(110) and (001) surfaces are shown in Fig. 2. The electron concentrations were 2×10^{17} cm^{-3} for the (110) surface and 5×10^{17} cm^{-3} for the (001) surface, respectively. The peak intensities of the spectra were normalized. As mentioned above, these two samples were placed on the same holder.

It was found that the valence band edge as well as As 3d and Ga 3d core level peaks for (110) surface appeared at higher binding energy positions than those for the (001) surface. This corresponds to the difference of the surface Fermi level positions. Figure 3 shows surface Fermi level positions determined from valence band spectra as a function of electron concentration. The surface Fermi levels were found to lie at around $E_C-0.5$ eV for the (110) surfaces and at around $E_C-0.7$ eV for the (001) surfaces. This indicates that the band bending occurs on a macroscopic scale and surface acceptor states exist in the band gap at MBE-grown GaAs(110) surface as well as at the (001) surface, although the amount of the band bending at (110) surfaces was significantly smaller than those at (001) surfaces. This result is quite different from that for the cleaved (110) surface showing the flatband.

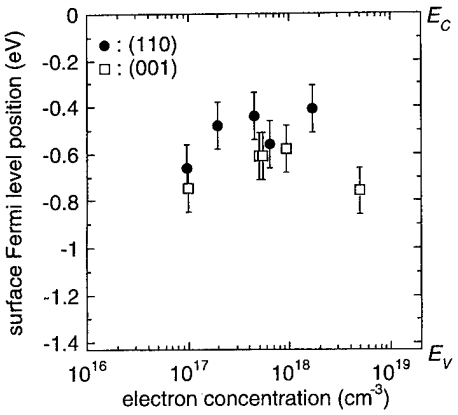


FIG. 3. Surface Fermi level positions measured by XPS as a function of electron concentration.

B. STM observation of the GaAs(110) surface

Prior to presenting the results of atomic scale UHV-STM observations, the results by *ex situ* atomic force microscopy (AFM) in air on the micrometer scale and *in situ* UHV-STM on the hundred nanometer scale are described first in order to show the degree of the flatness of the grown surfaces.

Figure 4(a) shows a typical *ex situ* AFM image of the MBE-grown (110) surface taken on the area of 5 μm×5 μm. Terrace regions, whose width was in the several hundred nanometer range, and steps dividing (110) terraces were observed. Multiatomic step formation (step bunching phenomenon) was also observed partially. Figure 4(b) shows a typical filled state STM image (sample bias; $V_S=-3.0$ V) of the MBE-grown GaAs(110) surface with an area of 100 nm×100 nm. In this region, the (110) terraces were divided by three steps basically running along $[\bar{1}1\bar{2}]$ or $[00\bar{1}]$ directions. From the images taken on the different areas, it was found that the width of the (110) terraces was not constant but random, often and no step was observed on different surface regions within the same observation area of 100 nm×100 nm. The height of the step was also found to be almost equal to 0.4 nm, corresponding to 2 monolayer (ML) height of the (110) surface possibly being related to the bilayer RHEED oscillation reported previously.[25] Taking account of the number of the steps observed on the larger area by AFM and the single step height of 2 ML, the angle of the misorientation of the substrate was estimated to be about 0.1°.

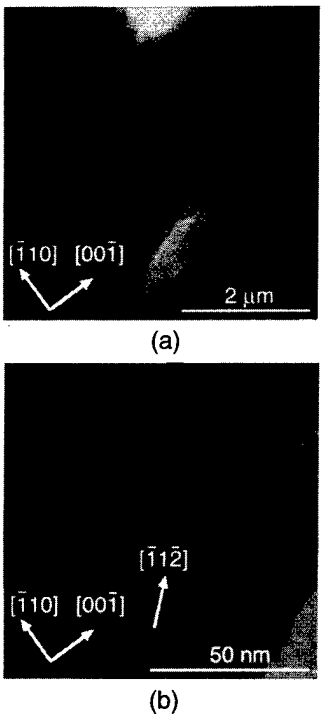


FIG. 4. (a) Typical *ex situ* AFM image of MBE-grown GaAs(110) surface taken on the area of 5 μm×5 μm and (b) typical filled state STM image ($V_S=-3.0$ V) of MBE-grown GaAs(110) surface with the area of 100 nm×100 nm.

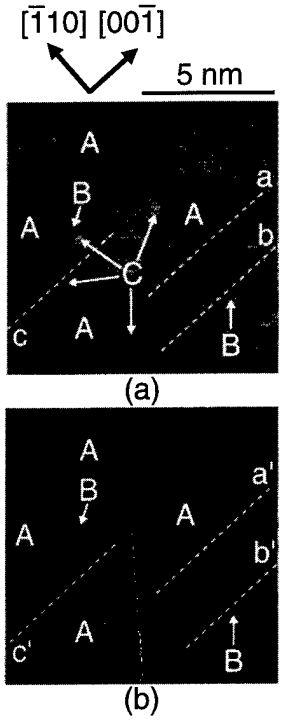


FIG. 5. (a) Empty state ($V_S=3.0$ V) and (b) filled state ($V_S=-3.0$ V) STM images taken on the GaAs(110) terrace region of a sample ($N_{Si}=5.0\times10^{18}$ cm^{-3}, $n=1.7\times10^{18}$ cm^{-3}). Images were taken at the nearly same position of the surface.

As for the atomic scale STM characterization, Figs. 5(a) and 5(b) show typical empty state ($V_S=3.0$ V, Ga sublattice) and filled state ($V_S=-3.0$ V, As sublattice) STM images taken on the (110) terrace region of a sample ($N_{Si}=5.0\times10^{18}$ cm^{-3}, $n=1.7\times10^{18}$ cm^{-3}), respectively. The images were taken at nearly the same position of the surface.

It was found that the surface basically possessed (1×1) regular arrays corresponding to the well-known surface relaxation structure being free from surface states. Although band bending occurred on this surface, kinks and their correlation with Si doping as seen on MBE-grown (001)-(2×4) surfaces[7–9] were not at all observed. This is

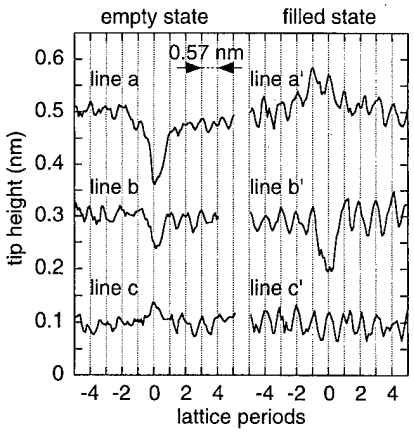


FIG. 6. Line-scan profiles taken along lines *a*, *b*, and *c* in Fig. 5(a) and those taken along the lines *a'*, *b'*, and *c'* in Fig. 5(b).

due to the obvious difference of atomic arrangements between the (110)-(1×1) relaxation structure without missing rows of atoms and the (001)-(2×4) reconstruction structure with missing rows of As dimers.

Instead, on the present (110) surfaces, three different characteristic features indicated by letters A, B, and C were typically observed on the basic (1×1) structures. Line-scan profiles taken along lines *a*, *b*, and *c* in Fig. 5(a) and those taken along lines *a'*, *b'*, and *c'* in Fig. 5(b) are summarized in Fig. 6.

The dominant feature which was most frequently observed is indicated by letter A and looks like a hole in the empty state image with the gradual decrease of the apparent height over several atomic distances toward the center, as shown in the top trace of Fig. 6. In the filled state image, it looks like a protrusion with gradual increase of the apparent height. Feature A was often observed at the step edges, as shown in Fig. 7. It was also found that these apparent change in the height at feature A was enhanced when the observation was done under smaller bias voltages than those used above. These gradual changes in the apparent height were previously interpreted to be directly related to the existence of the charge.[8,12–16] The simplest interpretation of feature A would be Ga vacancy acting as a single deep acceptor with negative charge. In order to further clarify this point, a tunneling spectroscopy measurement was performed at feature A in order to investigate the correlation between feature A and gap states, whose result will be presented in Sec. III C, leading to a more complicated interpretation.

On the other hand, feature B looks like a hole in both empty and filled state images and does not possess the gradual change of the height, as shown in the line-scan profiles in the middle of Fig. 6. Therefore, feature B is not due to the presence of the charge but is likely to be due to the actual hole. It is likely that feature B corresponds to the uncharged complex of As and Ga vacancies. As shown in the line-scan profiles at the bottom of Fig. 6, in the empty state image feature C looks like a protrusion on a lattice point without gradual change of the height, and it disappears in the filled state images. It is thus likely that feature C represents an atom adsorbed on the Ga atom without charge.

Next, a more detailed study was made on the structure and density of the dominant feature A. Examples of line-scan profiles along [00$\bar{1}$] and [$\bar{1}$10] directions taken through feature A and existing at different positions of the surface are shown in Figs. 8(b) and 8(c) together with the STM image (V_S=2.3 V) and the positions of feature A through which the profile was taken. From Figs. 8(b) and 8(c), it was found that the shapes of the profile were similar to each other. The asymmetric feature of the profiles was probably not real but an artifact reflecting the fact that the line-scan was taken from the image-scan data in a direction different from the original imaging scan direction where the scan was done from the left side to the right side of the STM image. Thus, drift and fluctuation in the imaging scan lead to asymmetric line-scan profile.

Figure 9(a) shows the observed average density of feature A as a function of electron concentration *n*. The data points represent averages of the densities determined from more than ten STM images taken on the macroscopically far areas. The image sizes used for the determination of the values were between 20 nm×20 nm and 50 nm×50 nm. The error bars show the rms deviation values of the measured density data. The solid lines in Fig. 9(a) show the number density of ionized donors in the surface depletion region under the surface band bending caused by the indicated surface Fermi level positions.

As seen in Fig. 9(a), the density of feature A increased monotonously with the electron concentration of the grown layer. It was also found that the magnitudes of the observed

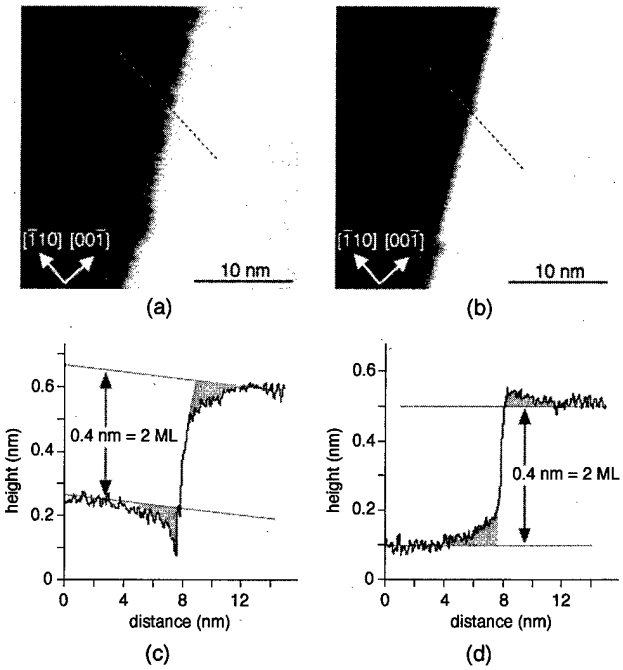


FIG. 7. (a) Empty state (V_S=3.0 V) and (b) filled state (V_S=3.0 V) STM images of GaAs(110) surface including a step. (c) Line-scan profile taken along the line in (a) and (d) line-scan profile taken along the line in (b). Images were taken at the nearly same position of the surface.

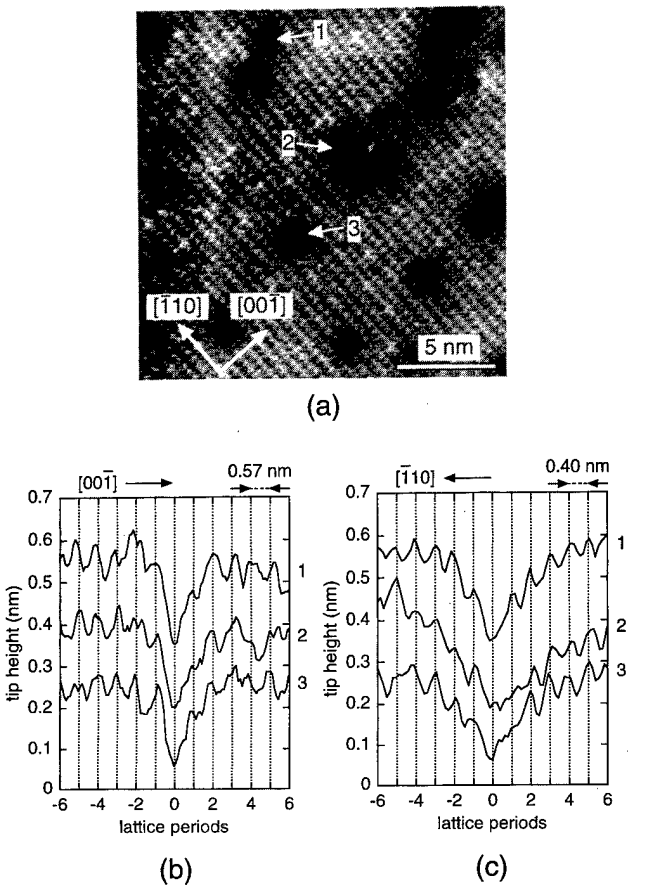


FIG. 8. (a) Empty state (V_S=2.3 V) STM image taken on MBE-grown Si(N_{Si}=5.0×10^{18} cm^{-3}, n=1.7×10^{18} cm^{-3})-doped GaAs(110) surface and examples of line-scan profiles along (b) [001̄] direction and (c) [1̄10] direction taken through feature A. Positions of feature A used for line-scan profiles are indicated in (a).

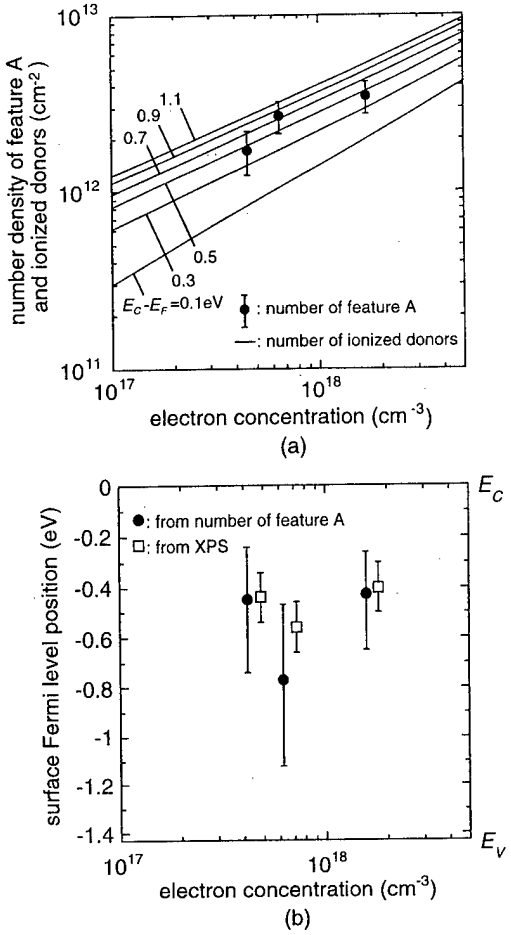


FIG. 9. (a) Number density of feature A and ionized donors as a function of electron concentration and (b) surface Fermi level positions expected from the density of feature A and those directly measured by XPS. The values of electron concentrations were slightly shifted horizontally.

densities of feature A were roughly appropriate to explain the band bending in terms of the surface acceptor model. Namely, the surface Fermi level positions expected from the density of feature A and those directly measured by XPS were compared in Fig. 9(b) where it was assumed that each feature A, acting as a pinning center, receives one electron from each Si donor within the surface depletion layer. For clearer presentation of the data, the values of electron concentrations were slightly shifted horizontally. As seen in Fig. 9(b), the expected and measured surface Fermi level positions agree with each other within accuracy of the measurement.

C. Tunneling spectroscopy

Typical tunneling spectra obtained at the basic (1×1) relaxation region is shown in Fig. 10(a). The estimated conduction and valence band edges, E_C and E_V, are also denoted in the figure. It was found that the surface Fermi level position at the (1×1) relaxation region was estimated to lie at about $E_C-0.4$ eV. This almost agrees with the result by XPS ($E_F \sim E_C-0.5$ eV).

Figures 10(b) and 10(c) show examples of the spectra obtained at the position on feature A. In Fig. 10(b), presence of gap states extending from the conduction band and the valence band were observed at feature A. On the other hand, no gap state was seen at feature A in Fig. 10(c). This result is very similar to a previous study on the kinks on the GaAs (2×4) surface.[8]

At feature A the spectrum was also found to be shifted up in sample voltage. This corresponds to the increase of the band bending and the local downward shift of the surface Fermi level position.

IV. DISCUSSION

Since the dominant feature A seems to be the dominant pinning center in the present study, the discussion is focused on the origin and pinning mechanism by this feature. The gradual change in the apparent height was previously interpreted to be due to the existence of the excess charge at the center of the feature.[8,12–16] Therefore, feature A is most likely related to the surface acceptor gap states which are negatively charged, receiving an electron from the Si donor within the surface depletion layer. It should be noted that

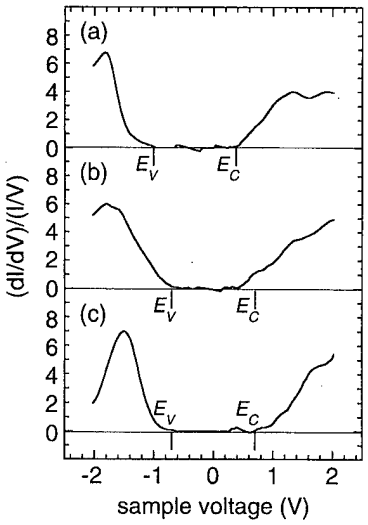


FIG. 10. Normalized conductivity vs sample voltage obtained at (a) basic (1×1) region, (b) the position on feature A showing the states extending from the conduction band and the valence band, and (c) the position on feature A showing no gap states. The estimated conduction and valence band edges, E_C and E_V, are also denoted.

feature A is not due to the presence of the positively charged Si donor itself near the surface region since the effect of the Si dopant appears as a protrusion rather than a hole in the empty state STM observation.[16]

The presence of a negative charge in the acceptor state modifies the local band bending in such a way to cause an upward shift of the density of empty states in the conduction band and to reduce the tunneling current from the tip to the empty state of the semiconductor, if the distance between the STM tip and the sample surface is maintained. Thus, when the constant current mode is used, the tip has to approach nearer the surface so that the effect of the Coulomb potential by the negative charge is canceled. This explains the locally observed gradual height change near feature A.

Thus, as already mentioned, the simplest interpretation is that the observed characteristic feature A corresponds to a Ga vacancy which acts as a single surface acceptor state with a deep discrete energy level and causes band bending through electron capture. This is also consistent with the report that the Ga vacancy acts as an acceptor to compensate the free electron in bulk GaAs.[26] Therefore, this seems to explain the observed behavior of the apparent heights. This interpretation by Ga vacancy was previously introduced by Cox *et al.*[13] and Lengel *et al.*,[12] and seems to be a special example of a UDM for the Fermi level pinning proposed by Spicer *et al.*[1]

However, with such a simple model, it is difficult to understand why no discrete level around the mid-gap was observed at feature A in the present STS measurement. Only a U-shaped gap state continuum was detected. This is similar to the previous STS study on the kinks on GaAs(001)-(2 $\times4$) surfaces.[8]

The fact that feature A remains as a protrusion in the filled state image seems also contrary to the simple acceptor model. This is because, if feature A represents the usual deep acceptor, it should give up the electron up to the tip under the strong positive tip bias, and the feature should vanish. If feature A represents, on the other hand, a fixed negative charge, it is possible to explain its protrusion under the strong positive tip bias in terms of the presence of the local upward shift of the surface potential. However, the observed gap state continuum remains difficult to explain, since a fixed charge should not give any density of states.

The As antisite defects used in the advanced UDM (AUDM)[2] or the As cluster used in the effective work function (EWF)[6] model do not seem to correspond to the present feature A, because they cannot be seen as an apparent hole in the empty state STM observation according to the STM observation of the As antisite defects near the cleaved (110) surface of the LT-GaAs(001) layer.[14]

Taking into account the result that the U-shaped gap states, extending from the conduction band and the valence band, were observed at feature A, the following interpretation based on the DIGS model[5] can be offered, which explains the various features of A observed by STM and XPS. Namely, feature A basically originates from Ga vacancy, but around feature A, slight random deviation of the bond length and/or the bond angle from those of the regular relaxation structure exists due to absence of the Ga atom at the center of feature A. Then, such surface disorder introduces a continuous U-shaped gap state spectra involving both acceptor and donor states in the band-gap region. Although the gap states around the surface Fermi level (zero sample voltage) could not be detected in this study, we believe that it was due to low sensitivity and rather rapidly ramped nature of our STS measurement and the gap state continues even in the middle of the band-gap region. Such the U-shaped gap state continuum restricts movement of the local Fermi level under both positive and negative biases at around the charge neutrality level or the sp^3 hybrid orbital energy.[5] Then, this limitation of the movement of the Fermi level causes the almost constant upward band bending under any bias conditions. Therefore, feature A can be observed under both bias polarities. The observed macroscopic pinning position of 0.4–0.5 eV is determined through electrostatic balance between the disorder-induced pinning centers of feature A with a gap state continuum and the gap state free (1×1) surface relaxation regions. The experiment indicated that each feature A with its surroundings receives, on average, one electron or so into its gap state contiuum.

Finally the reason why the number density of the Ga vacancy-derived feature A increases with the electron concentration may be explained in terms of the strongly Fermi level dependent defect formation energy of Ga vacancy proposed by Walukiewicz.[26]

V. CONCLUSIONS

Atomic level structure and Fermi level pinning on MBE-grown the n-GaAs(110) surface were studied *in situ* by STM, STS, and XPS. The main conclusions are listed below.

(1) The surface Fermi levels lay at around $E_C - 0.5$ eV for (110) surfaces and at around $E_C - 0.7$ eV for (001) surfaces, indicating the presence of the band bending on the MBE-grown n-GaAs(110) surface.
(2) The dominant feature observed on the surface looks like a hole in the empty state STM image with a gradual decrease of the apparent height over several atomic distances toward the center. In the filled state image, it looks like a protrusion with a gradual increase of the apparent height toward the center. The density of the feature agrees roughly with the required number density of acceptor type pinning centers to explain the observed band bending.
(3) Locally at the dominant feature, tunneling spectroscopy measurements detected a continuous U-shaped gap state continuum extending from the band edges with downward shift of the surface Fermi level.
(4) The observed properties of the dominant defect structure cannot be explained by a simple model in which they are due to Ga vacancy acting as a single discrete deep acceptor. Instead, they can be explained by the DIGS model where Ga vacancy, accompanying surface disorder such as random deviation of the bond length and/or the bond angle from those for the relaxation structure, leads to a U-shaped gap state continuum, causing Fermi level pinning under both positive and negative tip bias of STM.

ACKNOWLEDGMENTS

The authors would like to thank Associate Professor Junichi Motohisa and Associate Professor Tamotsu Hashizume for their valuable discussions. One of the authors (Y.I.) was financially supported by a Research Fellowship of the Japan Society for the Promotion of Science (JSPS).

[1]W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979); W. E. Spicer, I. Lindau, P. Skeath, and C. Y. Su, *ibid.* **17**, 1019 (1980).
[2]W. E. Spicer, Z. Liliental-Weber, E. Weber, N. Newman, T. Kendelewicz, R. Cao, C. McCants, P. Mahowald, K. Miyano, and I. Lindau, J. Vac. Sci. Technol. B **6**, 1245 (1988).
[3]V. Heine, Phys. Rev. **138**, 1689 (1965).
[4]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[5]H. Hasegawa and H. Ohno, J. Vac. Sci. Technol. B **4**, 1130 (1986).
[6]J. M. Woodall and J. L. Freeouf, J. Vac. Sci. Technol. **19**, 794 (1981).
[7]M. D. Pashley and K. W. Haberern, Phys. Rev. Lett. **67**, 2697 (1991).
[8]M. D. Pashley, K. W. Haberern, and R. M. Feenstra, J. Vac. Sci. Technol. B **10**, 1874 (1992).
[9]M. D. Pashley, K. W. Haberern, R. M. Feenstra, and P. D. Kirchner, Phys. Rev. B **48**, 4612 (1993).
[10]Y. Ishikawa, T. Fukui, and H. Hasegawa, J. Vac. Sci. Technol. B **15**, 1163 (1997).
[11]R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987).
[12]G. Lengel, R. Wikins, G. Brown, and M. Weimer, J. Vac. Sci. Technol. B **11**, 1472 (1993); G. Lengel, R. Wikins, G. Brown, M. Weimer, J. Gryko, and R. E. Allen, Phys. Rev. Lett. **72**, 836 (1994).
[13]G. Cox, K. H. Graf, D. Szynka, U. Poppe, and K. Urban, Vacuum **41**, 591 (1990).
[14]R. M. Feenstra, J. M. Woodall, and G. D. Pettit, Phys. Rev. Lett. **71**, 1176 (1993).
[15]M. B. Johnson, O. Albrektsen, R. M. Feenstra, and H. W. M. Salemink, Appl. Phys. Lett. **63**, 2923 (1993).
[16]J. F. Zheng, X. Liu, N. Newman, E. R. Weber, D. F. Ogletree, and M. Salmeron, Phys. Rev. Lett. **72**, 1490 (1994).
[17]R. M. Feenstra, Phys. Rev. Lett. **63**, 1412 (1989).
[18]R. M. Feenstra and P. Mårtensson, Phys. Rev. Lett. **61**, 447 (1988); P. Mårtensson and R. M. Feenstra, Phys. Rev. B **39**, 7744 (1989).
[19]J. A. Stroscio, R. M. Feenstra, and A. P. Fein, Phys. Rev. Lett. **58**, 1668 (1987).
[20]L. Pfeiffer, K. W. West, H. L. Stormer, J. P. Eisenstein, K. W. Baldwin, D. Gershoni, and J. Spector, Appl. Phys. Lett. **56**, 1697 (1990).
[21]D. M. Holmes, J. G. Belk, J. L. Sudijono, J. H. Neave, T. S. Jones, and B. A. Joyce, Surf. Sci. **341**, 133 (1995).
[22]B. X. Yang, Y. Ishikawa, T. Ozeki, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **35**, 1267 (1996).
[23]Y. Ishikawa, T. Fukui, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **36**, 1749 (1997).
[24]H. H. Farrell and C. J. Palmstrøm, J. Vac. Sci. Technol. B **8**, 903 (1990).
[25]D. M. Holmes, J. G. Belk, J. L. Sudijono, J. H. Neave, T. S. Jones, and B. A. Joyce, Appl. Phys. Lett. **67**, 2848 (1995).
[26]W. Walukiewicz, Appl. Phys. Lett. **54**, 2094 (1989).

Atomic-scale compositional structure of InAsP/InP and InNAsP/InP heterostructures grown by molecular-beam epitaxy

S. L. Zuo, W. G. Bi,[a] C. W. Tu, and E. T. Yu[b]
Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California 92093-0407

(Received 21 January 1998; accepted 26 May 1998)

Cross-sectional scanning tunneling microscopy (STM) has been used to characterize the atomic-scale structure of $InAs_{0.35}P_{0.65}$/InP and $InN_{0.01}As_{0.35}P_{0.64}$/InP strained-layer multiple quantum well structures grown by gas-source molecular-beam epitaxy. Atomically resolved STM images of the (110) cross-sectional plane reveal nanoscale clustering within the $InAs_xP_{1-x}$ alloy layers, with the boundaries between As-rich and P-rich regions in the alloy layers appearing to be preferentially oriented along the $[\bar{1}12]$ and $[1\bar{1}2]$ directions in the (110) plane. $(1\bar{1}0)$ cross-sectional images reveal that considerably less compositional variation appears within the $(1\bar{1}0)$ plane; features elongated along the [110] direction are observed, but few $\langle 112\rangle$ boundaries are seen. These observations suggest that the boundaries between As-rich and P-rich clusters may form preferentially within the $(1\bar{1}1)$ and $(\bar{1}11)$ planes. Comparisons of filled-state images of $InAs_xP_{1-x}$/InP and $InN_xAs_yP_{1-x-y}$/InP heterostructures suggest that N incorporation increases the valence-band offset in $InN_xAs_yP_{1-x-y}$/InP compared to that in $InAs_xP_{1-x}$/InP. © *1998 American Vacuum Society.* [S0734-211X(98)11104-6]

I. INTRODUCTION

The $InAs_xP_{1-x}$/InP and, more recently, $InN_xAs_yP_{1-x-y}$/InP heterostructure material systems have shown considerable promise for lasers and other optoelectronic devices operating at 1.06, 1.3, and 1.55 μm.[1–4] For $InAs_xP_{1-x}$/InP quantum well structures, the compressive strain in the $InAs_xP_{1-x}$ layer leads to a smaller valence-band effective mass that facilitates population inversion in lasers.[5] Furthermore, the large conduction-band offset in this material system[6] ($\Delta E_c \cong 0.75\Delta E_g$) leads to efficient electron confinement and reduced leakage current in laser diodes, thereby minimizing the threshold current in $InAs_xP_{1-x}$/InP lasers.[1] Finally, the composition in the $InAs_xP_{1-x}$/InP system is easier to control than that in the $In_xGa_{1-x}As_yP_{1-y}$/InP quaternary system, which has been explored extensively for optoelectronic device applications at wavelengths of 0.98–1.55 μm.[2,7] More recently, $InN_xAs_yP_{1-x-y}$ alloys have generated considerable interest, because incorporation of N at low concentration into the $InAs_xP_{1-x}$ alloy layers has been shown to produce a substantial decrease in band gap,[8] can partially compensate for strain due to As present in the alloy, and may possibly increase the conduction-band offset even further.[9] Above room-temperature lasing has been realized in $InN_xAs_yP_{1-x-y}$/$In_xGa_{1-x}As_yP_{1-y}$ quantum well microdisk lasers, with the improved performance compared to $In_xGa_{1-x}As$/$In_xGa_{1-x}As_yP_{1-y}$ quantum well lasers possibly due to an enhanced conduction-band offset coming from nitrogen incorporation.[8]

A significant concern in ternary and quaternary III–V semiconductor alloys is the possible presence of ordering, clustering, and/or compositional modulation, phenomena that have been observed to occur in a wide range of III–V material systems[10–12] and that can exert considerable influence on crystal quality, interface quality, and other electronic as well as optoelectronic properties such as band gap,[13] band-edge discontinuities, and transport properties.[14–17] In addition, the $InN_xAs_yP_{1-x-y}$/InP material system is of considerable interest because the effects of N incorporation on material properties are not clearly understood, but appear to differ considerably from the effects of alloying with other group V elements.[18,19] Detailed characterization and understanding of atomic-scale compositional variations and the effects of N incorporation on compositional structure and band offsets are therefore of great importance for optoelectronic and electronic devices based on these materials.

In this article, we describe scanning tunneling microscope (STM) studies of pseudomorphic $InAs_{0.35}P_{0.65}$/InP and $InN_{0.01}As_{0.35}P_{0.64}$/InP multiple quantum well heterostructures grown by gas-source molecular-beam epitaxy (MBE) on (001) InP substrates. Characterization of $InAs_xP_{1-x}$/InP and $InN_xAs_yP_{1-x-y}$/InP in these studies has provided detailed information about heterojunction interface properties, nanoscale clustering in $InAs_xP_{1-x}$ alloys, and the qualitative influence of N incorporation on heterojunction band alignments for $InN_xAs_yP_{1-x-y}$/InP.

II. EXPERIMENT

The samples used for these studies were grown by gas-source MBE. A 2500 Å InP buffer layer was grown initially on a (001) n^+ InP substrate, followed by a coherently strained heterostructure consisting of 50 Å $InAs_{0.35}P_{0.65}$ alternating with 100 Å InP for five periods, and then five periods of 50 Å $InN_{0.01}As_{0.35}P_{0.64}$ alternating with 200 Å InP. All epitaxially grown layers were doped n type (n

[a]Present address: Hewlett-Packard Company, 3500 Deer Creek Rd., MS 26M-7, Palo Alto, CA 94304.
[b]Electronic mail: ety@ece.ucsd.edu

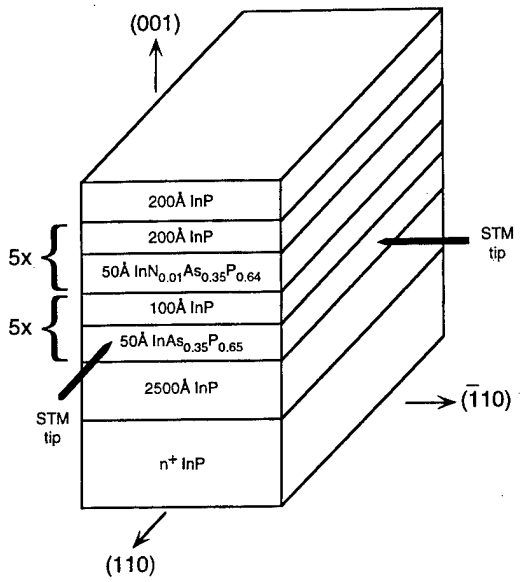


FIG. 1. Schematic diagram of the sample structure and STM geometry used in these studies. STM imaging is performed on both (110) and ($\bar{1}$10) cross-sectional planes.

$\sim 10^{16}$–10^{17} cm^{-3}). The substrate temperature during growth was 460 °C; other details concerning the growth chamber and procedures are described elsewhere.[20,21] Figure 1 shows a schematic diagram of the sample structure and STM geometry employed in our work. STM studies were performed on both (110) and ($1\bar{1}0$) cross-sectional surfaces exposed by *in situ* cleaving in an ultrahigh vacuum (UHV) STM chamber at a pressure of $\sim(7-9)\times 10^{-11}$ Torr. Electrochemically etched W tips cleaned *in situ* by electron bombardment were used for these studies.

III. RESULTS AND DISCUSSION

Figure 2(a) shows a 205 Å×205 Å (110) constant-current STM image of the $InAs_xP_{1-x}$/InP multiple quantum well structure, obtained at a sample bias of −2.4 V and a tunneling current of 0.1 nA. Electronically induced contrast between the $InAs_xP_{1-x}$ and InP layers is clearly visible. Because the valence-band edge of InAs is higher than that of InP, we interpret the brighter features as being associated with As and darker features with P within the $InAs_xP_{1-x}$ layer. Variations in composition at the nanometer scale are clearly visible in the $InAs_xP_{1-x}$ alloy layer in Fig. 2(a), allowing us to investigate in detail the nature of clustering in $InAs_xP_{1-x}$. From Fig. 2(a), it is apparent that within the $InAs_xP_{1-x}$ alloy layer there exist brighter As-rich clusters and darker P-rich clusters, as indicated by the arrows. The clustering of As and P within the $InAs_xP_{1-x}$ alloy layers leads to a marked asymmetry in interface quality in the (110) plane—the InP-on-$InAs_xP_{1-x}$ interfaces are considerably rougher and less abrupt than the $InAs_xP_{1-x}$-on-InP interfaces. In Fig. 2(a), dotted lines delineate two As-rich clusters, which appear to possess approximately triangular cross

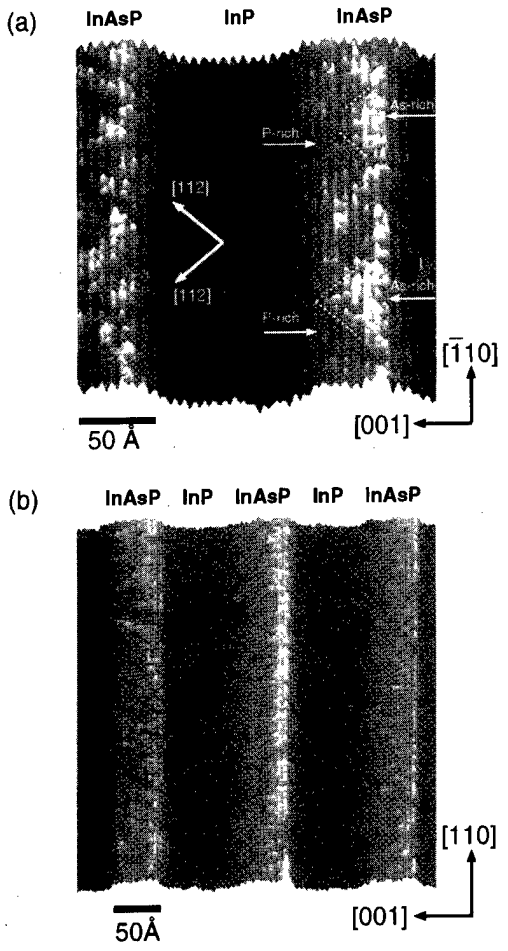


FIG. 2. (a) Three-dimensional rendering of a 205 Å×205 Å (110) constant-current STM image of the $InAs_{0.35}P_{0.65}$/InP multiple quantum well structure, obtained at a sample bias voltage of −2.4 V and a tunneling current of 0.1 nA. Major directions are indicated by arrows. Two triangular As-rich regions, bounded by dotted lines, and two P-rich regions are indicated. (b) Three-dimensional rendering of a 400 Å×400 Å constant-current ($1\bar{1}0$) STM image of the $InAs_{0.35}P_{0.65}$/InP multiple quantum well structure, obtained at a sample bias voltage of −2.4 V and a tunneling current of 0.1 nA.

sections in the (110) plane with bases along the [$1\bar{1}0$] direction and sides appearing to be preferentially oriented along the [$\bar{1}12$] and [$1\bar{1}2$] directions.

If the boundaries between As-rich and P-rich regions assume the form of simple planes, then the observation that the intersections of the (110) plane and the boundaries between As-rich and P-rich regions within the $InAs_xP_{1-x}$ layers are oriented preferentially along the [$\bar{1}12$] and [$1\bar{1}2$] directions implies that the boundary plane indices *(hkl)* should satisfy the equation

$$\mp(h-k)+2l=0. \tag{1}$$

The simplest solutions to Eq. (1) correspond to ($\bar{1}11$) and ($1\bar{1}1$) planes in the crystal, suggesting that the boundaries between As-rich and P-rich regions may form preferentially within these planes.

Imaging of the ($\bar{1}10$) cross-sectional plane provides further information about the compositional structure present

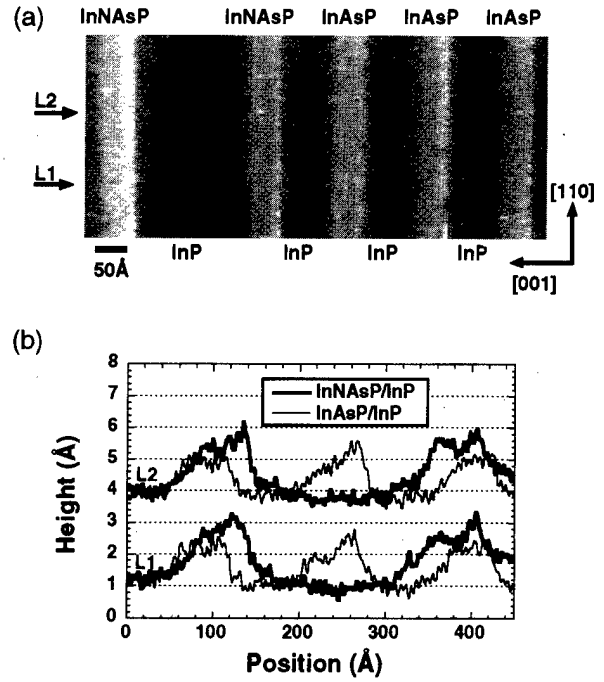


FIG. 3. (a) 860 Å×370 Å constant-current $(1\bar{1}0)$ STM image of $InAs_{0.35}P_{0.65}/InP$ and $InN_{0.01}As_{0.35}P_{0.64}/InP$ multiple quantum well structures, obtained at a sample bias voltage of −2.4 V and a tunneling current of 0.1 nA. (b) Line scans extracted from locations *L1* and *L2* indicated in (a). $InN_{0.01}As_{0.35}P_{0.64}/InP$ and $InAs_{0.35}P_{0.65}/InP$ profiles are superimposed to compare the topographic contrast observed for each material system.

within the $InAs_xP_{1-x}$ alloy layers. Figure 2(b) shows a 400 Å×400 Å $(\bar{1}10)$ STM image obtained at a sample bias of −2.4 V and a tunneling current of 0.1 nA. In Fig. 2(b), it is apparent that the As composition is graded in the [001] growth direction within the lower two $InAs_xP_{1-x}$ alloy layers: the image contrast is brighter (higher) near the $InAs_xP_{1-x}$-on-InP interfaces and darker (lower) near the InP-on-$InAs_xP_{1-x}$interfaces, corresponding to higher and lower As compositions, respectively. Furthermore, the $InAs_xP_{1-x}$ alloy layers show considerably less compositional variation in the [110] lateral direction than was evident along the $[1\bar{1}0]$ lateral direction in the (110) image, and ⟨112⟩ boundaries between clusters are not evident. These observations suggest that the As-rich and P-rich clusters tend to be preferentially elongated in the [110] direction, with approximately triangular cross sections in the (110) plane.

We have also obtained filled-state images of both the $InAs_xP_{1-x}/InP$ and the $InN_xAs_yP_{1-x-y}/InP$ heterostructures; these images provide information about the effect of N incorporation on the valence-band alignment in these material systems. Figure 3(a) shows a $(1\bar{1}0)$ filled-state STM image obtained at a sample bias of −2.4 V and a tunneling current of 0.1 nA. Figure 3(b) shows topographic line scans, each averaged across one atomic spacing, extracted from locations indicated in Fig. 3(a). From the image and the accompanying topographic line profiles, it is apparent that the contrast between the $InN_xAs_yP_{1-x-y}$ and InP layers is greater than that observed between the $InAs_xP_{1-x}$ and InP layers. A direct comparison of topographic contrast between these materials is valid since the comparison is made within the same image and therefore under identical tip and sample conditions. In Fig. 3(b), for each extracted line scan, the contrast between the $InN_xAs_yP_{1-x-y}$and the adjacent InP layer is generally 2 Å or more, while the contrast between $InAs_xP_{1-x}$ and InP layers is usually about 1.5–1.8 Å or less. These observations may indicate that the valence-band offset at the $InN_xAs_yP_{1-x-y}/InP$ heterojunction interface is somewhat greater than that for $InAs_xP_{1-x}/InP$.

IV. CONCLUSION

We have performed atomically resolved constant-current cross-sectional STM imaging of $InAs_xP_{1-x}/InP$ and $InN_xAs_yP_{1-x-y}/InP$ strained-layer multiple quantum well structures grown by gas-source MBE. These studies have revealed the presence of nanoscale clustering of As and P within the $InAs_xP_{1-x}$ alloy layers. A clear asymmetry in interface quality is visible in the (110) plane, with the $InAs_xP_{1-x}$-on-InP interfaces being considerably smoother than the InP-on- $InAs_xP_{1-x}$ interfaces. The boundaries between As-rich and P-rich regions in the $InAs_xP_{1-x}$ alloy layers are oriented preferentially in $[\bar{1}12]$ and $[1\bar{1}2]$ directions in (110) images, forming As-rich regions with approximately triangular (110) cross sections. In $(1\bar{1}0)$ cross-sectional images, we observe that the As composition in $InAs_xP_{1-x}$ layers varies much less in the lateral direction than in the (110) image. These observations suggest that within the $InAs_xP_{1-x}$ layers, the boundaries between As-rich and P-rich regions may form preferentially within the $(\bar{1}11)$ and $(1\bar{1}1)$ planes, and that the As-rich and P-rich clusters tend to be elongated along the [110] direction. Finally, $(1\bar{1}0)$ cross-sectional filled-state imaging of both $InAs_xP_{1-x}/InP$ and $InN_xAs_yP_{1-x-y}/InP$ heterostructures shows that N incorporation into the $InAs_xP_{1-x}$ layer increases the valence-band offset at the $InN_xAs_yP_{1-x-y}/InP$ interface compared to that for $InAs_xP_{1-x}/InP$.

ACKNOWLEDGMENTS

Part of this work was supported by DARPA (Optoelectronics Technology Center) and NSF (ECS 95-07986). E.T.Y. would like to acknowledge receipt of a Sloan Research Fellowship.

[1]T. Fukushima, A. Kasukawa, M. Iwase, T. Namegaya, and M. Shibata, IEEE Photonics Technol. Lett. **5**, 117 (1993).
[2]T. K. Woodward, T.-H. Chiu, and T. Sizer II, Appl. Phys. Lett. **60**, 2846 (1992).
[3]H. Q. Hou, A. N. Cheng, H. H. Wieder, W. S. C. Chang, and C. W. Tu, Appl. Phys. Lett. **63**, 1833 (1993).
[4]J. F. Carlin, A. V. Syrbu, C. A. Berseth, J. Behrend, A. Rudra, and E. Kapon, 1997 IEEE 9th International Conference on Indium Phosphide and Related Materials, Cape Cod, MA, 11–15 May, 1997, p. 563.
[5]E. Yablonovitch and E. O. Kane, J. Lightwave Technol. **6**, 1292 (1988).
[6]M. Beaudoin, A. Bensaada, R. Leonelli, P. Desjardins, R. A. Masut, L. Isnard, A. Chennouf, and G. L'Espérance, Phys. Rev. B **53**, 1990 (1996).
[7]H. Q. Hou and C. W. Tu, J. Cryst. Growth **120**, 167 (1992).
[8]W. G. Bi, Y. Ma, J. P. Zhang, L. W. Wang, S. T. Ho, and C. W. Tu, IEEE Photonics Technol. Lett. **9**, 1072 (1997).
[9]L. Bellaiche, S.-H. Wei, and A. Zunger, Phys. Rev. B **54**, 17568 (1996).

[10]H. R. Jen, D. S. Cao, and G. B. Stringfellow, Appl. Phys. Lett. **54**, 1890 (1989).
[11]H. R. Jen, K. Y. Ma, and G. B. Stringfellow, Appl. Phys. Lett. **54**, 1154 (1989).
[12]Y.-E. Ihm, N. Otsuka, J. Klem, and H. Morkoç, Appl. Phys. Lett. **51**, 2013 (1987).
[13]K. A. Mäder and A. Zunger, Appl. Phys. Lett. **64**, 2882 (1994).
[14]G. B. Stringfellow and G. S. Chen, J. Vac. Sci. Technol. B **9**, 2182 (1991).
[15]A. Zunger and S. Mahajan, *Handbook on Semiconductors, Vol. 3B, Materials, Properties and Preparation*, edited by S. Mahajan (North-Holland, Amsterdam, 1994), p. 1399.
[16]P. K. Bhattacharya and J. W. Ku, J. Appl. Phys. **58**, 1410 (1985).
[17]A. Gomyo, T. Suzuki, and S. Iijima, Phys. Rev. Lett. **60**, 2645 (1988).
[18]L. Bellaiche, S.-H. Wei, and A. Zunger, Phys. Rev. B **56**, 10233 (1997).
[19]W. G. Bi and C. W. Tu, Appl. Phys. Lett. **70**, 1608 (1997).
[20]T. P. Chin, B. W. Liang, H. Q. Hou, M. C. Ho, C. E. Chang, and C. W. Tu, Appl. Phys. Lett. **58**, 254 (1991).
[21]H. Q. Hou and C. W. Tu, Appl. Phys. Lett. **60**, 1872 (1992).

Chemistry of $[(t\text{-butyl})GaS]_4$ on Si(100)-(2×1)

R. I. Pelzel,[a] A. B. Hopcus, J. H. G. Owen, B. Z. Nosho, and W. H. Weinberg
Department of Chemical Engineering and Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, California 93106

(Received 21 January 1998; accepted 26 May 1998)

We have studied the chemistry of $[(t\text{-butyl})GaS]_4$ on Si(100)-(2×1) in ultrahigh vacuum. The characterization techniques used in this study were Auger electron spectroscopy (AES), temperature programmed desorption (TPD), and low-energy electron diffraction (LEED). Exposing $[(t\text{-butyl})GaS]_4$ to Si(100)-(2×1) at temperatures below 700 K gives evidence of gallium, sulfur, and carbon on the surface, as judged by AES, while exposing at an elevated temperature (>700 K) results in nearly carbon free adsorption. TPD spectra measured after exposure at 200 K show that the hydrocarbon ligands of the $[(t\text{-butyl})GaS]_4$ undergo a β-hydride elimination reaction and desorb predominantly as isobutene. For low exposures, the GaS core of the $[(t\text{-butyl})GaS]_4$ molecule dissociates upon annealing, as judged by the SiS and gallium TPD peaks. TPD experiments conducted after adsorption of more than 20 Å of $[(t\text{-butyl})GaS]_4$ on the Si(100)-(2×1) surface show evidence of Ga_2S desorption. For deposition at 700 K, LEED shows that the (2×1) reconstruction remains intact for low and moderate exposures, and a (1×1) pattern develops for higher exposures. © *1998 American Vacuum Society.* [S0734-211X(98)12704-X]

I. INTRODUCTION

The passivation of semiconductor interfaces is an issue of fundamental interest and has far-reaching technological implications. The successful passivation of III–V materials has been a particularly elusive goal. Unlike the native oxide of silicon, the oxide formed on GaAs (and other III–V semiconductors) is detrimental to device performance. The fact that the surface oxide of GaAs can degrade or even dominate the electrical performance of GaAs-based devices has severely hindered the development of GaAs metal–insulator–semiconductor (MIS) technology and optoelectronic technology.[1–3]

It has been known for some time that treating the GaAs surface with sulfur is successful in passivating the surface.[1–9] However, until recently, it was not possible to use sulfur to provide long-term passivation. Specifically, researchers were able to use simple sulfur-containing reagents (i.e., H_2S and elemental sulfur) to provide electrical and chemical passivation, but the passivation effects were only short lived.[10–12]

Recently, however, long-term passivation has been realized through the growth of cubic phase gallium sulfide on GaAs achieved by depositing a single-source molecular precursor, $[(t\text{-butyl})GaS]_4$.[13–19] A ball and stick model of $[(t\text{-butyl})GaS]_4$ is shown in Fig. 1. This molecule was used to deposit a face-centered-cubic (fcc) GaS layer on GaAs, which was instrumental in the fabrication of a GaAs MIS field effect transistor (FET) with low interface trap density and high resistivity.[19] The thick sulfide layer was shown to be resistant to atmospheric attack; the photoluminescence (PL) intensities do not decline over the course of a few months.[13,15–18] Although an *ex situ* characterization of the metalorganic chemical vapor deposition (MOCVD)-grown film clearly demonstrated the effectiveness of passivation, many questions regarding the surface chemistry and growth physics remain unanswered. In order to characterize fully the surface chemistry of $[(t\text{-butyl})GaS]_4$, *in situ* characterization is necessary. For the most part, surface sensitive techniques are absent during MOCVD growth. Therefore, an ultrahigh vacuum (UHV) investigation of growth needs to be conducted in order to characterize comprehensively the surface chemistry of this novel precursor. In fact, our group has recently begun to unravel the complex surface chemistry of $[(t\text{-butyl})GaS]_4$, on GaAs(100).[20,21] As an extension to this work, we describe here an initial investigation of the chemistry of $[(t\text{-butyl})GaS]_4$ on Si(100)-(2×1) as a means of comparison to the chemistry on GaAs(100).

II. EXPERIMENTAL METHODOLOGY

Experiments for this study were conducted in two separate stainless-steel UHV systems designated here as system 1 and system 2. System 1 houses an Auger spectrometer (Perkin–Elmer PHI 3017), four-grid, rear-view low-energy electron diffraction (LEED) optics, a differentially pumped quadrupole mass spectrometer (UTI-100C) equipped with a cryoshoud for temperature programmed desorption (TPD) studies, and a commercial (LK-2000-14-R) high-resolution electron energy loss spectrometer (HREELS). System 1 is also equipped with a liquid nitrogen cooled cryostat which allowed routine cooling of the sample to a temperature of 120 K. System 2 contains a scanning tunneling microscope (STM), a differential pumped quadrupole mass spectrometer (Stanford Research Systems RGA 200), and Auger and LEED components identical to those housed in system 1.

The samples used for this study were cut from *n*-type, nominally flat (±0.25°) and vicinal (1° miscut in the [001] direction) Si(100) wafers (Virginia Semiconductor Inc.). Samples of approximately 1 cm×1 cm were degreased in a mixture of H_2SO_4 and H_2O_2 (50:50 by volume) before being

[a]Corresponding author; electronic mail: pelzel@engineering.ucsb.edu

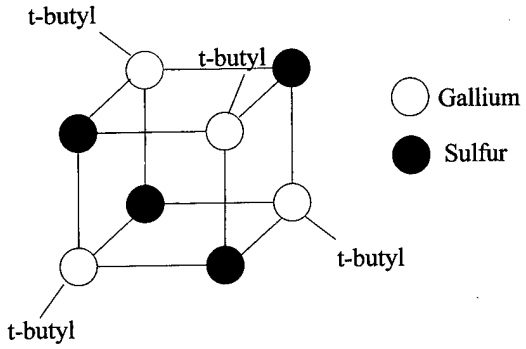


FIG. 1. Ball and stick representation of the $[(t\text{-butyl})GaS]_4$ precursor.

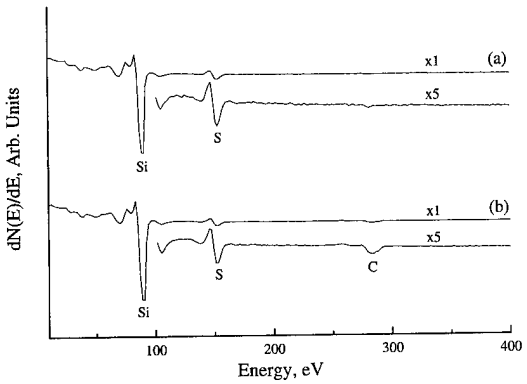


FIG. 2. AES spectra for a 30 s exposure of the Si(100)-(2×1) surface to $[(t\text{-butyl})GaS]_4$ at a surface temperature of (a) 700 and (b) 300 K. The spectra located below both (a) and (b) are magnifications to aid the reader.

mounted onto molybdenum sample holders, which had been cleaned using boiling HCl. Samples were introduced into both UHV adsorption/analysis chambers via loadlock systems. After degassing a sample for 24 h at 600–650 °C in UHV, the (2×1) reconstruction was prepared by repeatedly flashing the sample. Flashing refers to the procedure whereby the sample is annealed rapidly (less than 5 s) to 1200 °C, using electron bombardment, and then quenched to 650–675 °C (also less than 5 s). LEED was used to confirm the fact that the (2×1) reconstruction was well ordered, and Auger electron spectroscopy (AES) was used to verify that surfaces were clean prior to exposure to $[(t\text{-butyl})GaS]_4$. The AES spectra for the clean surfaces show no detectable carbon or oxygen on the surface as judged by the lack of Auger transitions at 271 and 503 eV.

The $[(t\text{-butyl})GaS]_4$ was synthesized at Rice University as discussed elsewhere.[22,23] The purity of the reagent as determined from nuclear magnetic resonance (NMR) was 99%. The $[(t\text{-butyl})GaS]_4$ was introduced into both UHV chambers using identical directional dosers.[21] The $[(t\text{-butyl})GaS]_4$ was housed in a hollowed stainless-steel cylinder (1.8 in. in length and 0.3 in. in diameter) with a removable faceplate (to facilitate replacement of the $[(t\text{-butyl})GaS]_4$ reagent). The faceplate contains five pinholes (each 0.2 mm in diameter) arranged in a square with one pinhole at the center of the square. The cylinder housing the $[(t\text{-butyl})GaS]_4$ was attached to the end of an linear translating arm (translation achieved via magnetic coupling). This arm was bolted to the loadlock of the analysis chamber, and the loadlock was pumped (using a turbomolecular pump) and lightly baked, achieving a pressure of 5×10^{-9} Torr with the $[(t\text{-butyl})GaS]_4$ source in the loadlock. To effect the adsorption experiments, the source was heated, and a gate valve between the main chamber and the loadlock opened so that the linear translator could be extended to provide directional exposure to the sample. Sample heating during adsorption (for both systems) was achieved using tungsten filaments behind the sample. Temperatures were estimated using thermocouples spot welded to the sample manipulators approximately 5 mm from the sample.

All exposures will be quantified as exposure times in front of the $[(t\text{-butyl})GaS]_4$ doser as well as by Si/S AES ratios. At this point, it is difficult to establish exact coverages since an accurate description of total coverage is dependent on whether the $[(t\text{-butyl})GaS]_4$ molecule is intact on the surface or whether decomposition has occurred. We are currently working on establishing a more quantitative representation of surface coverage.

III. RESULTS AND DISCUSSION

A. AES

From our group's work on GaAs, we discovered that it was possible to deposit $[(t\text{-butyl})GaS]_4$ carbon free by adsorbing it at elevated temperatures.[20,21] We conducted an AES study to determine if this is also possible for deposition on Si(100)-(2×1).

Figure 2 shows AES spectra for a Si(100)-(2×1) surface after a 30 s exposure at 300 K and a surface after an identical exposure time at 700 K. After a 30 s exposure at 300 K, AES indicates the presence of silicon, sulfur, and carbon which have transitions at 92, 152, and 271 eV, respectively. The Si/S ratio is 13. The gallium transition at 55 eV appears to be absent. However, this is not surprising since the intensity of this transition is expected to be weak in comparison to the silicon and sulfur transitions.[20,21,24] In subsequent experiments, we have confirmed the presence of gallium by taking scans at higher energies: the gallium transition is present at 1070 eV for all exposures studied. For an identical 30 s exposure at 700 K (cf. Fig. 2), the AES spectra again show the presence of silicon and sulfur (Si/S=10); there is only a negligible amount of carbon detected, however. Therefore, we are able to deposit $[(t\text{-butyl})GaS]_4$ on Si(100)-(2×1) nearly carbon free by adsorbing it at elevated temperatures.

B. TPD: Hydrocarbon desorption

Figure 3(a) shows the TPD spectra for mass 41 taken after a 10 s (Si/S=20) and a 60 s (Si/S=3) exposure of the Si(100)-(2×1) surface to $[(t\text{-butyl})GaS]_4$ at a surface temperature of 200 K. Mass 41 corresponds to the largest cracking fragment of isobutene which would result from a β-hydride elimination reaction of the *t-butyl* ligands from the $[(t\text{-butyl})GaS]_4$ molecule. We also monitored mass 43,

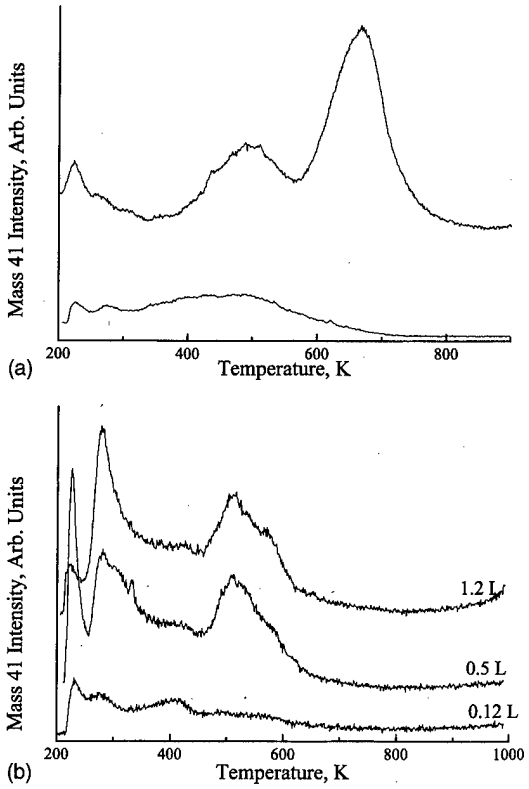


FIG. 3. (a) Isobutene desorption spectra for low and moderate exposures of $[(t\text{-butyl})GaS]_4$ at a surface temperature of 200 K. The lower curve corresponds to an exposure time of 10 s (Si/S=20), while the upper spectrum corresponds to an exposure time of 60 s (Si/S=3). (b) TPD spectra for isobutene desorption for the clean Si(100)-(2×1) surface for the various exposures indicated.

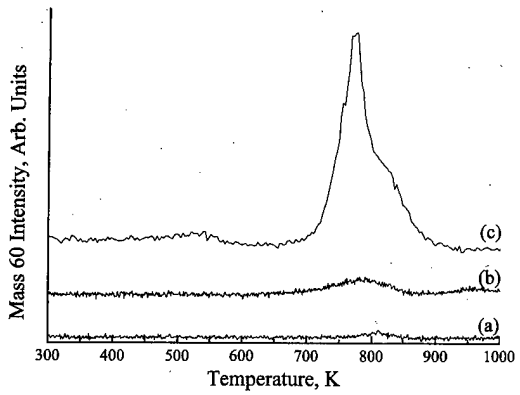


FIG. 4. TPD spectra for SiS desorption observed for (a) 10 s (Si/S=20), (b) 60 s (Si/S=3), and (c) 300 s (Si/S=0) exposures of the Si(100)-(2×1) surface to $[(t\text{-butyl})GaS]_4$ at a surface temperature of 200 K.

which is the major cracking fragment of isobutane. Although we did observe a small amount of isobutane, the intensity was markedly lower (by a factor of 5) than that observed for isobutene; isobutene is the major hydrocarbon desorption product. Figure 3(b) shows the mass 41 spectra for isobutene desorption from the initially clean Si(100)-(2×1) for various isobutene exposures at 200 K. The lowest temperature desorption peaks observable in both Figs. 3(a) and 3(b) (below 250 K) are attributable to desorption from the sample heating filament and are not the result of desorption from the surface.

At low exposures of $[(t\text{-butyl})GaS]_4$, isobutene desorbs over a broad temperature range with the distribution peaked at 280 and 500 K [cf. Fig. 3(a)]. In comparing the lower spectrum in Fig. 3(a) to the spectra for isobutene adsorption on the initially clean Si(100)-(2×1) surface [cf. Fig. 3(b)], we see some similarities. However, before this comparison is made, it is important to note that the spectra in Fig. 3(b) exhibit complexity beyond the scope of this article. Future work will address this complexity and its relationship to $[(t\text{-butyl})GaS]_4$ adsorption. The spectrum for low exposures to $[(t\text{-butyl})GaS]_4$ [cf. Fig. 3(a)] is most similar to the spectrum for a 0.12 L exposure of isobutene on the clean Si(100)-(2×1) surface [cf. Fig. 3(b)]. Comparing these two spectra we see that for both cases there is continuous desorption from 250 to 700 K. Furthermore, both spectra have a peak in the vicinity of 500 K. From this, we assert that the isobutene desorption seen for low exposures of $[(t\text{-butyl})GaS]_4$ may be attributable to desorption from the Si(100)-(2×1) surface. As mentioned above, the surface isobutene is generated by the β-hydride elimination reaction of the *t-butyl* ligands of the $[(t\text{-butyl})GaS]_4$ molecule.

For moderate exposures of $[(t\text{-butyl})GaS]_4$, the character of the isobutene TPD spectra is altered dramatically as compared to spectra observed for low exposures. The upper spectrum in Fig. 3(a) corresponds to an exposure of 60 s (Si/S=3), and the isobutene TPD is peaked at 500 and 670 K. The peak at 500 K is observed near the same temperature as the high temperature isobutene peak which is observed for the low exposure of $[(t\text{-butyl})GaS]_4$. In addition to the lower temperature peak, a higher temperature peak (670 K) is observed which is entirely absent from the isobutene desorption spectra from the clean surface [cf. Fig. 3(b)].

Although these isobutene TPD spectra cannot be used as conclusive evidence for a reaction mechanism, we can make some useful assertions and hypotheses based on the data of Fig. 3. For moderate $[(t\text{-butyl})GaS]_4$ exposures, the lower temperature peak may be attributable to isobutene desorption from the silicon surface resulting from a β-hydride elimination reaction. This supposition is based simply on the fact that the peak temperature observed for the lower temperature peak (500 K) at moderate coverages is in the vicinity of a peak which is present for isobutene desorption from the clean surface [cf. Fig. 3(b)]. The higher temperature isobutene TPD peak observed for moderate exposures is most likely due to a β-hydride elimination reaction, possibly one where the isobutene "desorbs" from gallium in the $[(t\text{-butyl})GaS]_4$ molecule rather than from the silicon surface. Such a reaction might occur in a multilayer of adsorbed $[(t\text{-butyl})GaS]_4$.

C. TPD: Gallium and sulfur desorption

Figures 4 and 5 show TPD spectra for masses 60 and 69, respectively, for a low (10 s, Si/S=20), moderate (60 s, Si/S=3), and high exposure (300 s, Si/S=0) of $[(t\text{-butyl})GaS]_4$ at a surface temperature of 200 K. Mass 60

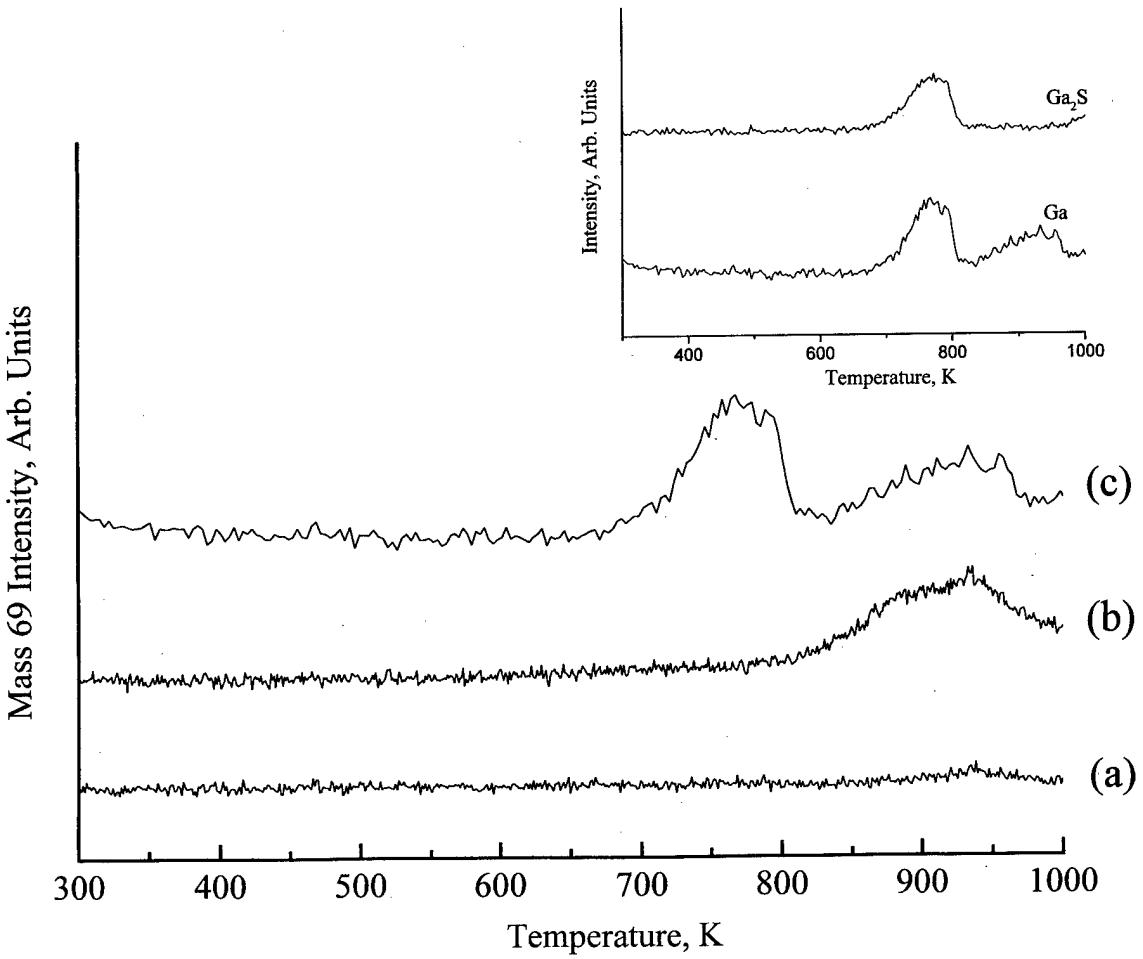


FIG. 5. TPD spectra for gallium desorption observed for (a) 10 s (Si/S=20), (b) 60 s (Si/S=3), and (c) 300 s (Si/S=0) exposures of the Si(100)-(2×1) surface to $[(t\text{-butyl})GaS]_4$ at a surface temperature of 200 K. The inset shows the gallium and Ga_2S TPD spectra for the 300 s exposure.

corresponds to SiS which is the major desorption product observed for elemental sulfur adsorption on Si(100),[25] and mass 69 corresponds to gallium which has been observed to be the primary desorption product for gallium adsorption on silicon.[26,27] We also observed desorption of Ga_2S (mass 172).

From Fig. 4, it is clear that we see evidence for SiS desorption. For low (10 s, Si/S=20) and moderate (60 s, Si/S=3) exposures, the SiS TPD exhibits a single peak centered between 790 and 810 K. The peak temperature shifts downward with increasing exposure, possibly due to surface adsorbate interactions. For the high exposure (300 s, Si/S=0), there appears to be a convolution of two peaks. The high temperature shoulder evident near 800 K resembles "monolayer" desorption while the larger peak at 770 K resembles what would be expected from multilayer desorption. The desorption temperature window we observe for SiS is approximately 70 K downshifted compared to SiS desorption observed for the deposition of elemental sulfur on Si(100).[25] This may well be due to adsorbate interactions present for $[(t\text{-butyl})GaS]_4$ deposition on Si(100) that are absent for elemental sulfur deposition. Furthermore, the single TPD study of elemental sulfur on Si(100)-(2×1) does not report multilayer SiS desorption.[25] However, this study focused only on adsorption up to 2 monolayers (ML) and our sulfur coverage is most likely in excess of this.

Figure 5 shows gallium TPD spectra for the low, moderate, and high exposures mentioned above. For low coverages, the spectrum contains a single gallium desorption peak centered at 940 K. For moderate coverages, the spectrum broadens with a peak temperature again of 940 K. The broadening of the spectrum for moderate coverages is indicative of the convolution of monolayer and multilayer desorption peaks observed for gallium adsorption on Si(100)-(2×1).[26,27] The temperature observed for gallium desorption is in good agreement with what has been reported in the literature.[26,27] For the highest coverage, the spectrum is bimodal, peaked at 780 and 940 K. Again, the character of the higher temperature desorption peak is indicative of monolayer and multilayer desorption. The additional gallium peak is a cracking fragment of Ga_2S observed for the high coverage regime (Si/S=0). The inset in Fig. 5 shows both gallium and Ga_2S desorption for the high exposure. As can be seen, the peak shape and maximum temperature coincide exactly. No Ga_2S desorption is observed for low and moderate $[(t\text{-butyl})GaS]_4$ exposures.

Based on these TPD data, we can assert that at low and moderate exposures of $[(t\text{-butyl})GaS]_4$ the precursor dissociates as judged by the desorption of isobutene, SiS, and gallium from the surface and the lack of Ga_2S. We were able to generate Ga_2S by condensing a thick layer of $[(t\text{-butyl})GaS]_4$ on the Si(100)-(2×1) surface and annealing it. Since the silicon AES peak is no longer observable for this thick multilayer, the thickness of this layer is estimated to be greater than 20 Å, based on the escape depth of an Auger electron.[28] At this point, it is unclear whether the gallium sulfide generated for the high exposure is cubic since Ga_2S desorption would be expected to observed if the gallium sulfide layer had a stoichiometry of Ga_2S or Ga_2S_3 neither of which is cubic gallium sulfide. STM studies are currently underway to address this point.

D. LEED

We conducted LEED studies for low to high exposures of $[(t\text{-butyl})GaS]_4$ at 700 K (negligible carbon adsorption). Our interest was motivated by extensive evidence that gallium induces a complex set of reconstructions on Si(100)-(2×1) for gallium coverages of less than 1 ML.[29–32] For all low exposures investigated, we observed no evidence that the (2×1) order was disrupted. The background intensity increased with exposure, but the LEED pattern remained (2×1). Eventually, the (2×1) reconstruction is lost as the exposure is increased, and a faint (1×1) pattern is observed. The fact that the reconstruction does not change for low exposures of $[(t\text{-butyl})GaS]_4$ suggests that gallium and sulfur rich domains do not form for submonolayer coverages. STM is currently being employed to confirm this.

IV. SUMMARY

We have used AES to demonstrate that we can deposit $[(t\text{-butyl})GaS]_4$ relatively carbon free at elevated temperatures (>700 K). The isobutene TPD spectra measured after exposures at 200 K show evidence for β-hydride elimination of the *t-butyl* ligands for the $[(t\text{-butyl})GaS]_4$ molecule. The TPD also shows the desorption of gallium and SiS for all exposures investigated giving strong evidence that the GaS core of the $[(t\text{-butyl})GaS]_4$ molecule completely dissociates for low and moderate exposures. At higher exposures (300 s, Si/S=0), we observe the desorption of Ga_2S, demonstrating that it may be possible to grow GaS by depositing thick layers (>20 Å) above a temperature of 400 K. The LEED studies show that the (2×1) reconstruction is intact for low and moderate exposures at elevated temperature (700 K). Eventually, a (1×1) LEED pattern is observed.

ACKNOWLEDGMENTS

The authors wish to thank C. J. Hagedorn and M. J. Weiss of the University of California, Santa Barbara and Dr. Andrew MacInnes of TriQuint Inc. for useful discussions. They also thank Dr. Andrew Barron and his group at Rice University for supplying them with the $[(t\text{-butyl})GaS]_4$ precursor. The financial support by QUEST, a NSF Science and Technology Center for Quantized Electronic Structures (Grant No. DMR 91-20007), by the National Science Foundation (Grant No. DMR-9504400), and by the W. M. Keck Foundation is gratefully acknowledged.

[1]C. R. Zeisse, L. J. Messick, and D. L. Lile, J. Vac. Sci. Technol. **14**, 957 (1977).
[2]L. G. Meiners, J. Vac. Sci. Technol. **15**, 1402 (1978).
[3]T. Izawa, T. Ishibashi, and T. Sugeta, Tech. Dig. Int. Electron Devices Meet. , 328 (1985).
[4]E. Yablonovitch, C. J. Sandroff, R. Bhat, and T. Gmitter, Appl. Phys. Lett. **51**, 439 (1987).
[5]M. S. Carpenter, M. R. Melloch, and T. E. Dungan, Appl. Phys. Lett. **53**, 66 (1988).
[6]Z. H. Lu, D. Landheer, J.-M. Baribeau, L. J. Huang, and W. W. Lau, Appl. Phys. Lett. **64**, 1702 (1994).
[7]C. J. Sandroff, R. N. Nottenburg, J.-C. Bischoff, and R. Bhat, Appl. Phys. Lett. **51**, 33 (1987).
[8]R. N. Nottenburg, C. J. Sandroff, D. A. Humphrey, T. H. Hollenbeck, and R. Bhat, Appl. Phys. Lett. **52**, 18 (1988).
[9]C. J. Sandroff, M. S. Hedge, and C. C. Chang, J. Vac. Sci. Technol. B **7**, 841 (1989).
[10]T. Tiedje, K. M. Colbow, D. Rogers, Z. Fu, and W. Eberhardt, J. Vac. Sci. Technol. B **7**, 837 (1989).
[11]H. Kawanishi, Y. Sugimoto, and K. Akita, J. Vac. Sci. Technol. B **9**, 1539 (1991).
[12]H. Kawanishi, Y. Sugimoto, and K. Akita, J. Appl. Phys. **70**, 805 (1991).
[13]A. R. Barron, Adv. Mater. Opt. Electron. **5**, 245 (1995).
[14]P. P. Jenkins, A. F. Hepp, M. B. Power, A. N. MacInnes, and A. R. Barron, Mater. Res. Soc. Symp. Proc. **282**, 111 (1993).
[15]A. N. MacInnes, M. B. Power, A. R. Barron, P. P. Jenkins, and A. F. Hepp, Appl. Phys. Lett. **62**, 711 (1993).
[16]A. N. MacInnes, M. B. Power, and A. R. Barron, Chem. Mater. **5**, 1344 (1993).
[17]A. N. MacInnes, M. B. Power, and A. R. Barron, Mater. Res. Soc. Symp. Proc. **282**, 659 (1993).
[18]M. Tabib-Azar, S. Kang, A. N. MacInnes, M. B. Power, A. R. Barron, P. P. Jenkins, and A. F. Hepp, Appl. Phys. Lett. **63**, 625 (1993).
[19]P. P. Jenkins, A. N. MacInnes, M. Tabib-Azar, and A. R. Barron, Science **263**, 1751 (1994).
[20]S. I. Yi, C.-H. Chung, J. H. G. Owen, A. Hopcus, and W. H. Weinberg, Surf. Sci. (submitted).
[21]S. I. Yi, C.-H. Chung, and W. H. Weinberg, J. Vac. Sci. Technol. A **16**, 1650 (1998).
[22]A. N. MacInnes, W. M. Cleaver, A. R. Barron, M. B. Power, and A. F. Hepp, Adv. Mater. Opt. Electron. **1**, 229 (1992).
[23]A. N. MacInnes, M. B. Power, and A. R. Barron, Chem. Mater. **4**, 11 (1992).
[24]H. Yamada, K. Ueno, and A. Koma, Jpn. J. Appl. Phys., Part 2 **35**, L568 (1996).
[25]A. Papageorgopoulos, A. Corner, K. Kamaratos, and C. A. Papageorgopoulos, Phys. Rev. B **55**, 4435 (1997).
[26]K. L. Carleton and S. R. Leone, J. Vac. Sci. Technol. B **5**, 1141 (1987).
[27]K. L. Carleton, B. Bourguignon, and S. R. Leone, Surf. Sci. **199**, 447 (1988).
[28]I. Lindau and W. E. Spicer, J. Electron Spectrosc. Relat. Phenom. **3**, 409 (1974).
[29]P. Sakamoto and H. Kawanami, Surf. Sci. **111**, 177 (1981).
[30]A. A. Basaki, J. Nogami, and C. F. Quate, J. Vac. Sci. Technol. A **8**, 245 (1990).
[31]B. Bourguignon, R. V. Smilgys, and S. R. Leone, Surf. Sci. **204**, 473 (1988).
[32]H. Sakama, A. Kawazu, T. Sueyoshi, T. Sato, and M. Iwatsuki, Phys. Rev. B **54**, 8756 (1996).

New model for reflection high-energy electron diffraction intensity oscillations

W. Braun[a)]
Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

L. Däweritz and K. H. Ploog
Paul-Drude-Institut für Festkörperelektronik, D-10117 Berlin, Germany

(Received 21 January 1998; accepted 28 May 1998)

We investigate the influence of inelastic processes on reflection high-energy electron diffraction (RHEED) oscillations by recording energy filtered RHEED intensity oscillations during homoepitaxy of (001)-oriented GaAs. The results clearly show that the dominant inelastic scattering process, plasmon inelastic scattering, does not influence the phase of the oscillations. It cannot therefore account for an independent process contributing a phase to the oscillations that is different from elastic scattering. As an alternative approach, we investigate a basic coherent scattering model introduced by Horio and Ichimiya. We compare its predictions with experiments in the one-beam condition for both GaAs and AlAs(001) homoepitaxy. The average crystal potential required for the fits can be determined independently by Kikuchi line fits, yielding a value of 10.5 ± 0.5 V for both GaAs and AlAs. This allows us to reduce the number of free parameters in the model to only the layer thickness. The theoretical fits of the model to the experimental data yield different layer thicknesses that are in good agreement with the surface reconstruction thicknesses for GaAs and AlAs. We therefore conclude that the phase of RHEED oscillations is determined by the surface reconstruction forming on top of the growing layer during crystal growth. This new model explains many experimentally observed RHEED oscillation phenomena in a unified approach. © *1998 American Vacuum Society.* [S0734-211X(98)12304-1]

I. INTRODUCTION

Reflection high-energy electron diffraction (RHEED) is widely used as an *in situ* characterization tool in molecular beam epitaxy (MBE) due to its compatibility with the growth process. Its main application is the determination of growth rates and the study of growth dynamics through RHEED intensity oscillations.[1–3] The intensity of the RHEED pattern oscillates with a period corresponding to the bulk plane spacing in the growth direction, allowing an *in situ* determination of the growth rate. Although RHEED intensity oscillations have been investigated for quite some time, many aspects are still not very well understood. Whereas the period of the oscillations can be explained in terms of a periodically changing surface configuration in the layer-by-layer growth mode, the diffraction mechanism actually leading to a modulated electron flux at the detector is still a matter of debate.

The phase of the oscillations varies strongly as a function of incidence angle[4] and as a function of As flux in GaAs homoepitaxy.[5] This has been explained by the superposition of two competing processes.[6–10] The first process is assumed to be the coherent Bragg diffraction between the top of the growing layer and the substrate surface it is growing on,[11] one that involves atoms at positions that define the long-range periodicity of the surface. The second process then is defined as incoherent scattering proportional to the density of step edge atoms and other defects varying in number with the evolving surface morphology.[12] At the so-called out-of-phase conditions between the Bragg peaks, the diffracted intensity would be dominated by the destructive interference between the two levels in the coherent process, leading to minima at half layer coverage. At the in-phase conditions or Bragg conditions, the coherent process would not produce any intensity variations because contributions from both levels are in phase. Instead, the incoherent process would dominate there, with a maximum at half layer coverage where the step density is largest. The phase variation could then be explained by the continuous variations between both extremes, alternating between a dominance of coherent and incoherent contributions.

Theoretical calculations still disagree as to whether the step density should contribute proportionally or antiproportionally to the observed RHEED intensity.[8,13] In this work, we investigate the possibility of inelastic scattering contributing to the second process during RHEED oscillations. Inelastic scattering destroys the coherence and is therefore a prime candidate as a second independent process contributing a different phase to the total oscillating intensity. A variation in the different energy loss processes should be detectable in the energy spectrum of the diffracted electrons during crystal growth.

II. ENERGY LOSS SPECTROSCOPY RHEED

To investigate the different energy loss processes that contribute to RHEED oscillations, we developed an energy loss spectroscopy (ELS-) RHEED detector that replaces the standard RHEED screen on the MBE chamber. A picture of

[a)]Electronic mail: braun@enpop2.eas.asu.edu, http://asumbe.eas.asu.edu/wolfgang/welcome.htm

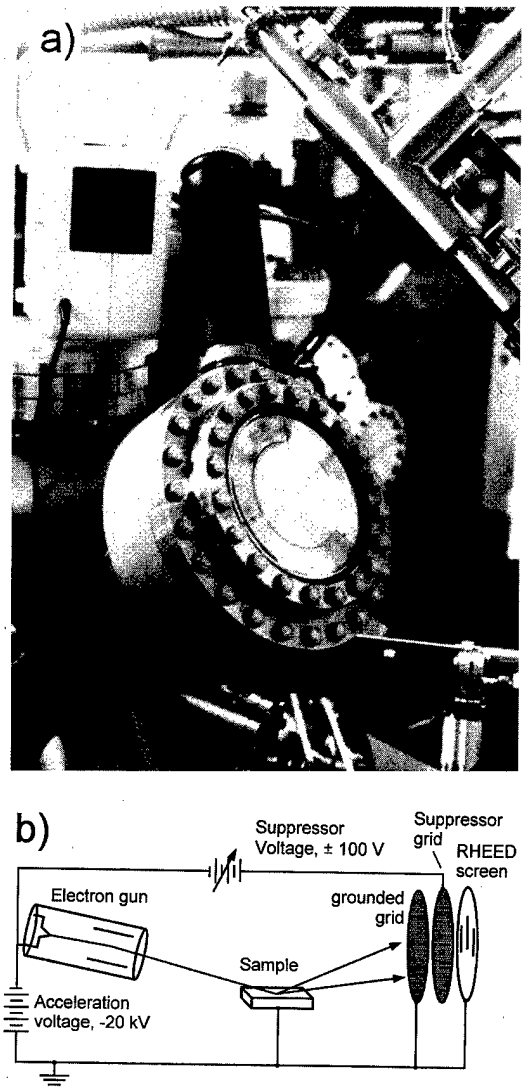


FIG. 1. ELS-RHEED detector. (a) The detector mounted on a standard MBE system. (b) Schematic of the ELS-RHEED setup. The RHEED screen and the grounded grid are fixed at zero potential, and the suppressor grid is kept at a variable potential close to the corresponding primary beam energy. All electrons that have experienced an energy loss larger than the potential difference are rejected by the suppressor grid. Differentiation of the intensity with respect to the suppressor grid potential yields the energy spectrum of the diffracted electrons.

the detector along with a schematic of its operating principle are shown in Fig. 1. The design is similar to the detector developed by Horio, Hashimoto, and Ichimiya.[14,15] Instead of three rather coarse spherical grids, we used two planar fine mesh Au grids with a period of 17 μm, resulting in better spatial and energy resolution. The energy filter rejects electrons with lower energy than the potential of the central grid. By varying the suppressor grid potential, only electrons with an energy loss less than the cutoff are detected on the phosphorous screen, allowing the acquisition of spectra. The main advantages of this detector are its rugged design that makes it compatible with practically any MBE system and the absence of channeltron or channelplate amplification that is usually incompatible with an As ambient. Details of the

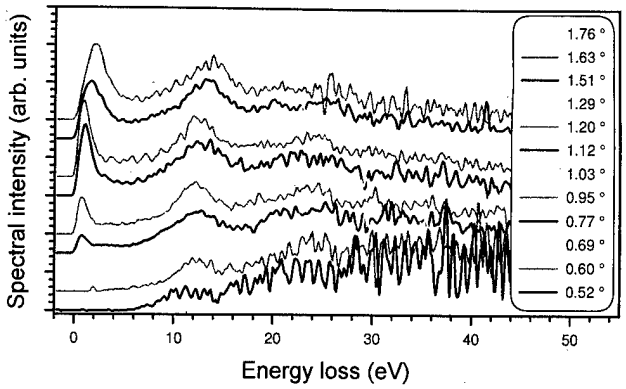


FIG. 2. Energy loss spectra recorded at different incidence angles on the specular spot. GaAs(001) $\beta(2\times4)$ reconstructed surface. The primary energy is 20 keV.

setup are described elsewhere.[16] Analysis of the intensity distribution on the screen was performed with a PC-based digital RHEED analysis system[17] equipped with a high sensitivity charge coupled device (CCD) camera and capable of acquiring data at a rate of 50 Hz.

A set of spectra for different incidence angles are shown in Fig. 2. In each case, the spectrum was measured on the specularly reflected (specular) spot. With increasing incidence angle, the amount of elastic scattering increases. Since the energy resolution of the ELS-RHEED detector is typically just below 2 eV, the instrument is unable to resolve inelastic processes involving less than 2 eV energy loss, which are some band-to-band transitions and phonon inelastic scattering. We therefore define "elastic" within the context of this article to also comprise phonon inelastic scattering. For the typical high-quality MBE-grown surfaces, the energy loss spectrum is dominated by multiple surface plasmon losses at $n\times11$ eV similar to surfaces of liquid metals.[18] At very low angles, all electrons undergo plasmon scattering and the elastic peak is not detectable. Around 1°, typically 10%–20% of the total intensity is elastic; the remainder is almost entirely plasmon inelastic. Loss processes involving larger energy transfer like core shell excitations are typically too small to contribute a significant amount to the total intensity. The most likely candidate for an inelastic second process that contributes a significant amount of the total intensity is therefore plasmon inelastic scattering.

We recorded RHEED intensity oscillations at different incidence angles and with different energy filter settings for GaAs(001) homoepitaxy. The results are shown in Fig. 3. The two examples were chosen to be close to the incidence angle with the least band-to-band losses, resulting in a narrow elastic peak. This allows a clear separation of the elastic contribution from the rest of the spectrum. Also, the curves of Figs. 3(a) and 3(b) are mutually out of phase with each other to rule out effects that could be unique to a specific phase position. The filter cutoff was set at $\leqslant$2, $\leqslant$15, and $\leqslant$30 eV energy loss to include only the elastic peak or the elastic peak plus the single and double surface plasmon losses. In both Figs. 3(a) and 3(b), the lower three curves represent the raw data with their actual intensity ratio,

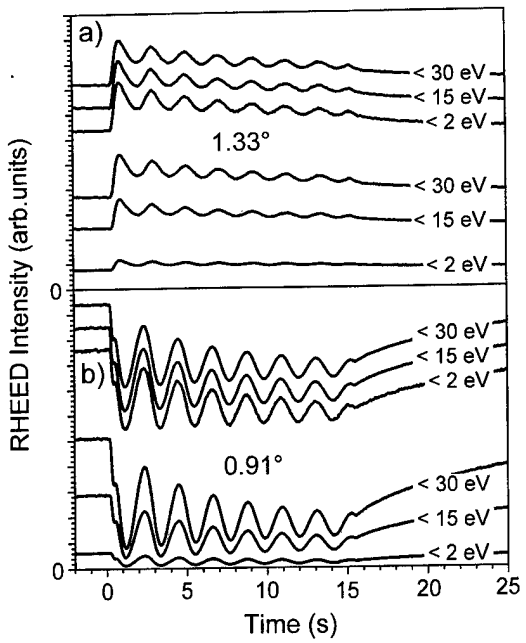


FIG. 3. RHEED oscillations recorded on the specular spot for GaAs(001) growth in the (2×4) surface reconstruction. The different curves represent different filter potential settings. The primary beam energy is 20 kV. Shown are the (a) original and (b) normalized curves for two cases that are out of phase with each other. In both cases, elastic and plasmon inelastic contributions to the spectrum maintain their relative intensities.

whereas the top three curves are normalized to their pregrowth intensity. In both Figs. 3(a) and 3(b), the curves are identical to within their linewidth. The same holds for the entire angular range of Fig. 2 that was studied. The top three curves in each panel of Fig. 3 are therefore shifted for clarity.

The identity of the normalized curves within measurement accuracy means that electrons having undergone plasmon inelastic scattering do not contribute a different phase to the RHEED oscillations. Instead, the process causing the oscillations is independent of the plasmon loss mechanism since both elastic and inelastic contributions produce the same curve shape. The diffraction process causing the oscillations must therefore be coherent in the sense that, as long as their direction is not changed significantly by the inelastic process, both the elastic and inelastic parts of the electron beam undergo the same process modulating its intensity. All electrons that are detected in the chosen area around the specular spot undergo the same process as far as the oscillations are concerned. We can therefore extend our results to phonon inelastic scattering. Plasmon inelastic scattering depends strongly on incidence angle, which is obvious from Fig. 2. Phonon inelastic scattering involves significant momentum transfer.[19] To overcompensate elastic scattering, the variation of phonon inelastic scattering with surface morphology would have to be large. If the amount of phonon inelastic scattering should depend on the surface imperfection density, it should therefore affect the distribution of elastic and plasmon inelastic scattering via the change in momentum. This effect should be detectable in our measurement. This is not the case. We therefore look for a different

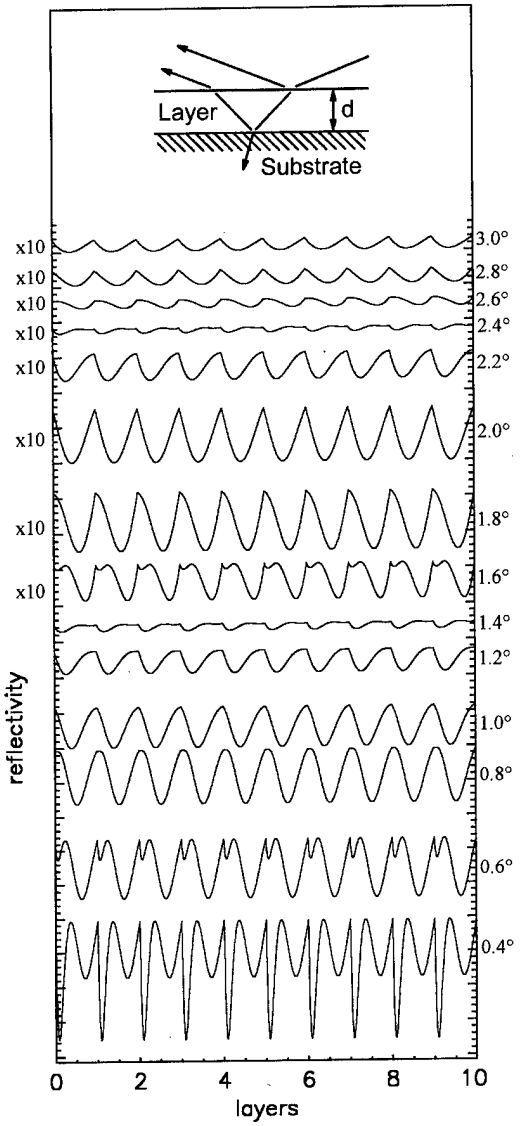


FIG. 4. Simulated RHEED oscillations for perfect layer-by-layer growth using the top layer interference model. A schematic of the model is shown in the inset. The layer thickness d is 0.282 nm with a potential V of 10 V. The absorption is neglected.

model to replace the two-process model. The RHEED oscillation mechanism we are looking for must depend sensitively on the electron's direction, but not very sensitively on its energy. It should also explain the phase dispersion as a function of incidence angle with a coherent scattering approach as a single process.

III. LAYER INTERFERENCE MODEL

We turn to a model proposed by Lehmpfuhl, Horio, and Ichimiya[20,21] that to our knowledge was only compared with Ge(111) growth experiments so far.[22] Simulated oscillations using this model are shown in Fig. 4. Growth is simulated by linearly increasing the potential of a layer on bulk material from zero to the bulk potential value,

$$V_{\text{layer}} = \theta V_{\text{substrate}}, \tag{1}$$

with θ being the layer coverage. When the bulk value is reached, the bottom interface vanishes and the process can be repeated periodically. The reflectance at the specular position is calculated from the interference of the beam reflected off the top growing layer surface with the beam refracted at the top and reflected at the bottom interface of the growing layer. Since the surface parallel component of the wave vectors remains unchanged throughout the process (elastic scattering), the geometry can be reduced to a one-dimensional problem, the familiar textbook example of a quantum mechanical particle incident on a twofold downward potential step. The nonvanishing probability of the quantum mechanical particle to get reflected back to the vacuum is largest at small incidence angles. This can be verified by the increasing small angle reflectance and oscillation amplitude in Fig. 4. The model completely neglects any lateral potential modulations and therefore is only valid in the so-called one-beam condition, where the azimuthal direction of the electron beam is along a low symmetry direction. This eliminates diffraction contributions due to the lateral potential variations. The one-beam condition in the experiment corresponds to an azimuth showing a minimum amount of Kikuchi lines, as will become clear from the discussion in Sec. IV.

For our present study, we simplify the model and then adapt the experimental conditions to match the approximations of the theory. Instead of trying to match an elaborate theoretical model to a chosen experiment, this has the advantage of keeping the number of free parameters in the model small, allowing a more stringent test of the basic mechanism the model describes. Growth is therefore simulated by one layer only to avoid the pitfalls of modeling growth on multiple levels.[23] Also, the substrate and final layer potentials are set identical and absorption is neglected. We have shown[23] that the inclusion of absorption through an imaginary part of the potential does not significantly affect the results. This layer interference model then has only two free parameters, namely the potential value V and the layer thickness d. For the simulations of Fig. 4, $V=10$ V and $d=0.28$ nm [the bulk (001) GaAs plane spacing] were used.

Analogous to Ref. 21, the reflectance is calculated as the square of the ratio of the reflected wave amplitude B to the incident wave amplitude A:

$$R=\left(\frac{B}{A}\right)^2 =\left(\frac{K^- a_{\mathrm{lay}} \exp(-iK^+ d)+K^+ b_{\mathrm{lay}} \exp(-iK^- d)}{K^+ a_{\mathrm{lay}} \exp(iK^- d)+K^- b_{\mathrm{lay}} \exp(iK^+ d)}\right), \quad (2)$$

where $K^+=K_\perp+k_{\perp\mathrm{lay}}$ and $K^-=K_\perp-k_{\perp\mathrm{lay}}$. a_{lay} and b_{lay} are the amplitudes of the downward and upward waves in the layer, given by

$$a_{\mathrm{lay}}=\frac{k_{\perp\mathrm{lay}}+k_{\perp\mathrm{sub}}}{2k_{\perp\mathrm{lay}}}\, a_{\mathrm{sub}} \quad (3)$$

and

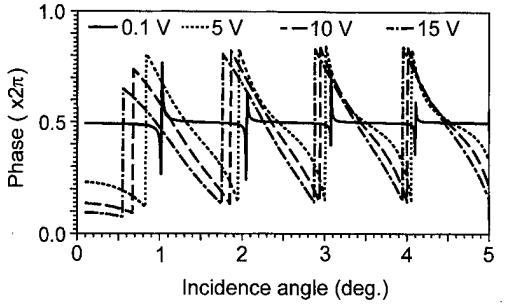


FIG. 5. Phase dispersion plots (position of the oscillation minimum vs incidence angle) for different potential values V of the layer interference model.

$$b_{\mathrm{lay}}=\frac{k_{\perp\mathrm{lay}}-k_{\perp\mathrm{sub}}}{2k_{\perp\mathrm{lay}}}\, a_{\mathrm{sub}}. \quad (4)$$

The surface normal component of the wave vector in the vacuum is denoted by $K_\perp$. The surface normal components of the wave vectors in the layer, $k_{\perp\mathrm{lay}}$, and the substrate, $k_{\perp\mathrm{sub}}$, are given by

$$k_{\perp\mathrm{lay}}=\sqrt{K_\perp^2+U_{\mathrm{lay}}} \quad (5)$$

and

$$k_{\perp\mathrm{sub}}=\sqrt{K_\perp^2+U_{\mathrm{sub}}}, \quad (6)$$

respectively, with

$$U_{\mathrm{lay}}=\frac{2me\theta V}{\hbar^2} \quad (7)$$

and

$$U_{\mathrm{sub}}=\frac{2meV}{\hbar^2}, \quad (8)$$

where m is the electron mass, e the electronic charge, θ the coverage of the layer, and V the average crystal potential.

The layer interference model has several attractive features. It reproduces the reduction in total intensity on the specular spot with increasing incidence angle. Also, it exhibits a periodical variation in the oscillation amplitude with increasing angle, which is also observed in the experiment.[11] In addition, the model features variations in the phase. The positions of the oscillation minima vary strongly. With increasing incidence angle, the minimum moves to the left, then a second minimum drops down close to integer coverages and subsequently moves left to repeat the process. The model always produces a maximum reflectance at integer coverages. This is a consequence of using only one layer. When the growth front is distributed among several layers, this peak vanishes and the derivative becomes continuous.[21] The additional sharp maximum is not artificial. It can sometimes be observed on very smooth surfaces during the first RHEED oscillation.[24]

When we plot the position of the oscillation minimum against the incidence angle, we obtain the curve shown in Fig. 5. With increasing potential value, the phase dispersion becomes increasingly sawtoothlike with the phase jumps moving to lower angles. These phase jumps correspond to

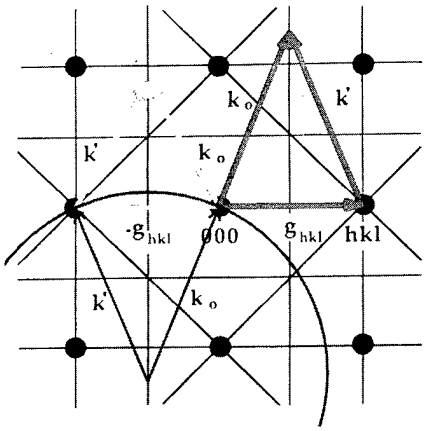


FIG. 6. Graphical representation of the diffraction condition (light gray vectors) and the Ewald sphere construction corresponding to the medium gray case in dark gray. The vectors originating at reciprocal lattice points end on Brillouin zone boundaries.

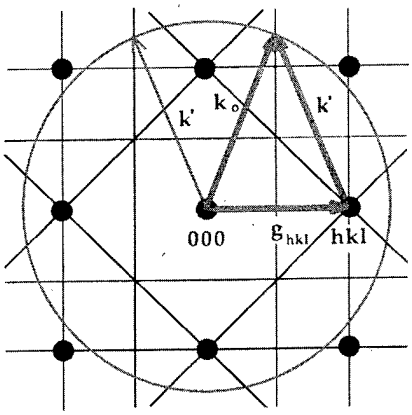


FIG. 7. Graphical representation of the sphere of reflections. The cut of this sphere with the Brillouin zone boundaries determines the Kikuchi line positions.

the crossover of the two oscillation minima and are located at local minima of the oscillation amplitude. The double minima at low angles have been observed experimentally.[25,26] For small potentials, the behavior continuously approaches that of the kinematical Bragg diffraction case with the phase jump positions coinciding with the (kinematical) in-phase Bragg conditions, where an x-ray measurement would produce intensity maxima. This is due to the fact that the layer interference model is a straightforward dynamical scattering extension of the Bragg diffraction model by including just the zeroth order Fourier component of the crystal potential. With increasing potential, the Bragg conditions move to lower angles due to refraction at the layer–vacuum interface. We therefore call them generalized Bragg conditions, since they take into account the refraction shifts.

The layer interference model still has two free parameters. One of them, the average crystal potential, can be determined independently by fitting the Kikuchi line pattern, reducing the parameters in the model to only one.

IV. KIKUCHI LINE FITS

Kikuchi lines[27] can be explained by a two-step scattering model.[28,29] In the first step, the incident electrons suffer collisions that randomize the direction of their wave vectors.

If the energy loss in this first collision is small compared to the electron energy, the resulting electron distribution corresponds to an electron source of (almost) the initial energy and isotropic emission within the crystal. To obtain a clear line pattern, the energy spread of this electron source must be small. The Kikuchi line pattern then is the diffraction pattern due to this isotropic electron irradiation. It can be determined using basic kinematical theory. The diffraction condition $\mathbf{k}' - \mathbf{k}_0 = \mathbf{g}_{hkl}$ is shown in Fig. 6. Here, $\mathbf{k}_0$ is the wave vector of the incident electron, $\mathbf{k}'$ the wave vector of the diffracted electron, and $\mathbf{g}_{hkl}$ is a reciprocal lattice vector. Diffraction occurs for any $\mathbf{k}_0$ that starts at the origin of the reciprocal lattice and ends on the plane bisecting the corresponding reciprocal lattice vector $\mathbf{g}_{hkl}$. These planes define the Brillouin zone boundaries of the crystal.

Usually, however, the diffraction geometry is treated using the Ewald sphere construction shown in dark gray in Fig. 6. In this case, the tip of $\mathbf{k}_0$ is attached to a reciprocal lattice point and its origin defines the center of the Ewald sphere. The Ewald sphere construction is not very suitable for the present case of wide-angle illumination, since the construction of the $\mathbf{k}'$ vectors consistent with the diffraction condition involves a continuum of Ewald spheres. Instead, an approach based on the basic geometry of the gray vectors in Fig. 7 yields a more direct result. Instead of the Ewald sphere, we define a sphere of radius k_0, centered at the origin of the reciprocal lattice. This sphere is called the "sphere of reflections".[29] As shown in Fig. 7, the diffraction condition is fulfilled for the intersection of the sphere of reflections with any Brillouin zone boundary.

Since the corresponding diffracted vectors $\mathbf{k}'$ for this condition also end at the zone boundary and start at a reciprocal lattice point, they can be translated to the origin by a reciprocal lattice vector, while still ending on the sphere. Like this, a vector $\mathbf{k}'$ exists for any vector $\mathbf{k}_0$ with the same properties, namely, that it ends at the intersection of the sphere of reflections with a Brillouin zone boundary. Due to the translational symmetry of the reciprocal lattice, the $\mathbf{k}'$ vectors scan the complete set of zone boundaries intersecting the sphere when $\mathbf{k}_0$ assumes all possible directions.

This construction provides us with a general geometrical procedure to produce the Kikuchi line pattern that holds for any type and geometry of diffraction. The first step consists of determining the intersection pattern of the sphere of reflections with the Brillouin zone boundaries of the crystal. This pattern determines the possible endpoints of the diffracted vectors. The visible Kikuchi line pattern then is the projection of these intersections to the RHEED screen with the origin of the reciprocal lattice as the center of the projection.

In the RHEED case, the procedure can be simplified by taking into account the small solid angle of the screen and

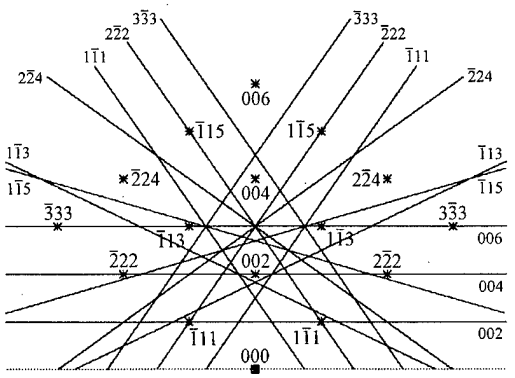


FIG. 8. Theoretical Kikuchi line pattern for the crystal surface geometry of a fcc (001) surface along the [110] direction. No surface potential corrections are included.

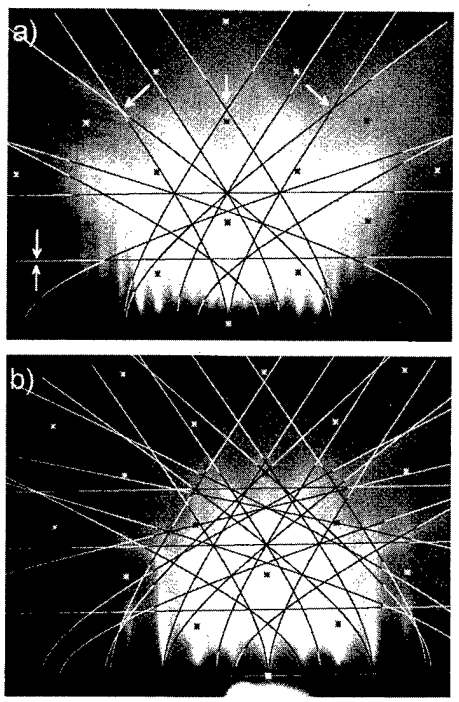


FIG. 9. Kikuchi line fits to the diffraction patterns of (a) GaAs $\beta(2\times4)$ and (b) AlAs $c(4\times4)$ reconstructed surfaces. The arrows indicate the most reliable points for fitting.

the large radius of the sphere of reflections. For the area visible on the screen, the cut of the sphere through the reciprocal lattice is almost planar and parallel to the screen. Let us also assume some broadening of the range of k_0's involved in the scattering. This causes all zone boundaries not parallel to the zone axis towards the screen to be blurred enough so that their intensities are indistinguishable from the background. No lines corresponding to these nonparallel zone boundaries are observed in the experiment.

Although the generation of the Kikuchi lines is a three-dimensional process, the construction of their projection on the screen has now been reduced to a two-dimensional problem. We plot the reciprocal lattice points in the plane parallel to the screen and construct the Brillouin zone boundaries. The geometry of the reciprocal lattice plane can often be directly verified in the RHEED image of a rough surface that displays a transmission pattern. The resulting line pattern is shown in Fig. 8 for a (001) face-centered-cubic (fcc) crystal surface looking along the [110] direction.

The Kikuchi lines are labeled according to the reflections they are derived from. In general, a few points close to the origin are sufficient to model the experimentally observed pattern, since the scattering potentials strongly favor scattering in the forward direction. We have to include a potential correction for the beams that cross the surface from the crystal to the vacuum since we assume that the pattern is generated in the bulk. The surface perpendicular component $k'_\perp$ of the diffracted wave vector $\mathbf{k}'$ outside the crystal is then given by

$$k'_\perp = \sqrt{k_\perp^2 - \frac{2me}{\hbar^2} V}. \tag{9}$$

The potential difference V is positive for an electron crossing the surface from the inside, leading to a reduction of its perpendicular wave vector component. Only electrons with a $k'_\perp$ vector larger than the potential correction are able to leave the crystal. The length of the vectors is measured with respect to the shadow edge position. We therefore obtain V as well as the shadow edge position as fitting parameters from an adaptation of the calculated Kikuchi line pattern to the experiment. In this way we can simultaneously obtain the misorientation of the substrate crystal and the average value of the crystal potential.

A fit for an exactly oriented (001) GaAs $\beta(2\times4)$ surface along $[\bar{1}10]$ is shown in Fig. 9(a). The zone axis of the Kikuchi pattern is marked by the square symbol. It does not coincide with the (000) spot of the elastic pattern, since there is no single incident beam direction for the Kikuchi line case. The origin of the reciprocal lattice for the Kikuchi line pattern is locked to the crystal zone axis, whereas the origin of the reciprocal lattice for the elastic pattern is locked to the transmitted beam position. Both, however, have the same orientation and geometry, as required by their construction. As expected for an exactly oriented substrate, the crystal zone axis coincides with the center of the Laue circles and is located on the shadow edge. To obtain a reliable fit, in a first step the overall size of the pattern is determined by adjusting the positions of the reciprocal lattice points parallel to the surface so that they lie on the corresponding streaks. The adjustment of the pattern perpendicular to the sample surface is then performed using the position of the origin and the potential value as fitting parameters. It is advantageous to take reference points far away from the shadow edge as well as features close to it, like the ones marked by the arrows in Fig. 9(a). The 004 line in the present case is most sensitive to the potential since it is parallel and close to the shadow edge. The potential is then adjusted so that the relative distance of the features close to and far from the shadow edge agrees with the experiment. For the simulation of Fig. 9(a), a potential value of 10.5 ± 0.5 V is obtained.

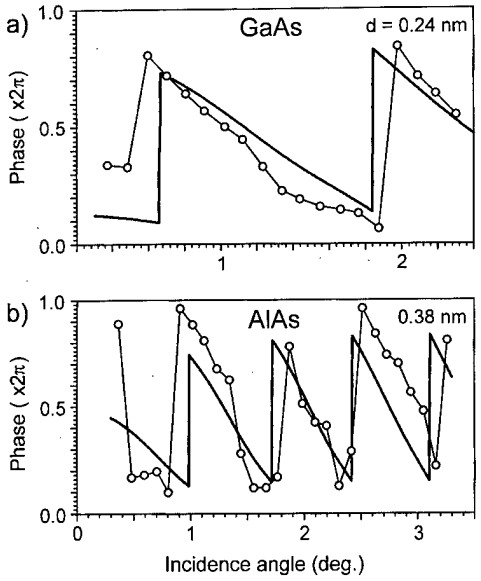


FIG. 10. Comparison of experimental data recorded along the [$\bar{2}$10] azimuth (circles) and thickness fits (thicker solid lines) obtained from the layer interference model. In (a), a value of $d=0.24$ nm is obtained; the result in (b) is $d=0.38$ nm. Potential $V=10.5$ V. GaAs was grown in the (2×4) surface reconstruction and AlAs in the $c(4\times4)$ surface reconstruction. The different thicknesses obtained for materials with identical bulk lattice constants lead to the model shown in Fig. 13.

The fit to an AlAs(001) $c(4\times4)$ surface along the same azimuth is shown in Fig. 9(b). Although the material and the surface reconstruction are different, the same potential value of 10.5 ± 0.5 V is obtained. We therefore use $V=10.5$ V in our simulations to model both GaAs and AlAs homoepitaxy.

V. PHASE DISPERSION IN THE ONE-BEAM CASE

To match an experiment to the layer interference model with its neglect of lateral potential modulation, experiments to be compared to the theoretical results of the layer interference model have to be performed at the one-beam condition. We chose a position close to the [$\bar{2}$10] azimuth of the GaAs/AlAs(001) surface. Measurements of the oscillation phase as a function of incidence angle are shown in Fig. 10 together with the theoretical results for both GaAs and AlAs homoepitaxy. The phase was determined by measuring the position of the oscillation minimum. GaAs was grown in the (2×4) surface reconstruction, and AlAs in the $c(4\times4)$. Apart from some discrepancies at very low angles, the agreement between model and experiment is excellent. The experiment not only explains the half period phase difference between generalized Bragg and off-Bragg conditions, but also matches the correct direction and approximate magnitude of the continuous phase shift inbetween the phase jumps. We are therefore confident that the model, despite its simplicity, correctly describes the main process governing the phase shift of RHEED oscillations.

The actual curve forms of the layer interference model and the experimental oscillations are compared in Figs. 11 and 12. Except for the spike at integer coverages that is always present, the curve forms match well. As mentioned

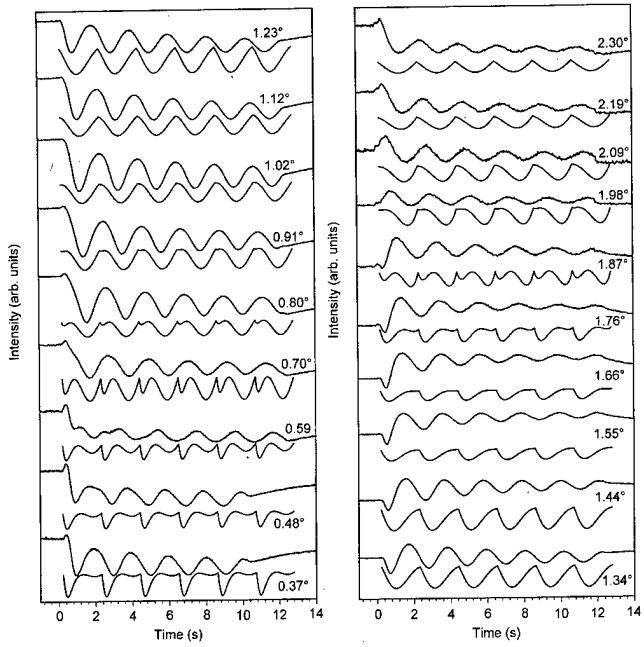


FIG. 11. Comparison of measured and simulated curve shapes for the GaAs data of Fig. 10(a).

before, the spikes could be removed by using a multilayer growth model. As can be inferred from the examples for Si and Ge,[21,22] a total of only a few percent of the coverage in adjacent layers effectively removes the spike, which results

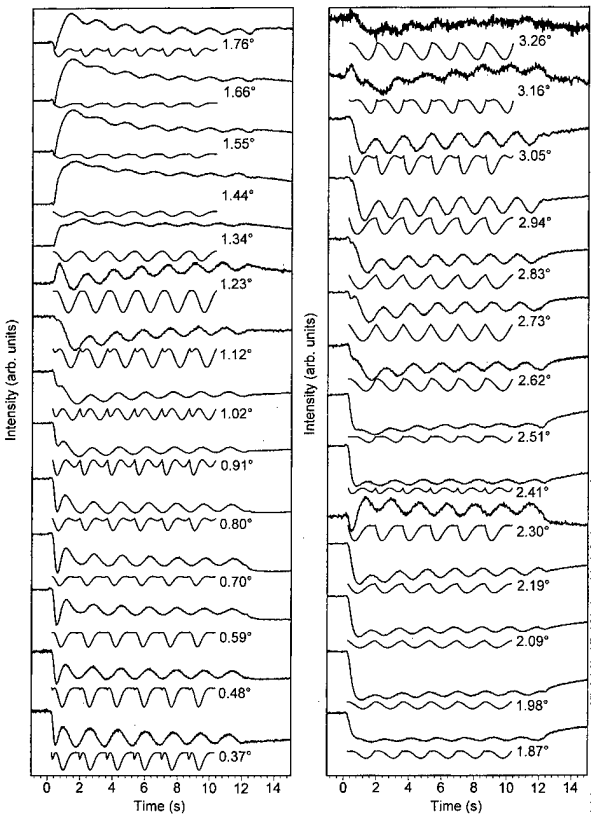


FIG. 12. Comparison of measured and simulated curve shapes for the AlAs data of Fig. 10(b).

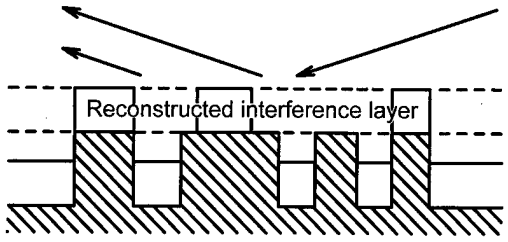


FIG. 13. Model for interference in the reconstructed layer on top of the growing bulk structure layer.

in a better fit. But even with single layer growth, all important features like the two minima at low angles for GaAs and the steady shift of the minima to the left are correctly reproduced.

GaAs and AlAs exhibit practically the same bulk lattice spacing. A fit with $d=0.283$ nm, however, does not agree with the experiments. Instead, a value of $d=0.24\pm0.02$ nm for GaAs and $d=0.38\pm0.02$ nm for AlAs is found. We believe that this is due to the surface reconstruction. The growing layer consists of both a layer with bulk structure and a surface reconstruction on top of the growing layer that forms on the islands as the layer is deposited. The reconstructed layer has a different structure and electronic configuration than the underlying bulk layer. In terms of the definition of an interface, the contrast is therefore more pronounced between a reconstruction and bulk structure growing layer than between the bulk structure growing layer and the substrate. If we approximate the crystal by one layer on a substrate like in our case, we therefore probably obtain a better result if we define the interface between the reconstruction and bulk structure layer than between the bulk structure layer and the substrate. This model is shown in Fig. 13. This model works well if we assume that the formation of the reconstruction on top of the growing bulk structure layer closely follows its coverage. If the reconstruction forms in patches smaller than the underlying islands, the match to the model should even be better, since then the approximation of a uniformly increasing average potential in the layer is met more effectively.

The values found for the layer thickness d agree remarkably well with structural data of the (2×4) and $c(4\times4)$ surface reconstructions,[30,31] which consists of ~1.0 and ~1.5 atomic bilayers, respectively. In both cases, the reconstructed layers are incomplete and slightly relaxed towards the bulk. The good match is another indication of the model's validity. The assumption of interference in the reconstructed layer is further supported by the behavior at the onset of the oscillations. For most incidence angles, the intensity during the oscillations is lower than for the static surface prior to growth. This behavior is inverted at the generalized Bragg conditions in both Figs. 11 and 12. Here, the average level of the oscillations is higher than the nongrowing level. The positions are around 1.9° for GaAs and 1.8° and 2.4° for AlAs. In the model, the generalized Bragg conditions correspond to a reflectance maximum at half layer coverage. This indicates that for a realistic surface, where the nongrowing surface is also covered by a reconstructed layer, the nongrowing surface should show a reflectance minimum due to destructive interference in the static reconstruction layer. When growth is initiated, the reflectance increases, resulting in higher intensity during growth.

Along high symmetry azimuths, the phase dispersion assumes a more complicated form,[4] since higher-order beams and surface resonances contribute much more strongly to the diffracted intensity. Also, the degree of multiple scattering changes for well-ordered surfaces along high-symmetry azimuths. Some special cases can even be approximated reasonably well by purely kinematical treatments.[32,33] RHEED oscillations in this case are very weak and only detectable at the kinematically forbidden reflection, again emphasizing the multiple scattering nature of the RHEED oscillation process.[23,34] For less ordered surfaces, kinematical contributions may still play an important role, accounting for features at unshifted Bragg positions that appear together with the shifted generalized Bragg conditions of the layer interference model, and they complicate the picture. Theoretical calculations[35] indicate that the potential in the top layer of the crystal is a strong function of azimuthal angle, which could account for the observed phase variation upon substrate rotation.[36] Last, not least, the layer interference model directly explains the dependence of the GaAs oscillation phase on As_4 overpressure,[5] since a variation in the As_4 pressure changes the surface reconstruction as well as the relative As content within one surface reconstruction.[30,37] This would directly change the phase of the oscillations through a change of d or V, or both.

VI. CONCLUSIONS

Based on ELS-RHEED measurements of RHEED oscillations and RHEED oscillations as a function of incidence angle, we propose a basic model for the occurrence of RHEED intensity oscillations. The model is based on a simplified version of a model introduced by Horio and Ichimiya.[21] We find that the experiments are explained by interference in the reconstruction layer on top of the growing bulk structure layer in the layer-by-layer growth mode. This layer interference model with only two free parameters, one of which can be determined by independent diffraction experiments, agrees remarkably well with measurements in the one-beam condition for both GaAs and AlAs. In addition, it explains several experimental phenomena in a unified coherent scattering approach.

The model not only explains the occurrence of different phases,[4] but also reproduces the correct direction and magnitude of the continuous phase shift between the generalized Bragg conditions in the one-beam condition. It reproduces the reduction in intensity with increasing intensity as well as the periodic variation in the oscillation amplitude.[11] It eliminates the necessity of postulating a second, independent mechanism that would lead to phase shifts when combined with Bragg scattering. This mechanism has been shown to be absent for inelastic scattering by recording RHEED oscillations as a function of energy loss. The layer interference

model also explains the occurrence of sharp additional maxima for RHEED oscillations on very smooth surfaces,[24] large amplitude oscillations with extra minima at low glancing angles,[25,26] and the dependence of the GaAs oscillation phase on As_4 pressure.[5] The model seems to be a promising starting point from which to develop more elaborate models for the many-beam case as well as models that treat surface reconstruction in a more realistic way.

ACKNOWLEDGMENTS

The authors would like to thank K. Hagenstein, M. Höricke, J. Müllhäuser, M. Reiche, and P. Schützendübe for valuable assistance and Y.-H. Zhang and A. Ichimiya for valuable discussions. This work was supported by the Volkswagenstiftung.

[1]J. J. Harris, B. A. Joyce, and P. J. Dobson, Surf. Sci. **103**, L90 (1981).
[2]C. E. C. Wood, Surf. Sci. **108**, L441 (1981).
[3]J. M. Van Hove, C. S. Lent, P. R. Pukite, and P. I. Cohen, J. Vac. Sci. Technol. B **1**, 741 (1983).
[4]J. Zhang, J. H. Neave, P. J. Dobson, and B. A. Joyce, Appl. Phys. A: Solids Surf. **42**, 317 (1987).
[5]F. Briones, D. Golmayo, L. Gonzalez, and J. L. De Miguel, Jpn. J. Appl. Phys., Part 2 **24**, L478 (1985).
[6]D. M. Holmes, J. L. Sudijono, C. F. McConville, T. S. Jones, and B. A. Joyce, Surf. Sci. **370**, L173 (1997), and references therein.
[7]S. L. Dudarev, D. D. Vvedensky, and M. J. Whelan, Phys. Rev. B **50**, 14 525 (1994).
[8]S. L. Dudarev, D. D. Vvedensky, and M. J. Whelan, Surf. Sci. **324**, L355 (1995).
[9]J. H. Neave, B. A. Joyce, P. J. Dobson, and N. Norton, Appl. Phys. A: Solids Surf. **31**, 1 (1983).
[10]P. J. Dobson, B. A. Joyce, J. H. Neave, and J. Zhang, J. Cryst. Growth **81**, 1 (1987).
[11]P. I. Cohen, P. R. Pukite, J. M. Van Hove, and C. S. Lent, J. Vac. Sci. Technol. A **4**, 1251 (1986).
[12]J. Sudijono, M. D. Johnson, C. W. Snyder, M. B. Elowitz, and B. G. Orr, Phys. Rev. Lett. **69**, 2811 (1992).
[13]U. Korte and P. A. Maksym, Phys. Rev. Lett. **78**, 2381 (1997).
[14]Y. Horio, Y. Hashimoto, K. Shiba, and A. Ichimiya, Jpn. J. Appl. Phys., Part 1 **34**, 5869 (1995).
[15]Y. Horio, Y. Hashimoto, and A. Ichimiya, Appl. Surf. Sci. **100/101**, 292 (1996).
[16]W. Braun, L. Däweritz, and K. H. Ploog, Proceedings of the Conference on Modulated Semiconductor Structures 8, Santa Barbara, CA, July 1997 (to be published).
[17]"SAFIRE" by VTS Schwarz GmbH, Hindenburgstrasse 12, D-76332 Bad Herrenalb, Germany; electronic mail: vts.js@t-online.de
[18]C. J. Powell, Phys. Rev. **175**, 972 (1968).
[19]J. L. Beeby, in *Reflection High-Energy Electron Diffraction and Reflection Electron Imaging of Surfaces*, edited by P. K. Larsen and P. J. Dobson, NATO ASI Series B, Vol. 188 (Plenum, New York, 1988), p. 77.
[20]G. Lehmpfuhl, A. Ichimiya, and H. Nakahara, Surf. Sci. **245**, L159 (1991).
[21]Y. Horio and A. Ichimiya, Surf. Sci. **298**, 261 (1993).
[22]Y. Horio and A. Ichimiya, Jpn. J. Appl. Phys., Part 2 **33**, L377 (1994).
[23]W. Braun, Ph.D. thesis, Berlin, 1996.
[24]L. Däweritz, J. Cryst. Growth **127**, 949 (1993).
[25]Z. Mitura, M. Strózak, and M. Jalochowski, Surf. Sci. Lett. **276**, L15 (1992).
[26]Z. Mitura and A. Daniluk, Surf. Sci. **277**, 229 (1992).
[27]S. Kikuchi, Jpn. J. Phys. **5**, 83 (1929).
[28]P. Dobson, *Surface and Interface Characterization by Electron Optical Methods*, edited by A. Howie and U. Valdrè (Plenum, New York, 1988), p. 159.
[29]M. Gajdardziska-Josifovska and J. M. Cowley, Acta Crystallogr., Sect. A: Found. Crystallogr. **47**, 74 (1991).
[30]T. Hashizume, Q.-K. Xue, A. Ichimiya, and T. Sakurai, Phys. Rev. B **51**, 4200 (1995), and references therein.
[31]D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L.-E. Swartz, Phys. Rev. B **41**, 5701 (1990).
[32]Y. Horio and A. Ichimiya, Surf. Sci. **219**, 128 (1989).
[33]W. Braun, O. Brandt, M. Wassermeier, L. Däweritz, and K. Ploog, Appl. Surf. Sci. **104/105**, 35 (1996).
[34]D. Lüerssen, A. Dinger, H. Kalt, W. Braun, R. Nötzel, K. H. Ploog, J. Tümmler, and J. Geurts, Phys. Rev. B **57**, 1631 (1998).
[35]A. Ichimiya (private communication).
[36]J. Resh, K. D. Jamison, J. Strozier, A. Bensaoula, and A. Ignatiev, Phys. Rev. B **40**, 11 799 (1989).
[37]L. Däweritz, Superlattices Microstruct. **9**, 141 (1991).

Anisotropic microstructure development during the reaction of Mg with GaAs

S. W. Robey[a)]
National Institute of Standards and Technology, Gaithersburg, Maryland 20899

(Received 21 January 1998; accepted 28 May 1998)

Reaction of Mg with the GaAs(001) surface occurs at substrate temperatures above 550 K and leads to three-dimensional (3D) growth of an epitaxial, cubic reaction phase with a lattice constant of 0.62 ± 0.02 nm, 9% larger than GaAs. The resulting anisotropic, 3D heteroepitaxial microstructure was studied using *in situ* electron diffraction and Auger spectroscopy and *ex situ* atomic force microscopy (AFM). Three-dimensional structures develop that are elongated along the GaAs[110] direction with aspect ratios (length to width) up to 20. These structures appear to be composed of isotropic islands from which elongated, tapered ridges extend in one direction along [110]. Analysis of AFM images and size distributions suggest a critical size for the initiation of ridge growth. We speculate that these structures develop in response to misfit strain in the reaction phase. [S0734-211X(98)12404-6]

I. INTRODUCTION

The purpose of this article is to describe an interesting growth morphology we believe occurs due to strain when an epitaxial overlayer is formed by reaction of Mg with a GaAs(001) substrate. Previously, Mg was found to react with GaAs(001) at temperatures greater than 550 K,[1] forming a three-dimensional (3D) epitaxial reaction phase with a large lattice mismatch with the substrate. Three-dimensional growth occurred under both isothermal conditions (deposition of Mg at substrate temperatures above the reaction temperature) and isoconcentration conditions (annealing previously deposited Mg overlayers). Under isoconcentration conditions, the resulting reaction developed a highly anisotropic surface mophology comprised of linear "islands" oriented with a long growth direction along GaAs[110] and length-to-width ratios up to 20. We suggest that strain is a factor leading to this anisotropic growth.

II. EXPERIMENTAL PROCEDURES

The experiments described here were performed in an ultrahigh vacuum (UHV) molecular beam epitaxy (MBE) and analysis system. Reflection high energy electron diffraction (RHEED) and Auger spectroscopy, performed with a single pass cylindrical mirror analyzer (CMA), were the primary *in situ* analysis techniques. Together they provided thin film compositional and structural information. *Ex situ* atomic force microscopy (AFM) measurements provided information on surface morphology.

The GaAs(001) surface was prepared by H_2 plasma cleaning with a remote plasma source operating at 2–3 Pa and 20–50 W of rf power at 13.56 MHz. Exposure for about 20 min at substrate temperatures between 500 and 550 K resulted in complete removal of the native surface oxide. The substrate was subsequently introduced into the deposition/analysis chamber where it was annealed to 800 K to produce a well-ordered 4×2 diffraction pattern indicative of a Ga-stabilized surface.

Deposition of Mg on GaAs(001) was performed at pressures below 5×10^{-8} Pa using a conventional MBE source with a boron nitride crucible. The Mg flux was determined with a quartz crystal microbalance (QCM) at the sample position. Typical deposition rates were in the range 0.2–0.5 nm/min. Reaction of Mg with the GaAs substrate was induced either by deposition at substrate temperatures above about 550 K (isothermal conditions) or by Mg deposition below this temperature, followed by annealing (isoconcentration conditions). The substrate temperature was monitored continuously using a thermocouple attached to the sample holder.

The AFM images were acquired in contact mode using silicon nitride tips with radii of approximately 20 nm. Except where noted, we will report surface feature dimensions that are uncorrected for this finite resolution. Measurements were repeated with several tips to reduce the opportunity for misinterpretation due to tip artifacts. Alignment using the cleaved edges of the samples, coupled with RHEED observations of the 4×2 reconstruction, allowed determination of the GaAs[110] and [$1\bar{1}0$] directions to within a few degrees during imaging. Imaging was initiated within 1 h after removal of the samples from the UHV chamber to minimize the effects of oxidation.

III. RHEED AND AUGER RESULTS

Mg deposition on GaAs surfaces at substrate temperatures below 500 K resulted in a loss of intensity in the GaAs diffraction and an attendant increase in the background. Sharp transmission diffraction spots, indicating the formation of a 3D surface morphology, became apparent after annealing the Mg overlayers to temperatures above 550 K. RHEED data illustrating this development are provided in Fig. 1. The diffraction exhibited fourfold azimuthal symmetry indicating a cubic structure. In addition to the strong diffraction features evident in Fig. 1, very weak spots were often visible. Some of these are can be seen in Fig. 1 at half-order posi-

[a)]Electronic mail: steven.robey@nist.gov

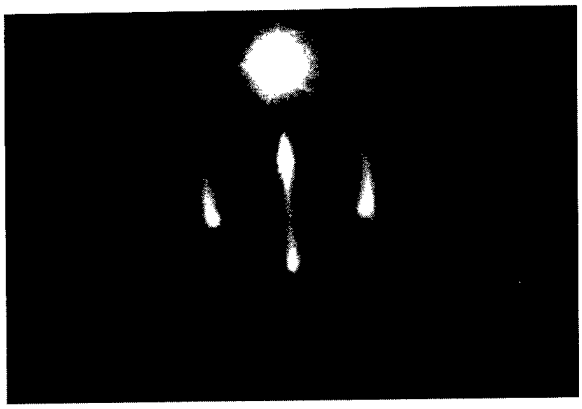


FIG. 1. RHEED images along the [1$\bar{1}$0] azimuth after deposition of 1 nm of Mg at 293 K followed by annealing to above 550 K to induce a reaction with the GaAs substrate.

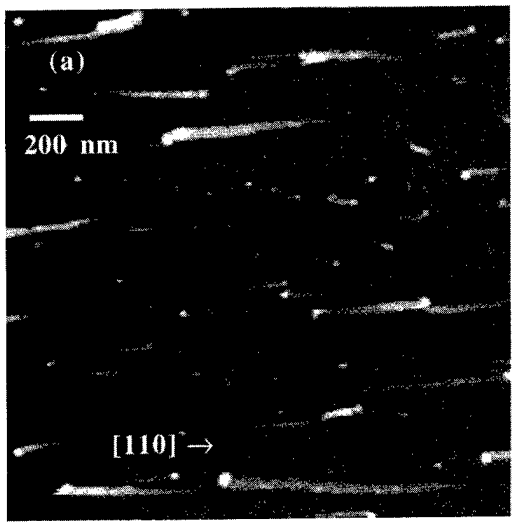


FIG. 2. (a) 2000 nm×2000 nm contact AFM image of the GaAs(001) surface after deposition and reaction of 1 nm of Mg. The light-to-dark contrast in the z scale spans a range of about 20 nm. The horizontal direction of the image (the direction of elongation of the surface structures) is [110].

tions relative to the strongest diffraction features while other weaker spots are not visible except at much higher incident beam intensity. All of the diffraction could be related a single cubic lattice parameter. No evidence for multiple phases was observed. For the purpose of the following discussion, we have chosen to index the diffraction in terms of the most intense spots. This choice produces a face-centered-cubic (fcc) structure with a lattice constant of 0.62 $\pm$0.02 nm. For thick films, Auger spectroscopy continued to indicate significant quantities of Ga and As, suggesting a Mg–Ga–As ternary compound. AFM revealed rough surfaces for these thick films, but no evidence for regions of exposed substrate to produce Ga and As intensity. No potential Mg–Ga–As ternary compounds could be discovered in the literature, but Mg–silicide and –germanide, Mg_2Ge, do occur in a cubic antifluorite structure with lattice constants of 0.63 (Mg_2Si) and 0.64 nm (Mg_2Ge).[2]

IV. AFM MEASUREMENTS OF SURFACE MORPHOLOGY

The origin of the transmission diffraction in RHEED is readily identified in AFM images. The growth of large 3D features, with their striking characteristics, is illustrated in the 2×2 μm^2 AFM image provided in Fig. 2. The most distinctive features are large anisotropic structures elongated along the GaAs[110] direction. Deposition of, in this case, 0.7 nm of Mg with subsequent annealing led to the formation of these 3D (typical heights are 5–15 nm) elongated structures. Some of the larger structures are 500–1000 nm in length along [110] and have mean widths varying from 40 to 100 nm, leading to in-plane aspect ratios up to 20. Typically the length-to-width aspect ratio decreases with the size of the feature, with many of the smallest structures appearing to be nearly isotropic.

The asymmetry of the features in Fig. 2 does not arise simply by a elongation in both directions along [110]. Instead, the elongation is more in the form of a wedgelike or needle shape extending in one direction along [110]. Typically there appears to be a nearly isotropic "island" at the largest end, from which a tapered "ridge" extends along [110]. The ridge reduces in height and narrows in width toward the end. Although, with few exceptions, a given structure extends in only one direction along [110], the overall symmetry along [110] is preserved because, on average, equal numbers extend in either direction.

The image in Fig. 2 reveals further interesting characteristics. In many cases there is an offset of about 10–20 nm between the ridge and the center of the terminating island in the direction perpendicular to the length of the ridge, i.e., along [1$\bar{1}$0]. Once again, there is no detectable bias in the direction of this shift, either in one direction or the other along [1$\bar{1}$0]. Another interesting characteristic is the quasi-periodic variation in height along the length of the ridge, which produces a segmented appearance with a mean segment length on the order of 60 nm. This can be seen with appropriate adjustment of the contrast in the AFM images, but it is most easily illustrated in magnified images such as those discussed below.

Smaller features observed in the AFM images provide clues to the nucleation and development of this unique growth morphology. The image provided in Fig. 3 shows detail of one such small feature that, as in the case of larger structures, consists of an isotropic island, about 50 nm in diameter and 10 nm high, with a ridge 100 nm long. The offset between the ridge and the island center is clearly evident. The ridge again consists of segments each approximately 50–60 nm in length, close to the mean segment length typically observed for larger features. The first step, or segment, is approximately 38–40 nm wide and 5 nm high while the second is 30 nm wide and 2 nm high. A cross section along the line $A-A'$, more clearly illustrating this steplike structure, is provided below the image. These observations suggest a discontinuous, stepwise development in 50–60 nm segments to produce elongation along [110].

Examples of incipient ridge development can also be seen in the AFM images. These features, along with the nearly

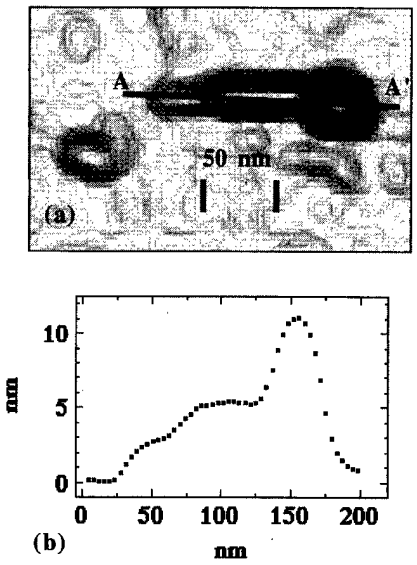


FIG. 3. (a) AFM image showing a small structure in the early stages of elongation. The line profile provided in (b) again illustrates the segmented development seen in previous figures. This small structure consists of a nearly isotropic end "island," 60 nm in diameter, and a ridge composed of two segments 50 nm in length.

isotropic character of most of the small islands, suggest that the initial growth is isotropic until, at a critical size, growth of the attached ridge is triggered. Based on measured diameters of small, isotropic features, this critical diameter is around 30–60 nm. To further test this possibility, the island base diameter (the diameter of the terminating island for elongated structures) versus the length of the ridge is plotted in Fig. 4. The trend in the data again suggests a critical value for the diameter at which elongation along [110] begins. A simple linear fit appears to approximate the data reasonably well and indicates a critical diameter of 40 ± 5 nm. (This value does not account for the finite resolution of the AFM tip. Assuming a Gaussian convolution with a conservative estimate of the tip resolution of 20 nm leads to an actual critical diameter of 35 nm.)

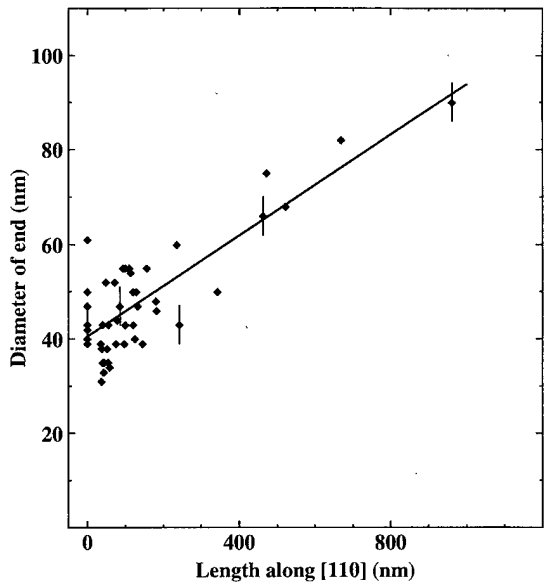


FIG. 4. Plot of the diameter of the isotropic, terminating island vs the length of the attached ridge. The trend in the plot indicates a critical diameter at which ridge development is initiated. A simple linear plot produces a value of 40 nm for this critical diameter.

V. DISCUSSION

From the observations related above, we now suggest a scenario for the development of the microstructure during reaction. Based on the isotropic nature of the smallest observable features, the initial growth does not experience any strong degree of anisotropy. Growth continues in this isotropic fashion until a critical size of about 40 nm in diameter and 4 nm in height. At this point, anisotropic growth along [110] is initiated and thereafter proceeds in a rapid fashion compared to continued growth of the original nucleus. Growth along [110] is not smooth and continuous but, instead, occurs in a steplike fashion creating segments approximately 50 nm long (along [110]) that sequentially develop and grow.

A variety of processes has been suggested as sources leading to the formation of anisotropic structures during thin film growth, including anisotropic accommodation at a particular facet or step direction,[3–5] alignment along step edges,[6] and strain.[7,8] There are several observations that suggest the importance of strain in this case. The epitaxial alignment to the substrate with a lattice constant 9% larger indicates there must be significant compressive strain at the interface. Also, the tendency to form large aspect ratio 3D structures is consistent with the minimization of interfacial contact and elastic relaxation observed in other strained layer systems.[9–12] The height-to-width aspect ratio was observed to vary significantly with the size of the island/structure. This is consistent with strain as a predominant factor in producing 3D growth, as opposed to a large difference in substrate and overlayer surface energies. We also note a strong similarity in the morphology observed here, particularly for thicker films, with the "ripple" structure that develops to relieve strain in InGaAs/GaAs (Ref. 13) and SiGe/Si(100).[14] These observations lead us to consider strain as the primary source of anisotropy. In the following we suggest possible explanations based on anisotropic relaxation of strain during growth.

For the purpose of further discussion, the basic premise will be that there is an asymmetry between [110] and [$1\bar{1}0$] in the ability to relieve strain. For instance, there may be a lower barrier for nucleation of a dislocation/dislocations to relieve strain in the [110] direction than for the [$1\bar{1}0$]. We note that an asymmetry of this type is present in sphalerite structures, such as GaAs, where the barrier to nucleation of 60° dislocations is much lower in the [110] direction than in the [$1\bar{1}0$].[15] Elongated growth in the [110] direction is then a consequence of an increased rate of strain relaxation in this direction.

Building on this premise, a number features of the growth can be explained. During the initial growth, strain is accommodated primarily by elastic relaxation in the 3D islands. As continued growth occurs, the islands reach a point where the accumulated strain is large enough to overcome the kinetic

barrier to formation of dislocations, as is observed in other systems.[9,12,16,17] This occurs when the islands are approximately 40 nm in diameter and 4 nm high, although variation in this critical size may occur with conditions such as temperature, etc. The (assumed) relative ease of formation of dislocations which relieve strain in the [110] direction leads to subsequent, faster growth along [110]. The observed critical size is thus explained as the onset above which predominantly dislocated growth occurs to form elongated ridges.

A discontinuous, stepwise growth has been observed previously in transmission electron microscopy (TEM) observations of strained layer growth.[10,18] LeGoues *et al.*[18] followed the development of Ge islands on Si(100) in real time under conditions of constant flux. Initially the islands relaxed elastically, maintaining coherence with the substrate. Eventually, accumulated strain led to nucleation of dislocations and subsequent growth at a rate faster than surrounding purely coherent islands. Growth then proceeded in a discontinuous manner as each additional dislocation was introduced. The islands grew primarily vertically until introduction of a new dislocation. Nucleation of a new dislocation caused a rapid evolution to a new, lower energy shape, with the island spreading approximately 20 nm beyond the point of dislocation nucleation. We note, however, that in this case the mean observed segment length is approximately one order of magnitude too large compared to the separation expected for pure edge dislocations to completely relax the misfit strain.

We can also suggest a potential explanation for the offset observed between the ridge and the isotropic end island. The strain profile in the elastically relaxed island is not homogeneous. Finite element calculations have shown that there is a significant increase in strain near the island edges.[19–21] Therefore, nucleation of dislocations can be expected to occur preferentially at the island edges. This expectation has been corroborated by cross-sectional TEM measurements.[22,23] Thus, if initiation of ridge growth is correlated with strain relaxation, for instance, by nucleation of a dislocation, then we would expect that this would occur predominantly on the periphery of the island, producing an offset.

Another possibility for the formation of these elongated features is suggested by the observation of similar ridge structures during the early stages of heteroepitaxy of InGaAs on GaAs(001).[13] Elongated growth, prior to the development of surface rippling, was observed and correlated with an enhanced growth rate in the shear field above and along the length of underlying dislocations (see, for example, Fig. 5 of Ref. 13). Although the underlying cause is still strain, the suggested mechanism is somewhat different from that speculated above. Ridge formation was observed preferentially along the [110] direction, again due to the asymmetry between [110] and $[1\bar{1}0]$ for dislocation nucleation and propagation in the sphalerite structure. Alignment of 3D islands above pre-existing dislocation lines has also been observed in Ge, SiGe/Si(001).[7] At this point there is not enough information to favor one of these explanations over the other or to consider possible connections between them.

VI. SUMMARY

We have attempted to provide some insight into the factors producing the interesting and rather unique microstructure observed under some conditions for the reaction of Mg with the GaAs(001) surface. The reaction phase appears to be a ternary Mg–Ga–As compound that grows epitaxially on GaAs(001) with a lattice mismatch of approximately 9%. Examination of the anisotropic features and volume distributions indicates a critical size at which anisotropic growth is initiated. Strain is suggested a factor in producing anisotropic growth along the GaAs[110] direction. Although many questions still remain and it is possible that other factors play a role, the evidence available suggests a mechanism for the anisotropic growth rooted in an asymmetry in the strain relaxation for the reaction phase between the GaAs[110] and $[1\bar{1}0]$ directions.

[1]S. W. Robey, J. Vac. Sci. Technol. A (accepted).
[2]E. N. Nikitin, E. N. Tkalenko, V. K. Zaitsev, A. I. Zaslavskii, and A. K. Kuznetsov, Inorg. Mater. **4**, 1656 (1968); K. A. Bol'shakov, N. A. Bul'onkov, L. N. Rastorgaev, and M. S. Tsirlin, Russ. J. Inorg. Chem. **8**, 1418 (1963).
[3]Y.-W. Mo, B. S. Swartzendruber, R. Kariotis, M. B. Webb, and M. G. Lagally, Phys. Rev. Lett. **63**, 2393 (1989).
[4]L. Porte, P. Krapf, Y. Robach, M. Phaner, M. Gendry, and G. Hollinger, Surf. Sci. **352**, 60 (1996).
[5]T. Nomura, K. Ishikawa, K. Murakami, and M. Hagino, J. Cryst. Growth **127**, 584 (1993).
[6]M. Taskinen, M. Sopanen, H. Lipsanen, J. Tulkki, T. Tuomi, and J. Ahopelto, Surf. Sci. **376**, 60 (1997).
[7]S. Yu. Shiryaev, F. Jensen, J. L. Hansen, J. W. Petersen, and A. N. Larsen, Phys. Rev. Lett. **78**, 503 (1996).
[8]V. Bressler-Hill, A. Lonke, S. Varma, P. M. Petroff, K. Pond, and W. H. Weinberg, Phys. Rev. B **50**, 8479 (1994).
[9]D. J. Eaglesham and M. Cerullo, Phys. Rev. Lett. **64**, 1943 (1990).
[10]M. Hammar, F. K. LeGoues, J. Tersoff, M. C. Reuter, and R. M. Tromp, Surf. Sci. **349**, 129 (1995).
[11]M. Goryll, L. Vescan, K. Schmidt, S. Mesters, H. Luth, and K. Szot, Appl. Phys. Lett. **71**, 410 (1997).
[12]H. T. Johnson and L. B. Freund, J. Appl. Phys. **81**, 6081 (1997).
[13]A. G. Cullis, A. J. Pidduck, and M. T. Emeny, Phys. Rev. Lett. **75**, 2368 (1995); A. G. Cullis, A. J. Pidduck, and M. T. Emeny, J. Cryst. Growth **158**, 15 (1996).
[14]D. E. Jesson, K. M. Chen, S. J. Pennycook, T. Thundat, and R. J. Warmack, Science **268**, 1161 (1995).
[15]M. S. Abrahams, J. Blanc, and C. J. Buiocchi, Appl. Phys. Lett. **21**, 185 (1972); K. L. Kavanagh *et al.*, J. Appl. Phys. **64**, 4843 (1988).
[16]R. Hull and A. Fischer-Colbrie, Appl. Phys. Lett. **50**, 851 (1987).
[17]B. Voigtlander and A. Zinner, Appl. Phys. Lett. **63**, 3056 (1993).
[18]F. K. Legoues, M. C. Reuter, J. Tersoff, M. Hammar, and R. Tromp, Phys. Rev. Lett. **73**, 300 (1994).
[19]A.-L. Barabasi, Appl. Phys. Lett. **70**, 2565 (1997).
[20]Y. Chen and J. Washburn, Phys. Rev. Lett. **77**, 4046 (1996).
[21]S. Christiansen, M. Albrecht, H. P. Strunk, and H. J. Maier, Appl. Phys. Lett. **64**, 3617 (1994).
[22]S. Guha, A. Madhukar, and K. C. Krajkumar, Appl. Phys. Lett. **57**, 2110 (1990).
[23]H. L. Tsai and J. Matyi, Appl. Phys. Lett. **55**, 265 (1989).

Equilibrium critical thickness for strained-layer growth

K. E. Khor and S. Das Sarma
Department of Physics, University of Maryland, College Park, Maryland 20742-4111

(Received 21 January 1998; accepted 26 May 1998)

The critical thickness for two-dimensional layer growth of $Si_{1-x}Ge_x$ on Si and $In_xGa_{1-x}As$ on GaAs is about 1–3 monolayers (ML) for $x=1$, beyond which islanding begins. Under certain growth conditions, this thickness t would increase as some power of $1/x$. The reason for this is not clear; Snyder *et al.* argue that, under *equilibrium* conditions, this critical Stranski–Krastanov (SK) thickness t_c is independent of x and should remain at 1–3 ML, but that, under nonequilibrium growth conditions, $t \sim x^{-4}$. However, Osten *et al.* showed, experimentally, that even under equilibrium conditions there is an increase of SK thickness with $1/x$. We carry out calculations of energetics of large three-dimensional (3D) islands on substrates with varying thicknesses t of the epilayer and different coverages θ. We show that at low θ or when islands are small (or both) then the SK thickness is small ~1–3 ML, in agreement with the results of Snyder *et al.* At increasing coverages, when interisland separation l decreases to the point where $l \sim s$ (island size), we observe ΔE to decrease for the lower thicknesses $t=3,4,5,...8...$; until thicknesses $t>3$ become more favorable. There is considerable tension going deep into the substrate directly below islands. The larger an island becomes, the more favorable it is for a thicker layer beneath it to be of the same material as the island. It is known that the critical size s_c at which 3D islands first become favorable varies as x^{-2}. We argue from this that, at equilibrium, the average 3D island size increases with x^{-p}, p some exponent, and at high enough coverages, when interisland separation is small, SK thickness t_c increases. The experimental results of Osten *et al.* are consistent with exponents $2 \leqslant p \leqslant 4$. © *1998 American Vacuum Society.* [S0734-211X(98)12804-4]

I. INTRODUCTION

We will look at the energetics of island growth on thin films deposited on lattice mismatched substrates. It is known that many of these systems grow in the Stranski–Krastanov (SK) mode, with three-dimensional (3D) islanding starting after layer by layer growth of a few monolayers (ML). The critical thickness for the growth of Ge on Si(001), with lattice mismatch of 4%, is about 3 ML (Refs. 1–4) and for InAs on GaAs (mismatch 7%) it is 1–2 ML.[5] Beyond the critical thickness 3D islands appear and may grow into the ''hut'' clusters' observed by Mo *et al.*,[1] with well defined (105) facets, tilted 11.3° to the surface, or into larger islands with $(11n)$-like faces,[6] depending on the growth conditions. Although the critical thickness for the above systems is 1–3 ML, it is known that for the growth of the corresponding less strained systems, $Si_{1-x}Ge_x$ on Si and $In_xGa_{1-x}As$ on GaAs films, the critical thickness t_c may exceed 3 ML, increasing as x decreases, depending on the growth conditions.[7,8] Snyder *et al.*[8] showed that $t_c \sim x^{-4}$ under *nonequilibrium* growth conditions; there is a minimum growth temperature T_c, below which 3D islanding is not seen.[7–9] At equilibrium or near equilibrium conditions they argue that t_c should be independent of x and remain at 1–3 ML for all x. The reason for this is that equilibrium SK critical thickness is apparently dependent only on interface energetics; the energy of an *isolated* island, $E=E_s+E_r$ (E_s is the extra surface energy of the island, over the energy of the surface it is on, and E_r is the energy of elastic relaxation) is almost independent of the thickness. We have seen in previous work[10] that in layer by layer growth of Ge on Si(001) the energy cost of adding an extra layer for 1–3 ML is much lower than that (~25 meV/atom) for subsequent layers. This is also true for layer by layer growth of $Si_{1-x}Ge_x$ on Si, suggesting that islanding under equilibrium conditions should be independent of x and should occur after the first few ML's if this SK thickness is due to interface energetics only. However there is experimental evidence to indicate that even under equilibrium (or near equilibrium) conditions the critical SK thickness can be greater than 3 ML for $x<1$. Osten *et al.*[11] showed that after the deposition of 30–60 ML of $Si_{1-x}Ge_x$ on Si(001) and then annealing at 600 °C for 30 min, the average SK thickness increases with x^{-1}. Hansson *et al.*[6] used liquid phase epitaxy to grow psuedomorphic $Si_{1-x}Ge_x$ on Si(001) at 800 °C; this technique allows growth close to thermodynamic equilibrium. The growth that they obtained was of the SK type and islands are truncated tetrahedral pyramids, with (111) side facets and axes along the [110] directions. For $x=0.85$, they showed that SK thickness was 8–9 ML. As has been used by the various groups above,[6,8,11] we take the term ''SK critical thickness'' to mean the thickness t_c just before islanding begins under nonequilibrium growth conditions or the interisland thickness of the wetting layer under equilibrium growth conditions.

The aim of our work here is to investigate the origin of this enhanced SK critical thickness under equilibrium or near equilibrium conditions. The results of the two groups above[6,11] indicate that enhanced thickness is found under conditions of coverage θ, high enough for interisland separation to be of the order of the size of the islands themselves. Harker *et al.*[12] have carried out finite element (FE) calcula-

tions of stress and strain in stripes of $Si_{0.5}Ge_{0.5}$ on a Si substrate. They show that there is considerable relaxation going deep into the substrate directly beneath the stripes. Similar results have also been observed by Christiansen *et al.*[13] for (111)-faceted Ge islands on Si using FE methods. For stripes with aspect ratios (height/width) of the order of 1, similar to the aspect ratios of islands seen in the experiments of Hansson *et al.*[6] and of Osten *et al.*,[11] Harker *et al.*[12] showed that the average stress in the Si substrate directly underneath the stripe is tensile and even as deep as one tenth of the height of the stripe this stress has a magnitude 20%–40% of the stress in a flat $Si_{0.5}Ge_{0.5}$ layer. (For the sizes of islands seen experimentally,[6,11] one tenth of the height ~20 layers.) This stress in the substrate becomes compressive at and just beyond the edges of the stripe with smaller average magnitude. We expect similar behavior for the region of the substrate beneath islands, so that the larger and taller they are the more energetically favorable it is for a *thicker* layer beneath the island to be of the same material as that of the island. This should, in general, tend to increase the SK critical thickness; the effect should be most observable at high island density and it should tend towards zero with this density.

In this work, we first show that SK critcal thickness t_c remains at about 3 ML for values of $x<1$ for low island densities (or low coverage θ). Recent calculations[10] have shown that $t_c=3$ ML for $x=1$, so we need to carry out a similar calculation for a smaller value of x, say, 0.5. Then we display results for the energetics of (low density) islanding for a small fixed small coverage ($\theta=6$ ML), for islands on 3, 4, and 5 ML of flat Ge. We show that once 3D islanding becomes favorable, the most favorable configuration is that of islands on 3 ML. Finally we show results for high coverage $\theta=30$–40 ML when for a fixed island size, similar to those seen in the experiments above,[6,11] interisland separation is comparable to island size. We show that, indeed, configurations with islands on Ge thickness $t>3$ ML can become more favorable. We discuss the results and show how they are relevant to the observation of equilibrium SK thickness increasing with decreasing x.

II. RESULTS AND DISCUSSION

We calculate the energetics of the growth of Ge on Si(001) by means of molecular dynamics simulations. We employ empirical potentials that we have developed for Si–Ge problems.[9,10] These potentials have been tuned to reproduce the energetics and surface stresses of a number of reconstructions on the Si(001) surface.[10] We follow previous studies[8,10,14] and carry out simulations for clusters in (1+1) dimensions [or equivalently 3D stripes in (2+1) dimensions since we use periodic boundary conditions in the (001) plane]. Since the primary effects to be studied are essentially results of elastic relaxation, our conclusions should be valid for 3D islands in (2+1) dimensions. This claim is further strengthened by the results of Tersoff and Tromp[15] who have shown that for (2+1) dimensions, each dimension contributes an additive term of exactly the same form to

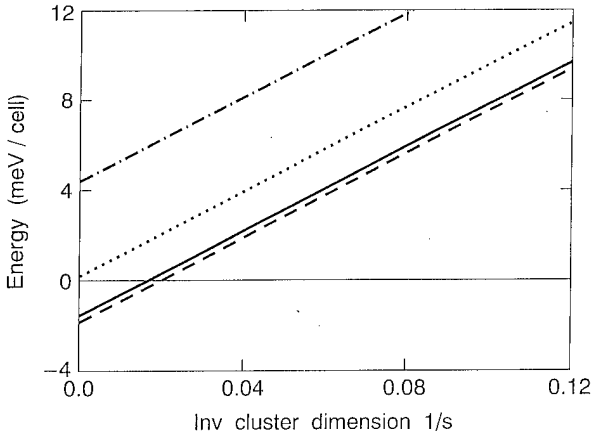


FIG. 1. Island energies in a meV/surface cell, plotted against s^{-1}, where s is the island size at midheight in surface cell units. The energies are of islands on $(Si_{0.5}Ge_{0.5})_nSi(001)$ substrates referred to as energies of the $(n+1)$ layer strained flat surface Dot-dashed line: $n=2$; dotted line: $n=3$; solid line: $n=4$; dashed line: $n=5$.

island energetics and Khor and Das Sarma[10] have shown that this expression, slightly modified, applies to tall (111)-faceted islands. Experimental and theoretical results[6,10,14] have indicated that (111)-faceted islands are the most energetically favorable; our simulations are carried out with these islands, on substrates thick enough (typically more than 60 ML) that the energy has converged to less than 0.1 meV/atom.

In Fig. 1, we plot the results of simulations for $Si_{0.5}Ge_{0.5}$ island clusters on substrates made up of $n=3$, 4, and 5 ML of $Si_{0.5}Ge_{0.5}$ on Si(001). In each case we compare the energy of islands sitting on n MLs of $Si_{0.5}Ge_{0.5}$ with the energy of the corresponding $(n+1)$ layer strained flat surface. As we have done previously,[10] to compare the energies of islands of different sizes, we need to keep coverage fixed; for an island of size $2M$ atoms, the energy to calculate is that of the configuration made up of this island on a substrate of length $L=M$ so that, when the atoms of the island are spread over this substrate, it is covered with exactly one extra monolayer of atoms. Under this condition, the island density is low, tending towards 0 as the size approaches infinity. We see that islanding is first favored when the thickness n is 3–4 ML, as in the case when $x=1$.[10] Within the accuracy of our calculations, there is no difference for the results of cases $n=4$ and 5. (Higher order n's do not change this conclusion.) This, together with the results of Fig. 2, will support the conclusion that, under conditions of low island density, the equilibrium SK thickness t_c for $Si_{1-x}Ge_x$ on Si(001) is independent of x and should remain at about 3–4 ML.

In Fig. 2, we plot the energetics of Ge islands laid down on $Ge_nSi(001)$ substrates for $n=3$, 4, and 5, under the condition of constant coverage $\theta=6$ ML. The energies are all referred to as the energy of the flat coherent $Ge_6Si(001)$ surface; at this coverage island densities are low; once 3D islanding becomes favorable, we see that the lowest energy configuration for a fixed coverage is that of islands on 3

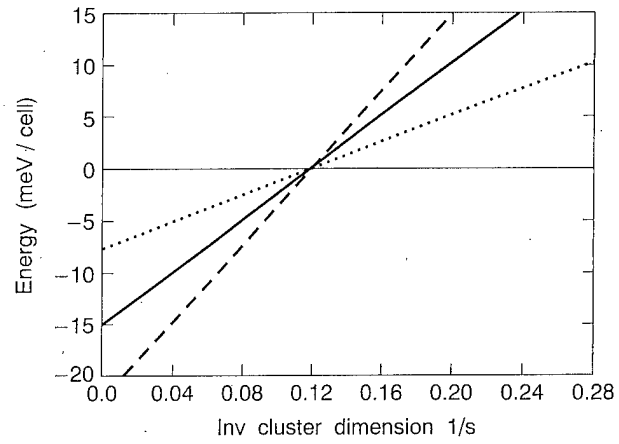


FIG. 2. Islanded surface energies at a constant coverage of $\theta=6$ ML in a meV/cell vs inverse cluster size. The islands are set on $Ge_nSi(001)$ substrates and the energies are referred to as the energy of the flat $Ge_6Si(001)$ surface. Dashed line: $n=3$; solid line: $n=4$; dotted line: $n=5$.

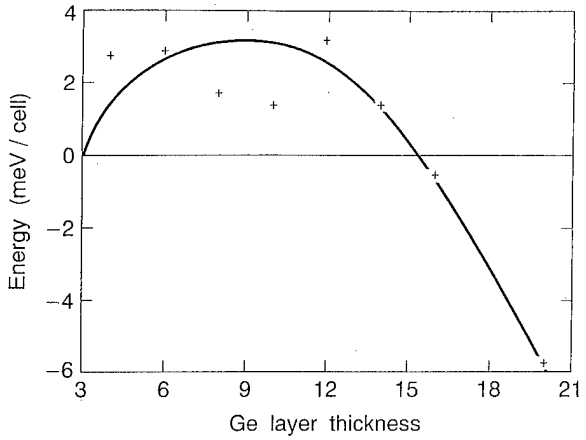


FIG. 3. Islanded surface energies at a constant coverage of $\theta=33$ ML in a meV/cell vs Ge layer thickness n. The energy sign has been reversed so that the positive region is the more stable. The islands are set on $Ge_nSi(001)$ substrates and the energies are referred to as the energy of the island on the $Ge_3Si(001)$ surface. The line is only a guide to the eye.

ML of Ge. The particular average size of islands at equilibrium would depend on energetic and entropic considerations. The results for low density $Si_{1-x}Ge_x$ islands on $(Si_{1-x}Ge_x)_nSi(001)$ substrates would be the same as can be inferred from Fig. 1. SK critical thickness at equilibrium should be about 3 ML and be largely independent of x.

We now look at the high coverage case. Osten *et al.*[11] used coverages of 30–60 ML and their results, together with those of Hansson *et al.*[6] show islands of width $w\sim50$ nm at the base ($\sim$130 surface cell units in the [110] direction) and interisland separation of about the same length. The results of Hansson *et al.* show all islands to have a constant aspect ratio [height (h): width (w)] of 0.5.[6,9] Using these parameters as our guide, we have carried out simulations with coverages $\theta\sim30$–40 ML on (111)-faceted islands with base width $w=70$–110 surface cell units and aspect ratio of 0.5. The results plotted in Fig. 3 are for islands with width $w=78$ and coverage $\theta=33$ ML (s is the width or size of the island at half height). In calculating the energies for Fig. 3, we have ignored the correction due to reconstructions on the (111) side facets of the islands[10] since we are only interested in differences of energy and the effect of the corrections on these would be negligible. We have also chosen the parameters w and θ so that the island is close to the transition point (see Fig. 2) at which island energies first become negative. The reason for this will be discussed later. In Fig. 3 the energies are for islands on $Ge_nSi(001)$ substrates, taken relative to that of an island on 3 ML of Ge (i.e., $n=3$). Figure 2 shows that under low coverage conditions this graph would be a straight line with negative gradient, passing through the point given by $E=0$ and $n=3$. However, Fig. 3 shows that under conditions of high island density the relaxation in the substrate below each island can have a substantial effect. The energies for islands on n MLs of Ge for a range of thicknesses, $3<n\leqslant16$, become more favorable than the energy for $n=3$. This is the likely mechanism for SK critical thickness to exceed 3 ML even under equilibrium conditions. Results remain substantially the same for small changes ($\sim$10% of θ and island size s). The feature of a range of n (>3) values being more favorable than $n=3$ remains. *At fixed coverage*, increasing island size s increases interisland distance; this has the result of decreasing the relative effect of substrate relaxation. Figure 2 shows as well that the energy difference between islands on n and $n+1$ ML would increase. As $1/s$ tends towards zero, at some point given by $s=s_0$, say, this energy would be greater than the energy gained due to substrate relaxation and islands on 3 ML of Ge will then be the most favorable configurations. We are interested in the interval $\Delta s=s_0-s_c$, where s_c is the size at which an island first becomes favorable. Now, we consider increasing island sizes while keeping them apart at a *constant* distance equal to w, their size at the base; the relaxation effect of the substrate will scale with the size of the islands, that is, the range of values of n (>3) for which island energies are more favorable than that for $n=3$ will increase and so will Δs. We can see from Figs. 1 and 5 of previous work[9] that as $x\rightarrow0$, the sizes s_c at which islands first become favorable get larger. From the following equation for the energy E of an island of side s and height h,[9,15]

$$E/V=\frac{2\Gamma}{s}-\frac{2c[1-\exp(-\alpha h/s)]}{\alpha\ln[se^{1.5}/h\cot(\theta)]},\tag{1}$$

where $\Gamma=\gamma_e csc(\theta)-\gamma_s\cot(\theta)$, $\gamma_{e,s}$ are, respectively, energies of the facet and the surface, θ, angle between the surface and facet, α is some constant, and $c\sim x^2$, we see that $s_c\sim c^{-1}\sim x^{-2}$. At low x, this transition point would move closer and closer to the abscissa (see Figs. 1 and 2), that is, s_c would tend towards very large values. It is reasonable to expect that, at this limit, average island sizes (this average is affected by energetic and entropic considerations) would be $\sim s_c$ and that the arguments we have given above for increased SK thickness would then apply. Then, in this limit $x\rightarrow0$, equilibrium SK thickness should scale as s_c and, therefore, as x^{-2}.

III. CONCLUSION

We have found for strained heteroepitaxial growth of semiconductors that, at high coverages, SK critical thickness can exceed 3 ML even under equilibrium growth conditions. The mechanism for this is the great elastic relaxation that goes deep into the substrate just beneath an island; this relaxation makes it favorable for a certain thickness of the material in the substrate to be of the same type as that in the island; this thickness scales with the size of the island. This effect is strongest when average island sizes are those near the transition point at which islanding just becomes favorable and the coverage is large enough for the interisland distance to be of the order of island size.

ACKNOWLEDGMENT

This work is supported by the U.S. Office of Naval Research.

[1]Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Phys. Rev. Lett. **65**, 1020 (1990).
[2]M. Asai, H. Ueba, and C. Tatsuyama, J. Appl. Phys. **59**, 2577 (1958).
[3]D. J. Eaglesham and M. Cerullo, Phys. Rev. Lett. **64**, 1943 (1990).
[4]F. K. Legoues, M. Copel, and R. M. Tromp, Phys. Rev. B **42**, 11 690 (1990).
[5]D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Denbars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993); D. Leonard, K. Pond, and P. M. Petroff, Phys. Rev. B **50**, 11 687 (1994); J. M. Moison, F. Houzay, F. Barthe, and L. Leprince, Appl. Phys. Lett. **64**, 196 (1994).
[6]P. O. Hansson, M. Albrecht, H. P. Strunk, E. Bauser, and J. H. Werner, Thin Solid Films **216**, 199 (1992).
[7]J. C. Bean, T. T. Sheng, L. C. Feldman, A. T. Fiory, and R. T. Lynch, Appl. Phys. Lett. **44**, 102 (1984).
[8]C. W. Snyder, J. F. Mansfield, and B. G. Orr, Phys. Rev. B **46**, 9551 (1992).
[9]K. E. Khor and S. Das Sarma, J. Vac. Sci. Technol. B **15**, 1051 (1997).
[10]K. E. Khor and S. Das Sarma, Phys. Rev. B **49**, 13 657 (1994); J. Tersoff, *ibid.* **45**, 8833 (1992).
[11]H. J. Osten, H. P. Zeindl, and E. Bugiel, J. Cryst. Growth **143**, 195 (1994).
[12]A. H. Harker, K. Pinardi, S. C. Jain, A. Atkinson, and R. Bullough, Philos. Mag. A **71**, 871 (1995).
[13]S. Christiansen, M. Albrecht, S. Strunk, and H. J. Maier, Appl. Phys. Lett. **64**, 3617 (1994).
[14]C. Ratsch and A. Zangwill, Surf. Sci. **293**, 123 (1993).
[15]J. Tersoff and R. M. Tromp, Phys. Rev. Lett. **70**, 2782 (1993).

Surface roughness, strain, and alloy segregation in lattice-matched heteroepitaxy

Catherine Priester[a)]
IEMN, Département ISEN, CNRS-UMR 9929, F-59652 Villeneuve D' Ascq Cedex, France

Genevieve Grenet
Ecole Centrale de Lyon, LEAME, CNRS-UMR 5512, F-69131 Ecully Cedex, France

(Received 21 January 1998; accepted 28 May 1998)

Spinodal decomposition of ternary semiconductor alloys during lattice-matched heteroepitaxy is considered here. It has been previously demonstrated that a perfectly flat surface (with no step) would forbid alloy demixing. The case of a rough surface is the purpose of this article. How the possibility of a better strain relaxation introduced by the surface roughness can favor alloy demixing is analyzed first. The present results are exemplified by the AlInAs lattice matched to the InP case. Second, a step-by-step model is proposed to simulate the growth process on a rough surface. This model leads to a description of the strain and alloy demixing during this growth. This study clearly shows how and why the atoms corresponding to binary materials with lower surface tension naturally tend to segregate towards bumped areas. © *1998 American Vacuum Society.*
[S0734-211X(98)12504-0]

I. INTRODUCTION

It is now well known[1–5] that most of the III–V alloys such as AlInAs or GaInP may show strong phase separation behavior when grown lattice matched to the substrate by molecular beam epitaxy (MBE). We have recently shown[6] that, from a theoretical point of view and contrary to what happens in bulk alloys,[7,8] when the alloy is deposited on a substrate lattice matched to the alloy lattice constant, such a spinodal decomposition is forbidden if the surface of the deposited film is perfectly flat: spinodal decomposition indeed would create strained areas which appear not to be favorable with regard to a random alloy where the mean strain is zero everywhere in the system, even if one takes into account the mixing enthalpy. Recently Guyer and Voorhees[9] proposed a decomposition mechanism which is active during epitaxial growth of strained alloy layers even when spinodal decomposition is no longer effective. This decomposition is associated with surface roughening. This analysis, which uses a continuous approach, concludes that alloy decomposition may stabilize surfaces under tension whereas it always destabilizes surfaces under compression. In a more recent work[10] Tersoff considered the case of vicinal surfaces and proposed a mechanism which relates spinodal decomposition to the formation of step bunches. This latter phenomenon can lead to the spontaneous formation of superlattices. In the present work, we have chosen to consider the case of a nominal surface, but one with some roughness. A recent thermodynamic treatment of a stressed alloy with a free surface[11] studied the relationship between morphological and compositional instabilities. Here, we do not consider morphological instabilities, but have chosen to keep within the framework of roughness limited to a few monolayers (ML) observed experimentally.[6] In Sec. II we demonstrate, by making use of a simple model, how the presence of wires on the surface favors alloy decomposition, leading to a lateral alternation of Al-rich and In-rich wires. In Sec. II we completely ignore how decomposition takes place, and only compare the equilibrium total energy of segregated and random systems. A step-by-step modeling of the alloy's heteroepitaxial growth is proposed in Sec. III: this approach is basically similar to the one which has recently successfully described strain assisted atomic diffusion in inhomogeneously strained systems such as self-assembled three-dimensional (3D) islands.[12] In the present work, starting from a slightly rough surface, we analyze how the atoms segregate as one deposits successive monolayers.

II. ALLOWED SPINODAL DECOMPOSITION IN SYSTEMS WHICH PRESENT A ROUGH SURFACE ALONG ONE SINGLE DIRECTION: A VERY SIMPLE DESCRIPTION

As briefly recalled in Sec. I, we have previously shown[6] that, when one considers a composition fluctuation around the lattice-matched composition $x=0.526$ given by Vegard's law for an $Al_{1-x}In_xAs$ film on an InP substrate, the alloy film is either tensile or compressively strained, and that this "extrinsic" strain energy (due to the mismatch with the substrate) contravenes the "intrinsic" strain energy due to the mismatch between the two end point binaries within the alloy. The calculation of this balance energy for a perfectly flat film has proved that the extrinsic strain energy gets the better of the intrinsic energy, thus prohibiting any spinodal decomposition in such systems. However, as the growth front is actually rough, a key point is to check whether the roughness can be efficient enough in relaxing the extrinsic strain energy. If yes, the balance energy will be switched, and, as far as additional surface energy does not prohibit such a roughness, spinodal decomposition would be allowed. In order to answer this we have calculated the total energy

a)Electronic mail: priester@isen.fr

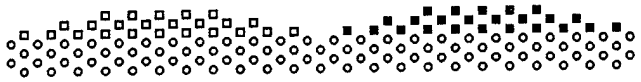


FIG. 1. Schematic cross-sectional view of a rough surface made of [113]-faceted truncated ridges. Only the fcc cation sublattice is represented. Open circles correspond to $Al_{0.474}In_{0.526}$ virtual cations, open squares to $Al_{0.474+x}In_{0.526-x}$ virtual cations, and closed squares to $Al_{0.474-x}In_{0.526+x}$ virtual cations.

(elastic+mixing+surface) for several series of systems with rough surfaces which are shown schematically in Fig. 1. The roughness is modeled here by considering a surface which presents an undulation in the direction perpendicular to the element-V dimers (the actual surface is assumed to be element V stabilized): the ridges are 3 ML high and [113] faceted. We have varied the widths of the ridges and the spaces between the ridges, and for every configuration we have calculated the total energy for alternation of In-rich and Al-rich surface ridges, varying the amount of demixing from zero to total demixing. In this simple model the elastic energy is calculated within the valence force field framework[13] and by using Keating's formulation.[14–16] The alloy atoms are here considered as virtual atoms. The excess enthalpy of the mixing density in each alloy area is approximated by $U(x) = x(1-x)\Omega$ where Ω, the interaction parameter, is 3.6 kcal/mol.[16] The surface energy is estimated by counting the additional dangling bonds due to the surface roughness (the additional energy of one dangling bond for an In–As terminated surface is estimated to 0.3 eV).[17]

Some of these results are reported in Fig. 2, which displays the variations of the reduced energy versus the demixing ratio (=0 for a random alloy, =1 for total demixing, viz, one wire being purely InAs and the other one being $In_{0.058}Al_{0.942}As$ since the total amount of deposited In corresponds to the lattice-matched composition). Each curve corresponds to a given size of wire, and the zero energy corresponds to the total energy of the equivalent flat random system (please note that, when one passes from one curve to the other, the size of the system varies).

From observation of Fig. 2, one can deduce the following points:

(1) for any rough surface the demixed system is more stable that the random one;
(2) one can define a critical roughness (typically 150 atoms/ridge which corresponds to 120 Å, 29 atoms, on the basis of each ridge), from which one can find demixed systems that are more stable that the flat random film;
(3) for a given roughness, one can define a critical demixing which is the minimal demixing for the demixed rough system to be more stable that the flat random film.

These qualitative results appear to not depend on the height of the roughness, nor on the lateral facets orientations. Only the quantitative values of the critical roughness and critical demixing are sensitive to these "design parameters."

This study has thus demonstrated that spinodal decomposition is favored by the surface roughness, and has confirmed the idea that, for a rough enough surface, this roughness

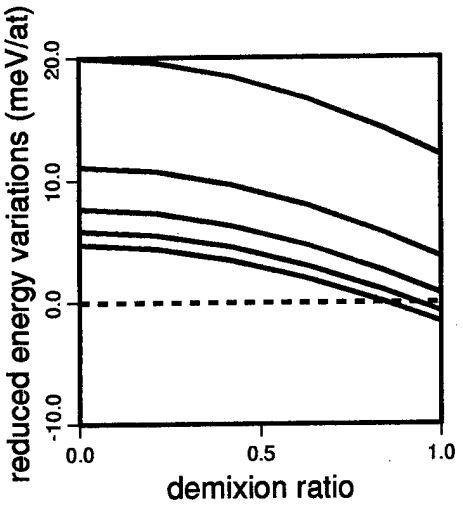


FIG. 2. Variations of the reduced energy (in meV/atoms in the ridges) vs the demixing ratio (from one ridge to its neighbor). Each curve corresponds to a given width of the ridges: from the upper to the lower curve the base of the ridge contains, respectively, 11, 17, 23, 29, and 35 atoms.

lowers the elastic energy in demixed areas strongly enough in order to balance the additional surface energy. However this study has only considered rather idealistic test systems (alternation of homogeneous demixed surface ridges). Since each surface ridge is considered homogeneous, the well known segregation of In towards the surface for such alloys cannot be described in the present model. As will be demonstrated in Sec. III, such In segregation in rough systems can strongly modify the strain distribution all over the system, and consequently modify the alloy decomposition. A more realistic description then requires one to take into account the possibility of a gradual composition modulation, and to study how the atoms can exchange during the growth process. This is the purpose of Sec. III.

III. STEP-BY-STEP DESCRIPTION OF THE GROWTH OVER A BIDIRECTIONALLY ROUGH SURFACE

As the existence of a roughness-assisted alloy segregation has been established, let us now no longer discuss virtual alloys (i.e., described using Vegard's law) and turn to random alloy systems [i.e., assuming a particular random cation distribution on their face-centered-cubic (fcc) sublattice sites].

In order to determine towards which alloy decomposition will energetically tend the system, we have used a procedure which has been successfully applied recently for describing the strain assisted atomic diffusion in self-assembled 3D GaInAs islands on a GaAs substrate.[12] The principle itself is the following.

At each step of the calculation, we define the class of cations which are "allowed to exchange" (see below) which will be referred in the following as the "moving class."

First, we calculate the stress field all over the system considered by replacing the cations of the moving class with virtual cations. Then we calculate, for any of these cation sites, the local (on two shells) dynamical matrices when this

cation is an In or Al. Dynamical matrices are derived from Keating's formulation of the elastic energy.[14,15] From these we deduce an estimation of the optimal cation distribution which minimizes the elastic+surface energy (for a given surface morphology).

Then we repeat the process (stress field+optimal cation distribution calculations) up to convergence (when no more atoms need to be exchanged in order to minimize the energy).

The key point here is to determine which atoms are "allowed to exchange." In order to take into account the fact that bulk diffusion is negligible at usual growth temperatures, we have decided to consider every "nonsurface" cation as frozen. "Nonsurface" cations mean cations which do not belong to the surface layer or the layer just underlying it. As we now consider completely or partially random alloys, the mixing enthalpy is directly included in our Keating description on equal footing as strain energy. The difference of surface tension between the two binaries, InAs and AlAs, which can be associated to the difference in dangling bond energy, is included in the calculation as an input parameter. Finally, note that the system we deal with shows a dimerized As-terminated surface (due to growth conditions). This As dimerization implies that the two surface directions, i.e., the [110] direction and the $[1\bar{1}0]$ direction, are not equivalent, even for a perfectly flat surface, exempt from steps.

We now model the initial roughness by means of a few ML high rectangular platelets with rather flat facets (from [112] to [114] orientation, whereas the growth axis is along the [001] direction). We have considered different platelet lengths and widths (which will be referred to as $n \times m$, n along the [110] direction and m along the $[1\bar{1}0]$ direction), but, for numerical requirements, the sample that contains a huge number of atoms is assumed to be periodic in the two directions perpendicular to the growth axis. One period contains two platelets along the [110] direction and one platelet along the $[1\bar{1}0]$ direction, which is the minimum number of platelets which allow the migration of cations from one platelet to another one. The presence of these platelets affects the local strain distribution all over the system. In order to model the growth process, we have chosen to capture a sequence of current images during this growth simulation. Between two consecutive images, 1 ML has been deposited on the surface, but, as III–V surfaces needs to be dimerized for surface tension requirements, we consider that the deposited monolayer results from covering every anion surface atom with entities made of an anion dimer and its four underlying cations. This means that when one passes from one step of the growth sequence to the next, the platelet that defines the roughness enlarges by one atom along the [110] direction and by two atoms along the $[1\bar{1}0]$ dimer direction (as can be seen in Fig. 4). This way the criterion, "platforms must keep able to be dimerized" is verified. Another consequence of this growth process modeling is that after deposition of several monolayers (the number depends on the distance between platelets at the starting point of the sequence), the roughness disappears and one recovers a flat surface on which new platelets will nucleate.

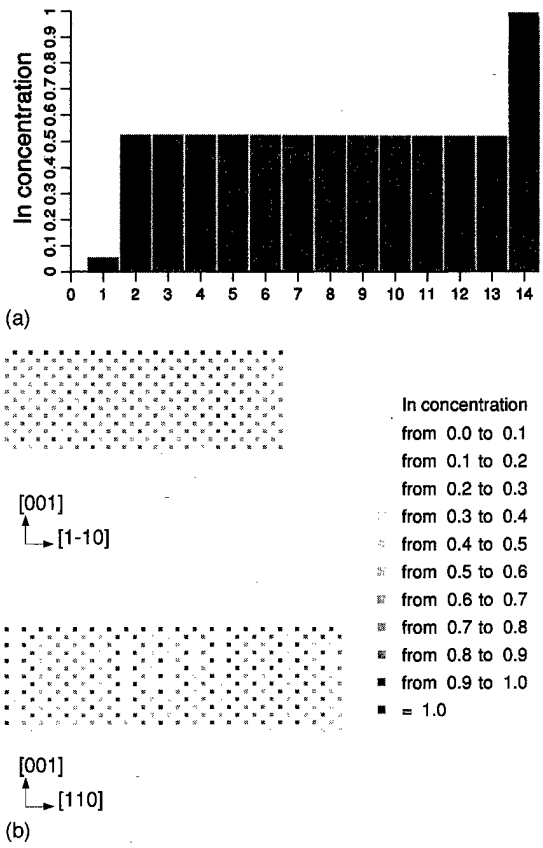


FIG. 3. System made of 14 $Al_{0.474}In_{0.526}As$ flat monolayers deposited on InP: (a) In concentration profile from the substrate (left side) to the surface (right side); (b) two "vertical" cross-sectional views: perpendicular to the [110] (bottom) and [110] (top) directions. In-rich atomic rows are represented by darker symbols and Al-rich rows by brighter symbols (see the inset).

Applying the procedure described above to a system with a perfectly flat surface leads to the variations of indium concentration given in Fig. 3. Figure 3(a), which displays the concentration profile as one goes from the substrate to the surface, illustrates nothing but the well known surface segregation.[18,19] Obviously this ideal abrupt profile would be obtained only for infinite diffusion constants or, equivalently, for very low growth rates. In realistic cases the resulting profile would be an intermediate situation between the purely homogeneous film and the "fully segregated" profile given in Fig. 3(a), but, for usual growth conditions, much closer to the latter[19] (this being just a bit smoothed). The question of kinetic limitations[20] goes beyond the scope of this article and will be briefly discussed in Sec. IV. The cross-sectional view is given in Fig. 3(b). The gray level indicates the mean concentration for each atomic row, with the darkest corresponding to pure In and the brightest corresponding to pure Al (see the inset). The cross section along the [110] axis (which displays the concentration along rows of the dimer direction) shows that the atoms are randomly distributed in this direction. The cross section along the $[1\bar{1}0]$ shows a light tendency towards structuring along the dimer direction, but only to a very short range, and with no periodicity.

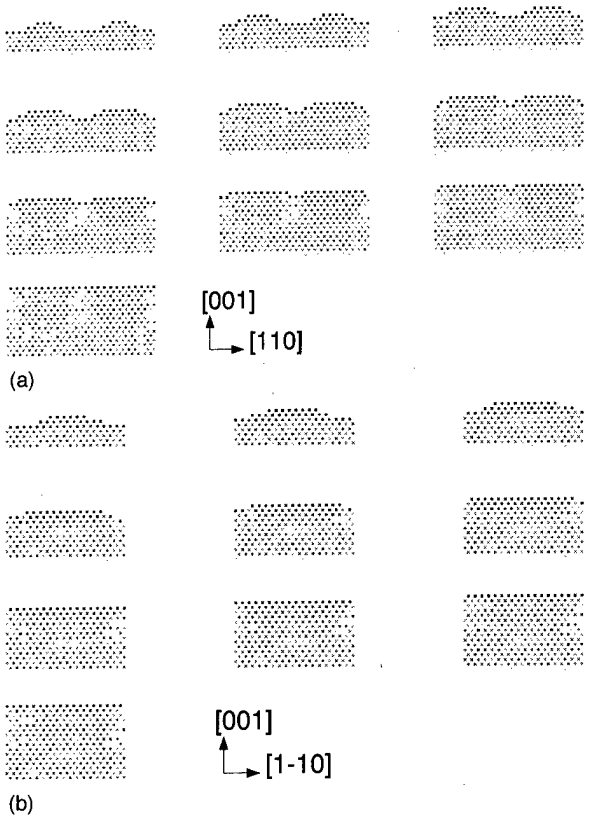


FIG. 4. Successive cross-sectional views of growth on a rough surface (0–9 ML deposited).

Let us now turn to systems with a rough surface. In Fig. 4 we show successive cross-sectional views of a system with an initial roughness corresponding to 2 ML high rectangular platelets containing 6×10 cations at their bases and 3×6 atoms in the other cation planes located in an 12×18 area. In order to allow cation transfer from one platelet to its neighbor, we have considered a 24×18 period, which contains two platelets (located at the center of each half period). For such a surface morphology, one recovers a flat surface after 9 ML has been deposited. In order to complete a global overview of the system, we show in Fig. 5 the corresponding successive "top views" where each square indicates the mean concentration for the part of the atomic row along the [001] axis located above the last entire [001] plane of the initial surface. Figures 4 and 5 clearly show In enrichment in the areas where the platelets were located. If one deposits more and more atomic layers, keeping the surface perfectly flat, one observes that the segregation slowly becomes smooth and tends to vanish. From this observation we can deduce that surface roughness is necessary to assist the alloy spinodal decomposition; in other words, even if one starts with a partially demixed alloy film, it is impossible to keep a trace of this initial demixing without the help of steps. Moreover, it is interesting to note that the surface never actually remains perfectly flat, and that new platelets or holes will nucleate, during the growth, not necessarily located exactly vertically above the previous ones. Studying this process would be an interesting extension of the present work, and

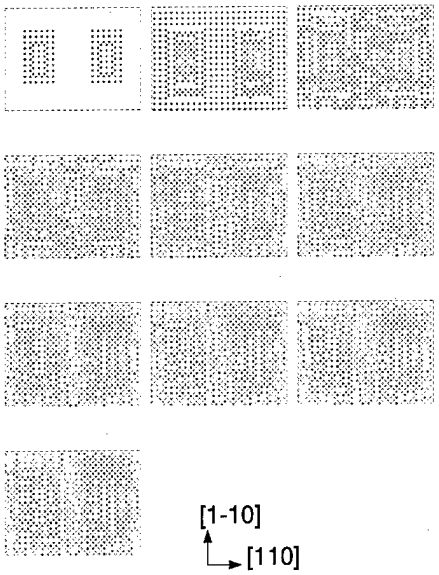


FIG. 5. Successive top views of a growth on a rough surface (0–9 ML deposited).

could, perhaps, provide an explanation of the staggered arrangement of the experimentally observed clusters.[6]

We have checked numerically that the segregation is mainly ruled by surface tension, as if one would consider a ternary alloy whose binary compounds would only differ in lattice parameter, a rough growth front would also lead to the decomposition of the alloy, but with no correlation between the location of initial platelets and Al-rich or In-rich areas.

Another interesting point we can deduce from our study is that, contrary to that was suggested in Sec. II, alloy decomposition does not result from exchanges from one bump to the other but, rather, from exchanges from "flat areas" to bumps. Actually, the surfactant behavior of In atoms forbids such bump-to-bump transfers.

We have also considered systems presenting a less natural starting roughness: this latter corresponds to platelets very close to one another, which can be viewed as crosshached prepatterned surfaces. In this case alloy decomposition is also strongly enhanced (compared to flat systems) but in a quite different way (e.g., in the case of a small enough square periodic pattern, one gets a clear composition modulation along the [100] direction, whereas in the case of nonsquare or large patterns the effect is less pronounced). This is due to the fact that, by the way we model the growth, the surface becomes flat again very quickly. For all the systems we have considered here, the roughness of the growth front appears as the *sine qua non* condition for alloy decomposition: it initiates and strongly influences alloy decomposition.

IV. CONCLUSION

In the present study, based on energetic considerations, as described by a Keating type model, we have clearly demonstrated the surface roughness to be a possible origin of alloy decomposition when growing an alloy lattice matched on a (nonvicinal) substrate. Obviously the "optimal cation distributions" we have calculated for several systems may not be reached due to kinetic limitations, and the actual resulting

partially demixed alloy will correspond to an intermediate situation between this optimal configuration and a completely random system. The influence of the finite growth rate could be introduced in our model as a phenomenological parameter (which would be deduced from experiment). Preliminary calculations have already shown that one could define a critical growth rate that wears away the alloy decomposition. Introducing the effect of growth temperature would require further considerations.

Here we have restricted our study to roughness corresponding to a bumpy surface; the case of a surface with hollows would also be of great interest, and will be the subject of a forthcoming study. The present study also warrants being extended to the case of low mismatched alloy growth. This would be the first step toward a more general understanding of the difference that can be observed when one grows tensile or compressively strained materials with different surface reconstructions.[21] Of course the key role of the surface roughness that we have demonstrated here for a given system (AlInAs lattice matched to InP) is not limited to this single system, and has to be considered as a possible origin of alloy segregation in other grown alloys.[4,5]

ACKNOWLEDGMENT

This work was partly supported by France Telecom (Contract No. 95 1B002).

[1]F. Peiro, A. Cornet, J. R. Morante, A. Georgakilas, K. Zekantes, and G. Halkias, Appl. Surf. Sci. **65**, 447 (1993).
[2]O. Ueda, T. Fujii, Y. Nakada, and I. Umebu, J. Cryst. Growth **95**, 38 (1989).
[3]A. Gomyo, K. Makita, I. Hino, and T. Susuki, Phys. Rev. Lett. **72**, 673 (1994).
[4]T. Okada, G. C. Weatherly, and D. W. McComb, J. Appl. Phys. **81**, 2185 (1997).
[5]R. R. Lapierre, T. Okada, B. J. Robinson, D. A. Thompson, and G. C. Weatherly, J. Cryst. Growth **155**, 1 (1995); **158**, 6 (1996).
[6]G. Grenet *et al.*, Appl. Surf. Sci. **123/124**, 324 (1998).
[7]J. W. Cahn, Acta Metall. **92**, 795 (1961).
[8]S.-H. Wei, L. G. Ferreira, and A. Zunger, Phys. Rev. B **41**, 8240 (1990).
[9]J. E. Guyer and P. W. Voorhees, Phys. Rev. Lett. **74**, 4031 (1995).
[10]J. Tersoff, Phys. Rev. B **56**, R4394 (1997).
[11]F. Glas, Phys. Rev. B **55**, 11 277 (1997).
[12]C. Priester and M. Lannoo, Appl. Surf. Sci. (to be published).
[13]M. J. P. Musgrave and J. A. Pople, Proc. R. Soc. London, Ser. A **268**, 474 (1962).
[14]P. N. Keating, Phys. Rev. B **30**, 6217 (1984).
[15]R. M. Martin, Phys. Rev. B **1**, 4005 (1970).
[16]J. L. Martins and A. Zunger, Phys. Rev. B **30**, 6217 (1984).
[17]C. Priester and M. Lannoo, Appl. Surf. Sci. **123/124**, 658 (1998).
[18]J. M. Moison, C. Guille, F. Houzay, F. Barthe, and M. Van Rompay, Phys. Rev. B **40**, 6149 (1989).
[19]O. Dehaese, X. Wallart, and F. Mollot, Appl. Phys. Lett. **66**, 52 (1995).
[20]J. E. Oh, P. K. Bhattacharya, Y. C. Chen, O. Aina, and M. Mattingly, J. Electron. Mater. **19**, 435 (1990).
[21]M. Gendry, G. Grenet, Y. Robach, P. Krapf, L. Porte, and G. Hollinger, Phys. Rev. B **56**, 9271 (1997).

Chemical bonding features for faultily stacked interfaces of GaAs{111}

Jun Nakamura[a] and Tetsuya Mishima
Kagami Memorial Laboratory for Materials Science and Technology, Waseda University, Shinjuku-ku, Tokyo 169-0051, Japan

Moto-hisa Masui, Mineo Sawayanagi, Sung-Pyo Cho, Masayasu Nishizawa, and Toyoaki Eguchi
Department of Materials Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-0072, Japan

Toshiaki Osaka
Department of Materials Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-0072, Japan and Kagami Memorial Laboratory for Materials Science and Technology, Waseda University, Shinjuku-ku, Tokyo 169-0051, Japan

(Received 21 January 1998; accepted 28 May 1998)

The electronic states for normally stacked and faultily stacked layers on the GaAs{111} *A*, *B* surfaces are calculated by use of the discrete variational *X* α cluster method and the plane wave nonlocal pseudopotential method. The results show that chemical bondings between atoms are not as ionic in the faultily stacked layer of (111) *B* as they are in the (111) *A* case, and that on the (111) *A* surface more attractive Coulomb interaction energy is gained in the faulty stacking layer than in the normal stacking one. These results explain well the more frequent emergence of in-plane faults in the (111) *A* surface, which is well known in GaAs{111} *A*, *B* growth experiments. The total energy calculations also provide quantitative interpretation of such growth features. © *1998 American Vacuum Society.* [S0734-211X(98)12604-5]

I. INTRODUCTION

In general, single crystal ingots of III–V compound semiconductors with a zinc-blende (ZB) structure have been grown on their {111}-oriented seed crystals by the liquid-encapsulated Czochralski (LEC) technique. A structural perfection of the {111} LEC-grown crystal is known to depend on the polarity of the seed crystal:[1–5] In the {111} end surface terminated with group V atoms, (111) *B*, the growth front is mostly composed of a fault-free layer with a stacking of *AaBbCc* inherent to the ZB-type structure. The letters *Aa*, *Bb*, or *Cc* represent a pair of group III and group V atoms. The other end surface that consists of group III atoms, (111) *A*, tends to grow with considerable amounts of in-plane faults such as twin and stacking faults, for example, including a locally stacked layer of *AaBbAa* based on a wurtzite (WZ) structure (Fig. 1). Such a polar surface dependence has been observed also in the III–V compound films prepared by molecular-beam epitaxy and metalorganic chemical vapor deposition.[6]

Most III–V compounds representative of GaAs have the ZB structure, in which the atoms are tetrahedrally and covalently bonded. However, of the III–V compounds, GaN, or InN is an example that has the other tetrahedrally bonded structure of WZ type, which is more ionic.[7,8] This gives us a key to the question of what causes the polar surface dependence mentioned at the beginning of this article: If the ionicity for bonds between the atoms at the {111} *A*, *B* surface layers becomes locally larger than the inherent ionicity in the bulk crystal, they easily grow in the WZ-type stacking of *AaBbAa* corresponding to an in-plane fault in ZB.

[a]Corresponding author; electronic mail: junj@mn.waseda.ac.jp

One of the purposes of this article is to examine the difference in ionicity between ZB-type and WZ-type structures using first-principles calculations. The other is to explore the origin of the frequent induction of in-plane faults in the LEC-grown film on the GaAs(111) *A* substrate on the basis of the difference in the "inherent" ionicity between the {111} *A*, *B* surfaces.

II. CALCULATIONS AND MODELS

A. Total energy calculations

The total energy calculations were performed by use of the first-principles local-density functional (LDF) method.[9,10] We adopted a momentum-space formalism[11] and *ab initio* norm-conserving pseudopotentials[12–14] for the total energy calculations. The wave functions were expanded in a plane-wave basis set. The total energy, E_{total}, is given by

$$E_{\text{total}}=E_{\text{Ewald}}+E_{\text{Hartree}}+E_{XC}+E_{\text{potential}}+E_{\text{kin}}, \tag{1}$$

where the terms represent, respectively, the Ewald (core–core Coulomb) energy, the Hartree (electron–electron) energy, the exchange and correlation energy, the potential (electron–core) energy, and the electronic kinetic energy.[15] The Wigner form[16] was used for the calculation of the exchange-correlation energy.

In this study, the following two types of slab geometries were used for the simple calculation. Slabs in one geometry were prepared for the ZB and WZ structures, which have supercells consisting of two Ga (group III) and two As (group V) atoms. In the present calculations the unit-cell vectors for the ZB and the WZ structures were as follows:

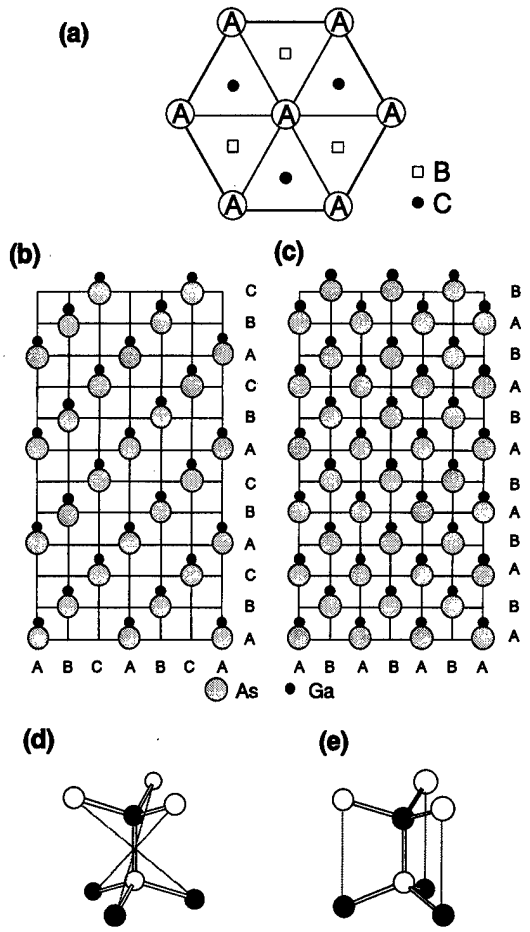


FIG. 1. (a) Projection of the most-closely packed layer. Stacking sequences of the (b) ZB-type and the (c) WZ-type GaAs. ZB and WZ have a stacking of $AaBbCcAaBbCc\cdots$ in the [111] direction and of $AaBbAaBb\cdots$ in the [0001] direction, which are in the (d) staggered and (e) eclipsed conformations, respectively. The letters Aa, Bb, or Cc represent a pair of Ga and As atoms.

$$\begin{cases} \mathbf{a}_{ZB}=(\frac{1}{2},0,\frac{1}{2})a_0, \\ \mathbf{b}_{ZB}=(0,\frac{1}{2},\frac{1}{2})a_0, \\ \mathbf{c}_{ZB}=(1,1,0)a_0, \end{cases} \tag{2}$$

$$\begin{cases} \mathbf{a}_{WZ}=(\frac{1}{2},\frac{1}{2},0)a_0, \\ \mathbf{b}_{WZ}=(0,\frac{1}{2},\frac{1}{2})a_0, \\ \mathbf{c}_{WZ}=(\frac{2}{3},-\frac{2}{3},\frac{2}{3})a_0. \end{cases} \tag{3}$$

Here, a_0 is the lattice constant for the ZB unit cell. As can be seen in Eqs. (2) and (3), the axial ratio for the WZ unit cell is assumed to be ideal, that is, $c/a=\sqrt{8/3}$ (where a and c are the lattice constants for the WZ unit cell). One reason is in the way that the supercells for both the structures have the same volume in order to comprise directly results of the total energy calculations for both systems. The other is that the use of the actual c/a causes only a small relaxation energy difference below 1 meV/atom.[17] For slabs in the other geometry, we employed the slab models shown in Figs. 2(a)–2(e), in which 8, 10, and 12 atomic layers of GaAs are combined with vacuum regions which have corresponding atomic layers of 12, 10, 8, respectively. The front side of these slabs is represented by the {111} A, $B-1\times1$ ideal surfaces, while the back side is terminated with a fictitious H atom which has 0.75 (1.25) nuclear and 0.75 (1.25) electronic charges for the (111) A [(111) B] surface models so as to eliminate artificial dangling bonds and to decouple both the surfaces.[18] In optimizing the geometries of the slab models, the atoms in the back side H layer and its adjacent Ga–As{111} bilayer were fixed.

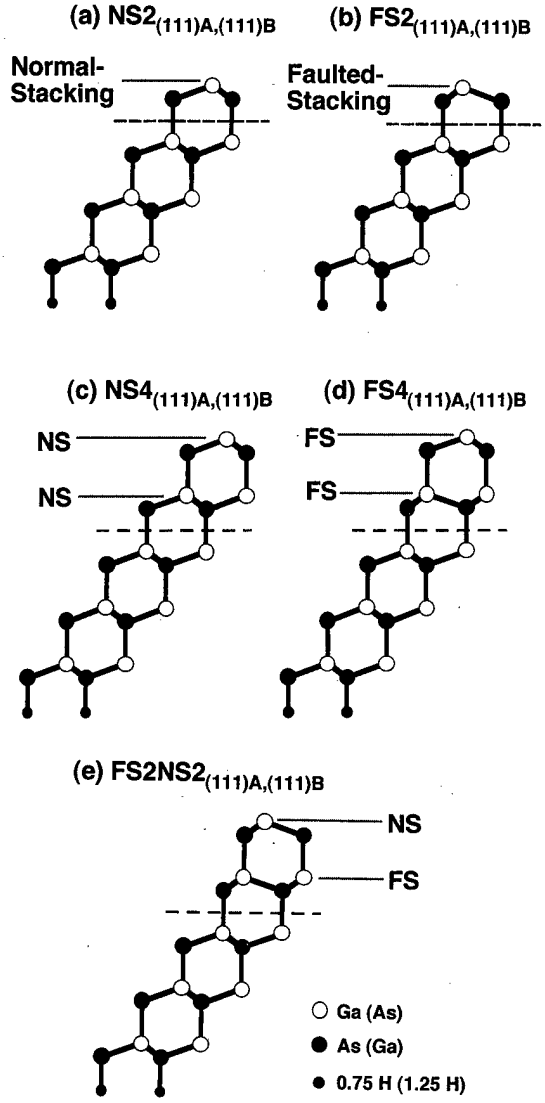


FIG. 2. Models employed for the total energy calculations. (a) $NS2_{(111)A}$ and $NS2_{(111)B}$ models for the NS layer (2 layers=1 bilayer) on the GaAs(111) A and GaAs(111) B surfaces, respectively. (b) $FS2_{(111)A}$ and $FS2_{(111)B}$ models for two faulted stacking layers. (c) $NS4_{(111)A}$ and $NS4_{(111)B}$ models for four normal stacking layers. (d) $FS4_{(111)A}$ and $FS4_{(111)B}$ models for four normal stacking layers. (e) $FS2NS2_{(111)A}$ and $FS2NS2_{(111)B}$ models for the NS layer (2 layers=1 bilayer) on the surfaces of the $FS2_{(111)A}$ and $FS2_{(111)B}$ models, respectively.

To save computational costs, we performed the total energy calculations in the following two steps: In the first step, we employed Kleinman–Bylander (KB) type separable pseudopotentials.[14] Here, both electronic and ionic degrees of freedom were optimized using the conjugate gradient

method. By using this, we can reduce the computational effort to calculate the potential matrix elements. However, since in using the KB type potentials a spurious bound state (i.e., ghost state) occasionally emerges,[19,20] we carefully chose the value of the kinetic energy cut off, i.e., 4.41–7.29 Ry, to remove it. In the second step, further optimization of the atomic structures with higher cutoff energy (up to 7.29–12.25 Ry) was carried out by using the Bachelet–Hamann–Schlüter type nonseparable pseudopotentials[13] until forces acting on the atoms are below $\sim 1\times 10^{-3}$ Ry/a.u.[21]

B. Cluster calculations

Numerical calculations were performed using the self-consistent charge discrete variational (SCC-DV)-$X\alpha$-cluster method, the details of which are reported elsewhere.[22,23] In this method, the Hartree–Fock–Slater (HFS) equation for a cluster is self-consistently solved by use of a localized exchange potential ($X\alpha$ potential). The adjustable (exchange-correlation) parameter α was taken to be 0.7.[23] Numerical $1s-4p$ atomic orbitals of Ga and As, which were obtained as solutions of the atomic HFS equations, were utilized as basis sets. We employed the Mulliken population analysis[24] to estimate the effective atomic charges from the orbital population.

We prepared the following four cluster models for the growing surfaces of GaAs{111} A, B, (1) and (2) are normally stacked (NS) surface models, NS-(111) A and NS-(111) B, and (3) and (4) are faultily stacked (FS) surface models, FS-(111) A and FS-(111) B, respectively. Schematics of these clusters are shown in Fig. 3. The {111} A, B surfaces are represented as embedded models where each cluster is set in the Madelung-type electrostatic exterior potential, V_{out}.[25] The effective Hamiltonian for the cluster is written as

$$H_{\text{eff}}=H_{\text{in}}+V_{\text{out}}, \tag{4}$$

where H_{in} is the Hamiltonian for the cluster. The V_{out} can be calculated as the lattice sum,

$$V_{\text{out}}(\mathbf{r})=\sum{}' \ \phi_n(\mathbf{r}), \tag{5}$$

where the prime means the exclusion of the sites inside the cluster from the summation. The contribution of the nth ion

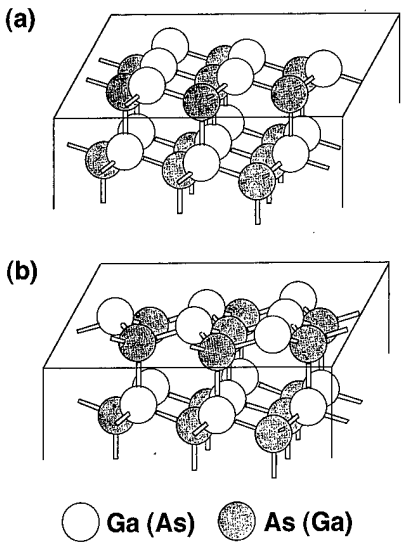


FIG. 3. Models employed for the cluster calculations. (a) NS surface models on the GaAs(111) A and (111) B surfaces, respectively; NS-(111) A, NS-(111) B. (b) FS surface models, FS-(111) A and FS-(111) B.

to the lattice sum, $\phi_n(\mathbf{r})$, is assessed by a model of the uniformly charged sphere with a finite radius r_{n0}, and is expressed as

$$\phi_n(\mathbf{r})=\begin{cases} Q_n/|\mathbf{r}-\mathbf{R}_n| & (|\mathbf{r}-\mathbf{R}_n|>r_{n0}), \\ Q_n(3r_{n0}^2-|\mathbf{r}-\mathbf{R}_n|^2)/2r_{n0}^3 & (|\mathbf{r}-\mathbf{R}_n|\leq r_{n0}), \end{cases} \tag{6}$$

where Q_n is the net charge of ions and $|\mathbf{r}-\mathbf{R}_n|$ is the distance from the center of the nth ion in position $\mathbf{R}_n$. The values of r_{n0} were chosen as the ionic radii for Ga and As atoms, 1.1716 and 1.3039 Å (Goldschmidt), respectively.[26]

III. RESULTS AND DISCUSSION

A. ZB(3C)–WZ(2H) polymorphism

As mentioned in Sec. I, we are very interested as to why, in the LEC-growth of GaAs{111} A, B crystals with ZB structures the (111) A surface favors growing with an in-plane fault of WZ-type stacking, $AaBbAa$. In order to explore the origin of the ZB–WZ polymorphism in GaAs, we have to understand significant differences in the electronic states of the ZB and the WZ structures. As the first step for

TABLE I. Energy differences between the ZB and WZ structures (meV/2atoms).

	GaAs $\Delta\epsilon_i$(ZB−WZ)	Ge $\Delta\epsilon_i(3C-2H)$	InSb $\Delta\epsilon_i$(ZB−WZ)	Sn $\Delta\epsilon_i(3C-2H)$
Ewald	−302.9	−321.9	−264.3	−280.7
Hartree	−310.8	−293.8	−291.1	−212.0
Exchange and correlation	+17.6	+4.7	+18.9	−22.2
Potential	+620.2	+575.1	+549.0	+440.1
Kinetic	−50.5	−3.2	−34.8	+63.6
Total	−26.5	−39.0	−22.4	−11.2

TABLE II. Energy differences between the models for GaAs(111) *A* (eV/slab).

	$\Delta\epsilon_i(\mathrm{NS2-FS2})_{(111)\,A}$	$\Delta\epsilon_i(\mathrm{NS4-FS4})_{(111)\,A}$	$\Delta\epsilon_i(\mathrm{NS4-FS2NS2})_{(111)\,A}$
Ewald	−90.1545	+202.4785	+154.5866
Hartree	−89.7156	+200.7304	+153.3762
Exchange and correlation	+0.3804	−0.7598	−0.4771
Potential	+180.1735	−404.3700	−308.6115
Kinetic	−0.7317	+1.8311	+1.0985
Total	−0.0479	−0.0898	−0.0273

this, we test the difference of the total energy between the ZB- and WZ-GaAs models, as shown in Table I, which is calculated as

$$\Delta\epsilon_{\mathrm{total}}(\mathrm{ZB-WZ}) = E_{\mathrm{total}}(\mathrm{ZB}) - E_{\mathrm{total}}(\mathrm{WZ}). \quad (7)$$

Here, the negative and the positive values in $\Delta\epsilon_{\mathrm{total}}$(ZB−WZ) prefer to form the ZB and the WZ structures, respectively. In Table I, the value of −26.5 meV/Ga1As1 manifests itself in higher stability for the ZB structure, which is in good agreement with that previously reported, −24.04 meV/Ga1As1 [$(c/a)_{\mathrm{WZ}}$ relaxed] or −24.74 meV/Ga1As1 [$(c/a)_{\mathrm{WZ}}$ unrelaxed].[17] Second, we try to decompose the total energy in Eq. (1) in order to examine a difference in the ionicity between the ZB and the WZ structures.[27] As shown in Table I, both Hartree potential and kinetic energy terms have negative values, indicating that electronic charges for the WZ structure are more localized. This localization makes the core–electron potential energy have a large positive value, reflecting larger ionicity in the WZ structure.

Third, it is of importance to know whether or not a relation between the components of the total energy and the ionicity in GaAs is generalized. Table I shows the $\Delta\epsilon_i(3C-2H)$ (i indicates each component) for Ge which closely resembles GaAs in a shell structure [in a homopolar semiconductor like Ge the ZB-type structure is often called a cubic-diamond structure ($3C$) and the WZ type a hexagonal-diamond structure ($2H$)]. In Table I, the absolute values of $\Delta\epsilon_{\mathrm{Hartree}}(3C-2H)$ and $\Delta\epsilon_{\mathrm{potential}}(3C-2H)$ for Ge are considerably small compared to those for GaAs, which is indicative of the larger ionicity of GaAs. Such a trend is seen also for a combination of Sn and InSb in the fifth row of the periodic table, as shown in Table I.

Here, the question should arise as to why a WZ structure has larger ionicity, in other words, why a ZB–WZ polymorphism is dominated by a degree of ionicity. Figures 1(d) and 1(e) show the stacking sequences for the ZB and WZ structures, $AaBbCcAaBbCc\cdots$ in the [111] direction and $AaBbAaBb\cdots$ in the [0001] direction, respectively. As seen in Figs. 1(d) and 1(e), the respective stackings result in the staggered and the eclipsed conformation, and in each of which the cation (group III)–anion (group V) distance between the third nearest neighbors is long and short. Thus, in one and the same III–V compound, attractive Coulomb energy between the cation and the anion is gained more in the WZ structure than in the ZB. GaAs is a case in which the attractive Coulomb energy is not the most effective factor for stabilizing the WZ structure. Of systems having larger ionicity than the GaAs, III nitrides such as GaN and InN have the WZ structure, and are dominated by this kind of energy.

B. Stability of faultily stacked layers on GaAs{111} *A*, *B*

Tables II and III show the differences in total energy, $\Delta\epsilon_i$(NS2−FS2), $\Delta\epsilon_i$(NS4−FS4), and $\Delta\epsilon_i$(NS4−FS2NS2) for the GaAs(111) *A* surface and those for the (111) *B* surface, respectively. The comparison between $\Delta\epsilon_{\mathrm{total}}(\mathrm{NS2-FS2})_{(111)\,A}$ and $\Delta\epsilon_{\mathrm{total}}(\mathrm{NS2-FS2})_{(111)\,B}$ addresses the issue that a faultily stacked layer is not as stable on the (111) *B* surface as on the (111) *A*, which reproduces well the experimental results. Looking more carefully into the respective energy components in Tables II and III, we notice the following: The values of the Hartree, potential, and kinetic components for the (111) *A* model are negative, positive, and negative, respectively, while for the (111) *B*

TABLE III. Energy differences between the models for GaAs(111) *B* (eV/slab).

	$\Delta\epsilon_i(\mathrm{NS2-FS2})_{(111)\,B}$	$\Delta\epsilon_i(\mathrm{NS4-FS4})_{(111)\,B}$	$\Delta\epsilon_i(\mathrm{NS4-FS2NS2})_{(111)\,B}$
Ewald	+28.9997	+11.6400	+3.7164
Hartree	+28.9390	+10.8345	+3.4004
Exchange and correlation	−0.1256	+0.1370	+0.0667
Potential	−58.1446	−22.1901	−7.0107
Kinetic	+0.2792	−0.5076	−0.2051
Total	−0.0523	−0.0864	−0.0323

TABLE IV. Averaged net charges for the models employed.

	All the atoms in the models $\lvert\delta\bar{Q}\rvert_{\text{overall}}$	Atoms in the 1st layer $\lvert\delta\bar{Q}\rvert_{\text{1st layer}}$
NS-(111) *A*	0.131	0.244
NS-(111) *B*	0.134	0.224
FS-(111) *A*	0.140	0.262
FS-(111) *B*	0.122	0.224

model, they have the opposite signs. These results imply that in the (111) *A* surface the FS2 layer has a larger ionic character than the NS2 layer. This is a plausible result, because the WZ-type stacking of *AaBbAa* including the FS2 layer induces larger ionicity, as mentioned in Sec. III A. Interestingly, in the (111) *B* surface, contrary to the (111) *A*, the FS2 layer is less ionic than the NS2 layer. In this case, the attractive Coulomb energy does not become a dominant component for stabilizing the FS2.

Next, we focus on the difference between the calculated energies for the secondly stacked bilayers on the (111) *A* and (111) *B* surfaces, that is, $\Delta\epsilon_i$(NS4−FS4) or $\Delta\epsilon_i$(NS4−FS2NS2). As seen in Tables II and III, since for both the surfaces $|\Delta\epsilon_{\text{total}}(\text{NS4}-\text{FS4})|$ is larger than $|\Delta\epsilon_{\text{total}}(\text{NS2}-\text{FS2})|$ or $|\Delta\epsilon_{\text{total}}(\text{NS4}-\text{FS2NS2})|$, the second layers always preserve the NS type independent of the stacking type of the first bilayer, NS or FS. Thus, whenever the FS layer becomes the outermost layer, a rotational-type twin boundary,[28] indicated by a broken line in Fig. 1(e), appears. More noteworthy for the (111) *A* model is the result that each of the energy components has an opposite sign between $\Delta\epsilon_i(\text{NS2}-\text{FS2})_{(111)\,A}$ and $\Delta\epsilon_i(\text{NS4}-\text{FS4})_{(111)\,A}$. This means that double bilayers having WZ-type stacking tend to weaken the ionic character even on the (111) *A* surface. From these results, we can say that if the outermost bilayer in the growth front results in smaller ionicity compared to the originally ZB-type stacked layer, then it would have little chance to grow in WZ-type stacking.

As one step toward the understanding of the local charge states and the chemical bonding character in the growing surface layers of GaAs, we performed the cluster calculations for the models mentioned in Sec. II B. Given in Table IV are the absolute values of net charges for the clusters, averaged over the models, $|\delta\bar{Q}|_{\text{overall}}$, which are calculated by use of the Mulliken population analysis.[24] Table IV shows that the net charge, i.e., ionicity, increases in the sequence of FS-(111) *A*>NS-(111) *B*>NS-(111) *A*>FS-(111) *B*. A similar tendency is found by the total energy calculations. What has to be noticed, further, is that the net charges for the atoms in the first layer, $|\delta\bar{Q}|_{\text{1st layer}}$, become large in the order of FS-(111) *A*>NS-(111) *A*>FS-(111) *B* =NS-(111) *B*. On the other hand, the chemical bonding features associated with the first layer can be characterized by the different charge density maps, $\Delta\rho=\rho$ (for the cluster) $-\Sigma\rho$ (constituent atom), for the NS and FS cluster models of the (111) *A* and (111) *B* surfaces. Figures 4(a)–4(d) show that the charge density for the dangling bond (indicated as "D" in Fig. 4) decreases (increases) in the first layer of the

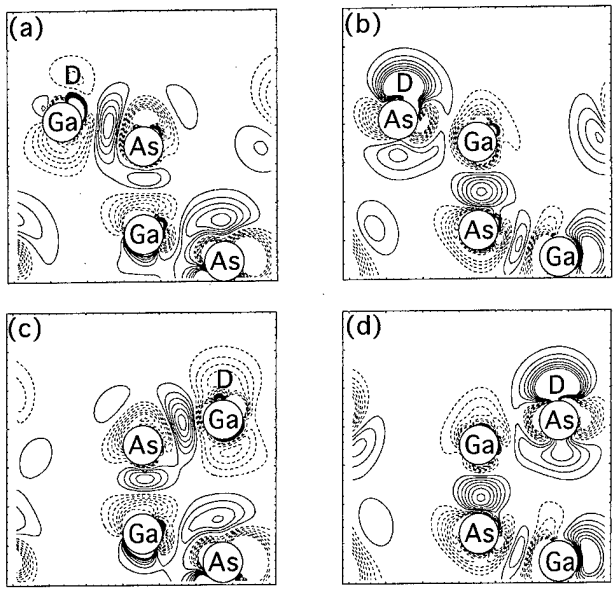


FIG. 4. Difference charge density maps for the models employed. The contours are drawn for the $(1\bar{1}0)$ plane perpendicular to the {111} *A*, *B* surfaces. The solid and the dashed lines represent accumulated and depleted charges, respectively.

(111) *A* [(111) *B*] surface. The following explanation has this situation to understand more deeply: At the outermost surface of GaAs{111}, if it does not exhibit any reconstruction, the partially filled sp^3 dangling bonds will remain unbonded; the dangling bond energy level for the electronegative As is lowered more than that for the electropositive Ga.[8,29–31] Hence, the surface energy of (111) *A* (*B*) would be reduced when dangling bonds states for Ga (As) become empty (filled). Next, we draw attention to the difference in the contributions of the NS and the FS to the charge states of the first layer. Looking at the different charge density maps for the (111) *A* surface again, we notice that the dangling bond charges at the FS surface are much less than those at the NS surface. For the (111) *B* surface, changes in the dangling bond charges are hardly recognized at either of the NS and FS surfaces, as seen in Figs. 4(b) and 4(d), which agrees well with the above population analysis result.

Finally, by combining these cluster calculation results with the total energy calculation ones, we can reach the following roots for the more frequent induction of in-plane faults in the LEC-grown film on the GaAs(111) *A* substrate. (1) The outermost layer of the (111) *A* surface inherently has larger ionicity than the (111) *B* one. (2) In the (111) *A* surface, when the outermost layer is of the faulted type, it gains enough attractive Coulomb energy to stabilize itself. These two roots make it possible to establish the conditions most suitable for growing in-plane fault-free GaAs crystals, when used with many other crystal growth techniques as well as with LEC.

IV. CONCLUSIONS

Electronic states for the growing surface layers of GaAs{111} *A*, *B* were calculated by use of the discrete varia-

tional $X\alpha$ cluster method and the plane wave nonlocal pseudopotential method. The results obtained are that a stacking fault in the GaAs(111) A surface layer, which includes a local stacking layer of $AaBbAa$ based on a WZ structure, is more easily introduced than in the (111) B surface. Such a polar surface dependence is attributed to the difference in the chemical bonding features of the {111} A, B surface layers: The atoms in the $AaBbAa$ stacked layer of the (111) A surface have a more ionic bonding character compared to the (111) B case because the former gains larger attractive Coulomb interaction energy. Further, the double bilayer stacking of WZ type leads to a decrease of ionicity in the bilayer on both the {111} A, B surfaces. These results bring us to conclude that, when the growing (111) A surface layer gains larger ionicity than the original surface by introducing any stacking faults, the WZ type of $AaBbAa$ is prevalent.

ACKNOWLEDGMENTS

The authors wish to thank Professor H. Adachi (Kyoto University) and Dr. K. Shiraishi (NTT Basic Research Laboratory) for use of the calculation programs. They would like to acknowledge fruitful discussions with Dr. A. Ohtake and C. Tsutsui. This work was partially supported by a Grant-in-Aid for Scientific Research (B) from the Ministry of Education of Japan and by a Grant-in-Aid for Research for the Future (Next-Generation Process Technologies: wafer-scale fabrication processes of quantum dots) from the Japan Society for the Promotion of Science.

[1]P. L. Moody, H. C. Gatos, and M. C. Lavine, J. Appl. Phys. **31**, 1696 (1960).
[2]A. Steinemann and U. Zimmerli, Solid-State Electron. **6**, 597 (1963).
[3]H. C. Gatos, P. L. Moody, and M. C. Lavine, J. Appl. Phys. **31**, 212 (1960).
[4]R. K. Mueller and R. L. Jacobson, J. Appl. Phys. **32**, 550 (1961).
[5]K. F. Hulme and J. B. Mullin, Solid-State Electron. **5**, 211 (1962).
[6]A. Ohtake (private communication).
[7]J. C. Phillips, *Bonds and Bands in Semiconductors* (Academic, New York, 1973), p. 13.
[8]W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980), p. 64.
[9]P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).
[10]W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).
[11]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[12]D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[13]G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[14]L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[15]K. Shiraishi and H. Kamimura, Mater. Sci. Eng., B **3**, 287 (1989).
[16]E. P. Wigner, Phys. Rev. **46**, 1002 (1934).
[17]C.-Y. Yeh, Z. W. Lu, S. Froyen, and A. Zunger, Phys. Rev. B **46**, 10 086 (1992).
[18]K. Shiraishi, J. Phys. Soc. Jpn. **59**, 3455 (1990).
[19]X. Gonze, P. Käckell, and M. Scheffler, Phys. Rev. B **41**, 12 264 (1990).
[20]X. Gonze, R. Stumpf, and M. Scheffler, Phys. Rev. B **44**, 8503 (1991).
[21]The maximum difference between $\Delta\epsilon_{\text{Total}}$ s for cutoffs of 7.29 and 12.25 Ry is of the order of $\sim 10^{-4}$ Ry, which is small enough to get the same optimized geometries.
[22]D. E. Ellis and G. S. Painter, Phys. Rev. B **2**, 2887 (1970).
[23]H. Adachi, M. Tsukada, and C. Satoko, J. Phys. Soc. Jpn. **45**, 875 (1978).
[24]R. S. Mulliken, J. Chem. Phys. **23**, 833 (1955).
[25]C. Satoko, M. Tsukada, and H. Adachi, J. Phys. Soc. Jpn. **45**, 1333 (1978).
[26]Small variations in the values of r_{n0} hardly affect the calculated results.
[27]J. Nakamura, H. Nakajima, and T. Osaka, Appl. Surf. Sci. **121/122**, 249 (1997).
[28]B.-T. Lee and J. Y. Lee, Appl. Phys. Lett. **57**, 346 (1990).
[29]J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw–Hill, New York, 1974), Vol. 4, p. 91.
[30]J. Nakamura, H. Konogi, and T. Osaka, Phys. Rev. B **51**, 5433 (1995).
[31]J. Nakamura, H. Konogi, H. Sato, and T. Osaka, J. Phys. Soc. Jpn. **66**, 1656 (1997).

Papers from the Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing

16–20 November 1997
Osaka, Japan

Sponsored by
The American Vacuum Society
The Japan Society for the Promotion of Science
The National Science Foundation

Editors for the Gamo Conference:
John Melngailis, Kenji Gamo, and Jon Orloff

Published by the American Vacuum Society through
the American Institute of Physics, New York, 1998

International Standard Serial Number: 0734-211X

International Standard Book Number: 1-56396-851-7

Copyright 1998 by the American Vacuum Society

Published by the American Institute of Physics
500 Sunnyside Blvd., Woodbury, NY 11797-2999

Printed in the United States of America

The Conference is published in both the 1998 Jul/Aug issue of the
Journal of Vacuum Science and Technology B, Vol. 16, No. 4 and in a special casebound edition available to participants

ORGANIZING COMMITTEE

Kenji Gamo
Osaka University

Jon Orloff
University of Maryland

John Melngailis
University of Maryland

Papers from the Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing

Preface 2438

General Interest

Material-wave nanotechnology: Nanofabrication using a de Broglie wave
Shinji Matsui and Jun-Ichi Fujita 2439

Economic and technical case for ion projection lithography
I. L. Berry 2444

Microelectromechanical tunneling sensor fabrication and post-processing characterization using focused ion beams
F. P. Stratton, R. L. Kubena, H. H. McNulty, R. J. Joyce, and J. Vajo 2449

Sources and Systems

Multipurpose nanobeam source with supertip emitter
S. Kalbitzer and A. Knoblauch 2455

Focused ion beam optical column design and consideration on minimum attainable beam size
Kiyoshi Sakaguchi and Tetsu Sekine 2462

Overlay accuracy tests for direct write implantation
S. Mogren and I. L. Berry 2469

Development of ion and electron dual focused beam apparatus for high spatial resolution three-dimensional microanalysis of solid materials
Zhaohui Cheng, Tetsuo Sakamoto, Masanori Takahashi, Yasuyuki Kuramoto, Masanori Owari, and Yoshimasa Nihei 2473

Key technologies of a focused ion beam system for single ion implantation
T. Matsukawa, T. Shinada, T. Fukai, and I. Ohdomari 2479

Development of wide range energy focused ion beam lithography system
M. Kinokuni, H. Sawaragi, R. Mimura, R. Aihara, and A. Forchel 2484

Current status of single ion implantation
Takahiro Shinada, Yoshinori Kumura, Jun Okabe, Takashi Matsukawa, and Iwao Ohdomari 2489

Beam Induced Chemical and Physical Effects

Chemically and geometrically enhanced focused ion beam micromachining
P. E. Russell, T. J. Stark, D. P. Griffis, J. R. Phillips, and K. F. Jarausch 2494

Focused ion beam technology applied to microstructure fabrication
Michael J. Vasile, Raja Nassar, and Jushan Xie 2499

Focused-ion-beam-assisted etching of diamond in XeF_2
Jun Taniguchi, Naoto Ohno, Shuuichi Takeda, Iwao Miyamoto, and Masanori Komuro 2506

Focused ion beam etching of resist/Ni multilayer films and applications to metal island structure formation
Masayoshi Nakayama, Fujio Wakaya, Junichi Yanagisawa, and Kenji Gamo 2511

Focused ion beam direct deposition and its applications
Shinji Nagamachi, Masahiro Ueda, and Junzo Ishikawa 2515

Microanalysis

Transmission electron microscopy observation of thin foil specimens prepared by means of a focused ion beam
H. Saka 2522

(Continued)

Dynamic Monte Carlo simulation for depth profiling by ion-sputter etching: Application to the AlAs/GaAs multilayered system
Hyung-Ik Lee and Ryuichi Shimizu **2528**

Proposals for exact-point transmission-electron microscopy using focused ion beam specimen-preparation technique
T. Ishitani, Y. Taniguchi, S. Isakozawa, H. Koike, T. Yaguchi, H. Matsumoto, and T. Kamino **2532**

Nanofabrication, Devices and Implantation

GaAs microcrystal growth on semiconductor surfaces by low energy focused ion beam
Toyohiro Chikyow and Nobuyuki Koguchi **2538**

AlGaAs/GaAs heterostructures grown on a focused-Be-ion-beam written backgate
Y. Hirayama and T. Saku **2543**

Formation of GaAs/AlGaAs constricted-channel field-effect transistor structures by focused Ga implantation and transport of electrons via focused ion beam induced localized states
H. Kim, T. Noda, and H. Sakaki **2547**

Hopping conduction and localized states in p-Si wires formed by focused ion beam implantations
Hirotaka Iwano, Shigeaki Zaima, and Yukio Yasuda **2551**

End-point detection using focused ion beam-excited photoemissions in milling deep small holes in large scale integrated circuit structures
Hiroshi Yamaguchi, Keiya Saitoh, and Michinobu Mizumura **2555**

Focused ion beam implantation for opto- and microelectronic devices
H. König, N. Mais, E. Höfling, J. P. Reithmaier, A. Forchel, H. Müssig, and H. Brugger **2562**

Direct writing of active loads by focused ion beams
C. Wiemann, M. Versen, and A. D. Wieck **2567**

***In situ* scanning tunneling microscope studies of high-energy, focused ion implantation of Ga into GaAs: Direct observation of ion beam profiles**
G. A. C. Jones, P. D. Rose, and S. Brown **2570**

Ion beam synthesis of cobalt disilicide using focused ion beam implantation
J. Teichert, L. Bischoff, and S. Hausmann **2574**

AUTHOR INDEX **2578**

PREFACE

On the following pages are papers presented at the 1997 Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing. The workshop was held at the Senri Hankyu Hotel in Osaka, Japan from 16–20 November 1997. It was the third workshop in this series of Japan/U.S. workshops devoted mainly to focused ion beam (FIB) technology and applications, and was sponsored by the Japan Society for the Promotion of Science and the National Science Foundation in the U.S. with co-sponsorship by the American Vacuum Society. Industrial sponsors from Japan were: EICO, Hitachi, JEOL, Kobelco, NEC, NTT, SELETE, Seiko, and Shimazu. Industrial sponsors from the U.S. were: FAI (now Nanofab), FEI, Micrion, and Schlumberger.

In the ten years since the first international workshop on "Focused Ion Beam Technology and Applications" the number of FIB systems used in industry has grown by more than tenfold. While ten years ago most FIB machines were in industrial or academic research labs, today the vast majority of the new machines that have been added are in use in the semiconductor industry for failure analysis, circuit rewiring, mask repair, or transmission electron microscopy sample preparation. Many of the topics discussed in the workshop ten years ago are now central to the practical applications, and no doubt the topics of this workshop will likewise serve to enhance the technology for industry as well as provide intellectual stimulation. The technical presentations covered a range of subjects including general beam technologies, ion sources and systems, beam induced chemistry, microanalysis, and ion implantation.

We are grateful to the authors and to the reviewers for the quality of the papers, to the staff and students of Osaka University for help in organizing the seminar, and to Valerie Mainwaring for the preparation of the proceedings.

John Melngailis
Kenji Gamo
Jon Orloff
Proceedings Editors

Material-wave nanotechnology: Nanofabrication using a de Broglie wave

Shinji Matsui[a] and Jun-Ichi Fujita
Fundamental Research Laboratories, NEC Corporation, Tsukuba 305, Japan

(Received 16 January 1998; accepted 20 April 1998)

An approach for nanofabrication using de Broglie wave has been developed. Line and dot patterns with 100-nm periodicity were delineated on a PMMA resist by electron-beam holography with a thermal field-emitter gun and an electron biprism. This technique allows us to produce nanoscale periodic patterns simultaneously. Furthermore, the possibility of nanostructure fabrication by atomic-beam holography has been demonstrated by using a laser-trap technique and a computer-generated hologram made by electron-beam lithography. © *1998 American Vacuum Society.* [S0734-211X(98)06504-4]

I. INTRODUCTION

Recent years have witnessed a number of investigations concerning nanostructure technology. The objective of research on nanostructure technology is to explore the basic physics, technology, and applications of ultrasmall structures and devices with dimensions in the sub-100 nm regime. Nanostructure devices are now being fabricated in many laboratories to explore various effects, such as those created by downscaling existing devices, quantum effects in mesoscopic devices, tunneling effects in superconductors, etc. In addition, new phenomena are being explored in an attempt to build switching devices with dimensions down to the molecular level.

Electron-beam (EB) lithography is the most widely used and versatile lithography tool used in fabricating nanostructure devices. Because of the availability of high-quality electron sources and optics, EB can be focused to diameters of less than 10 nm.[1–5] The minimum beam diameters of scanning electron microscopes (SEMs) and scanning transmission microscopes are 1.5 and 0.5 nm, respectively.[6,7] While focused-ion beam (FIB) can be focused close to 8 nm.[8] EB and FIB can be used to make nanoscale features in the 100 to 1 nm regime. Scanning tunneling microscope (STM) is used for atomic technology in the region of 1–0.1 nm. These current technologies using particle characteristics of electrons and atoms almost establish the ultimate resolution. On the other hand, wave characteristics have not been applied to nanostructure fabrication.

This article describes nanofabrication using de Broglie waves of electrons and atoms.

II. ELECTRON-BEAM HOLOGRAPHY

Holographic lithography has an advantage that it can produce a number of periodic patterns simultaneously. Electron holographic lithography was applied to nanofabrication. Electron interference fringes were recorded on a PMMA resist by using a W(100) TFE gun and an electron biprism, and the fabricated patterns were observed by a conventional transmission electron microscope (TEM) and atomic force microscope (AFM).[9,10]

[a]Electronic mail: matsui@selete.co.jp

The electron optics of TEM with a W(100) thermal field emitter (TFE) gun for electron holographic lithography is schematically illustrated in Fig. 1. An electron beam of 40 kV is focused above an electron biprism with two condenser lenses. The Möllenstedt-type electron biprism is constructed of two grounded plane electrodes and a fine-wire electrode, called a filament, between them. When a positive voltage V_B is supplied to the filament, electron waves traveling on both sides of the filament are deflected and superimposed to form interference fringes on an observation plane. A Pt wire of 0.6 μm in diameter was used as the filament. As is well known, two coherence waves overlapping at an angle of θ produce interference fringes with spacing s represented by

$$s=(\lambda/2)/[\sin(\theta/2)], \tag{1}$$

where λ denotes the de Broglie wavelength of 6.0×10^{-3} nm in this case.

Figure 2 shows four-wave interference fringes through an X biprism. Setting an X biprism below two condenser lenses instead of the Möllenstedt-type biprism, which has two filaments placed normal to each other and both are supplied V_B, four coherent waves produce fringes like a checkerboard below the intersection of filaments, with the same spacing s, as given by Eq. (1). Thus, electron holographic lithography would be, in principle, possible to generate line and dot patterns whose minimum spacing is $\lambda/2$, which is comparable to the crystal lattice spacing.

A 30-nm-thick PMMA, spin coated on a 50-nm-thick self-supporting SiN_x membrane and prebaked at 170 °C for 20 min, was set on the observation plane 70 mm below the biprism. The self-supporting nitride (SiN) membrane was about 60 μm square and used to place the PMMA below interference fringes appropriately. Electron exposure to produce line patterns was carried out for 18 s with a dose of 25 μC/cm^2, which was measured at the fringe part. Then, the PMMA was developed in MIBK:IPA=1:3 for 1.0 min and rinsed in IPA for 30 s. Similarly, PMMA dot patterns were exposed, at half the dose as that for the line patterns, in order to maintain whole dots. The electron exposure to produce dot patterns was carried out for 9.0 s with a dose of 13 μC/cm^2. The PMMA was developed in MIBK:IPA for 3.0 min and rinsed in IPA for 1.0 min.

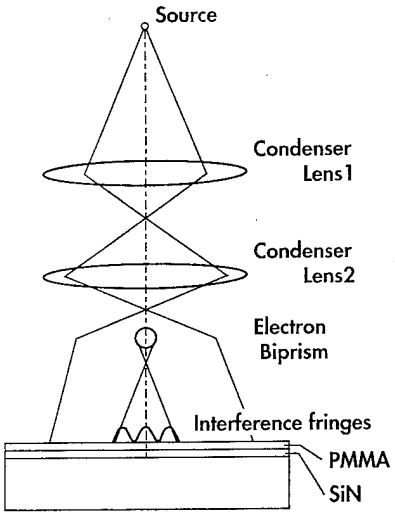


FIG. 1. Scheme of electron optics of TEM for electron holographic lithography.

Figure 3(a) shows interference fringes of the Möllenstedt-type electron biprism, which was magnified 530 times by the lenses below the observation plane and recorded on a photoplate with 1.0 s exposure. Figure 3(b) shows the AFM image of the same interference fringes as those in Fig. 3(a), which was recorded on PMMA. The thickness of PMMA is represented by a photocontrast in Fig. 3(b), and the thicker PMMA corresponds to the brighter part of the image. The supplied voltage to the filament of electron biprism, V_B, was 5.3 V and the spacing of fringes s was 108 nm in Figs. 3(a) and 3(b). Figure 4(a) shows interference fringes of the X biprism magnified and recorded on a photoplate, and Fig. 4(b) shows the AFM image of interference fringes recorded on PMMA. The supplied voltage to filament V_B, was 5.0 V and the spacing of fringes s, was 125 nm in Figs. 4(a) and 4(b). In Fig. 4(a), dot patterns are found at the intersection where four-wave interference occurred and line patterns around the dot patterns where two-wave interference occurred. In Fig. 4(b), about 10×10 dots are recognized, but lines are not observed, owing to the reduction of dose. Consequently, Figs. 3(b) and 4(b) show that line and dot patterns were fabricated successfully, and the dose needed for lines is

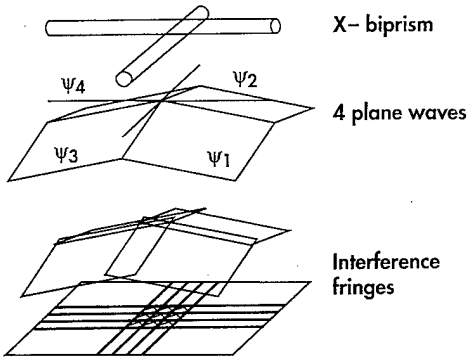


FIG. 2. Four-wave interference fringes through an X biprism.

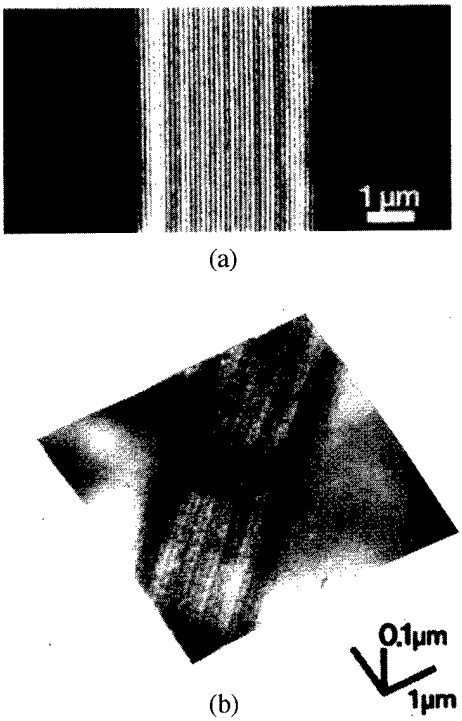


FIG. 3. (a) Two-wave interference fringes magnified and recorded on a photoplate. (b) Interference fringes corresponding to (a) recorded on PMMA; V_B: 5.3 V and s: 108 nm.

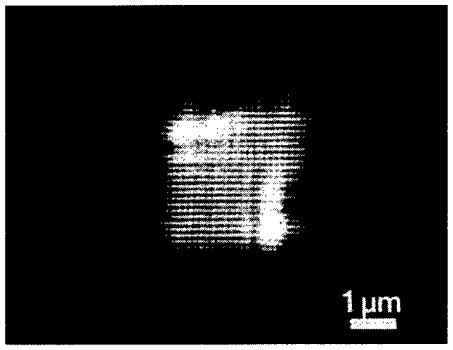
about twice as that for dots. More precise fabrication would be possible by optimizing the dose.

In order to produce finer patterns than 100 nm in a period, the larger overlapping angle θ, i.e., the larger supplied voltage to the filament V_B, should be selected. A simple assessment suggests that spacing s, becomes 1 nm when V_B is 2.4 kV with the same electron optics. Carbon contamination line patterns with a period of 23 nm fabricated by a 30 kV SEM were reported by using the same technique.[11] More complex periodic patterns except the line and dot will be also fabricated using a designed biprism on a SiN thin membrane made by EB lithography and dry etching.

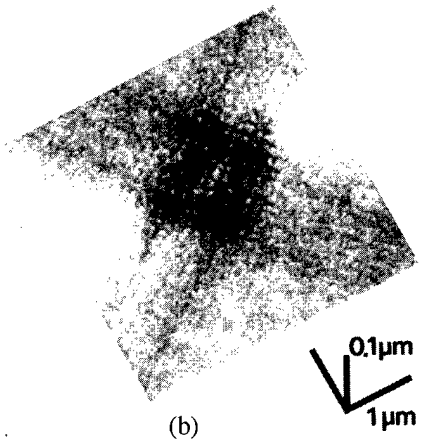
Atomic resolution of EB holography with a biprism will be demonstrated by using both a hydrogen monolayer adsorbed on a Si (111) or (100) surface as a resist and the STM tip as a coherent electron source in an UHV chamber in the future. Hydrogen atom desorption will be observed by using the same STM tip after EB holography atomic fabrication. After demonstration of atomic fabrication by EB holography, it may be possible to make artificially new lattice structures on the surface.

III. ATOMIC-BEAM HOLOGRAPHY

Atomic manipulation based on a holographic principle has been demonstrated by using a laser-trap technique and a computer-generated hologram (CGH) made by EB lithography.[12] One approximation of a CGH is the binary hologram, in which the hologram takes a binary value, either 100% transparent or 100% opaque. This hologram can be



(a)



(b)

FIG. 4. (a) Two- and four-wave interference fringes magnified and recorded on a photoplate. (b) Interference fringes corresponding to (a) recorded on PMMA; V_B: 5.0 V and s: 125 nm.

directly translated to a hologram for atomic de Broglie waves, by cutting out the pattern on a film that is equal to the pattern of the 100% transmission area of the binary hologram. A monochromatic atomic wave reconstructs an atomic pattern by passing the hologram.

The hologram used in this experiment was a Fourier hologram, which produced the Fourier-transformed wave-front of the object. When the hologram is illustrated with a plane wave, the far-field pattern of the diffracted wave produces an image of the object. The object used in this experiment was a transparent F-shaped pattern, in which the transparent portion had a constant amplitude and random phase distribution. The object was represented by the complex transmission amplitude at points on a 128×128 matrix covering the F-shape pattern. The two-dimensional array of numbers was Fourier transformed using a fast Fourier transform algorithm, and the resulting 128×128 complex areas (cells) of the Fourier hologram. The transmission function of each cell of the hologram was expressed by a matrix of 4×4 subcells.

A 100-nm-thick SiN membrane was used for the hologram. The binary pattern was transferred to a ZEP resist (Nippon Zeon Corporation) on the SiN membrane by an EB writing system. Subsequent CF_4 plasma etching created holes in the membrane. A scanning electron micrograph of the hologram is shown in Fig. 5. The size of the subcell was

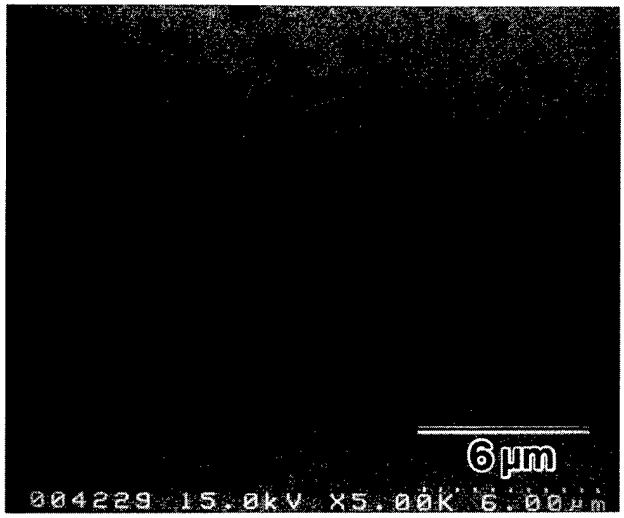


FIG. 5. Binary CGH hologram on SiN membrane made by EB lithography and dry etching.

0.3×0.3 μm square, so the size of the entire hologram was 153.6×153.6 μm. To increase the intensity of the deflected beam, the same pattern was repeated ten times along the x and y directions, making the overall size of the hologram 1.5×1.5 mm.

A schematic diagram of this experiment is shown in Fig. 6. The ultracold Ne atomic beam was generated by the reported method.[13] The cloud of Ne atoms in the trap was ~0.3 mm in diameter, and the one-directional average velocity of the atoms was 20 cm s^{-1}. The hologram was placed 40 cm below the trap and was mounted on the top of a 0.2-mm-diam diaphragm. The size of the diaphragm limited the resolution of the image of the Fraunhofer hologram. The position of the hologram was not adjusted because any small portion of the hologram could produce the same image. The average

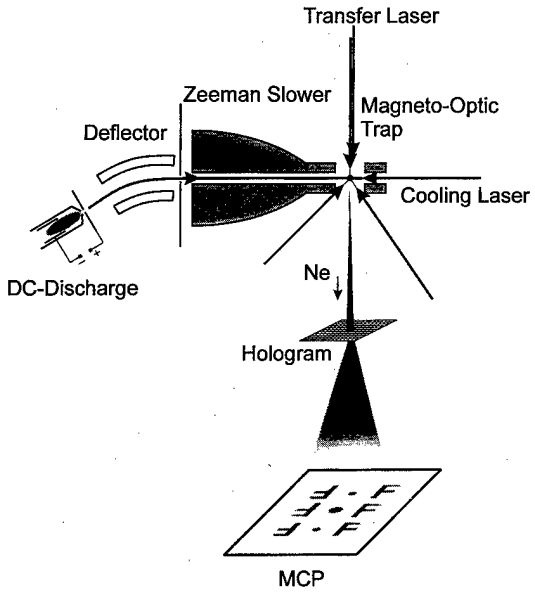
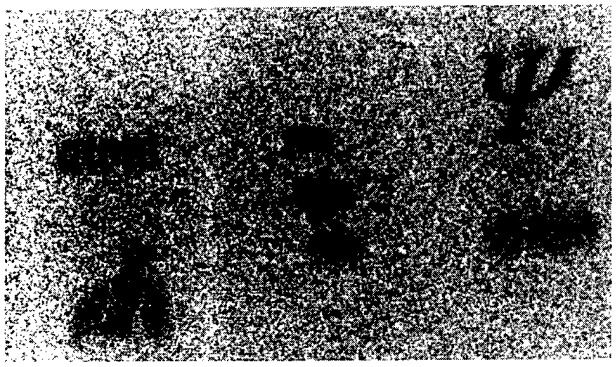


FIG. 6. Experimental apparatus of atomic-beam holography.

(a)



(b)

FIG. 7. Reconstructed image: (a) "F" pattern; and (b) "atom, Ne, and Ψ" pattern.

atomic velocity at the hologram was 2.8 ms^{-1}, corresponding to a de Broglie wavelength λ of 7.1 nm. The acceleration due to gravity reduced the relative velocity spread to ~0.28%. To detect the Fraunhofer diffracted pattern from the hologram, multichannel plate detector (MCP) was placed 45 cm below the hologram. Figure 7(a) shows the reconstructed F pattern. The data were accumulated for 10 h, and the total atom number of spots on the figure was 6×10^4. Figure 7(b) shows another example of reconstructed patterns which represents characters of "atom, Ne, and ϕ."

In this experiment, a focusing lense for imaging was not used, but it is possible to combine the function of a focusing lens into the hologram.[14] In such a hologram, the resolution is determined by the same rule as applies to an optical lens. The binary hologram does not control the phase and amplitude of the wave inside a hole. When the hologram is the sole component for atomic-beam manipulation, the practical limit is approximately the minimum size of holes, which is in the range 10–100 nm.

As a next step, there is a possibility of making any three-dimensional nanostructures by using CGH with an electron/atom de Broglie wave, as shown in Fig. 8. A three-dimensional image is obtained by atom direct deposition using CGH with the phase and amplitude control. In the case of electrons, it will be expected to form by EB chemical vapor deposition.

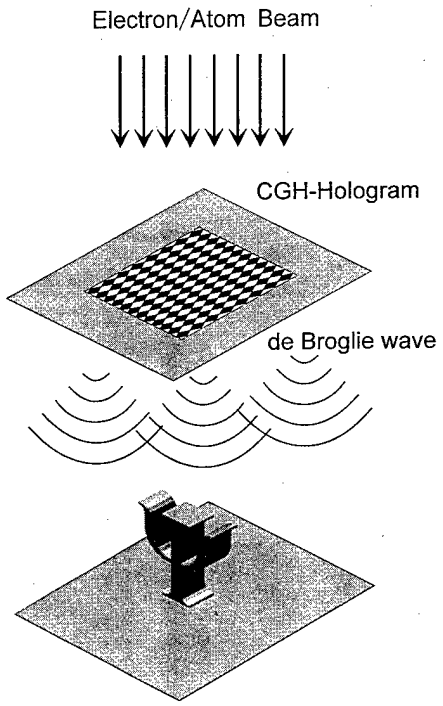


FIG. 8. Concept of three-dimensional nanofabrication by CGH with an electron/atom de Broglie wave.

IV. SUMMARY

Material-wave nanotechnology using electron- and atomic-beam holography has been developed. This technique allows nano- and atomic-scale structures to be produced simultaneously. Especially material-wave nanotechnology using atomic-beam holography with CGH is promising from the point of view of making any shape and direct deposition. We have to continue the research for the remaining important subjects such as atomic resolution and CGH demonstration in EB holography, resolution limit of atomic-beam holography, three-dimensional patterning and opening to other fields (e.g., material-wave communication as an analogy of optical communication) in material-wave nanotechnology.

[1]F. Emoto, K. Gamo, S. Namba, N. Samoto, and R. Shimizu, Jpn. J. Appl. Phys., Part 2 **24**, L809 (1985).
[2]W. Chen and H. Ahmed, Appl. Phys. Lett. **63**, 1116 (1993).
[3]K. Kurihara, K. Iwadate, H. Namatsu, M. Nagase, H. Takenaka, and K. Murase, Jpn. J. Appl. Phys., Part 1 **34**, 6940 (1995).
[4]J. Fujita, H. Watanabe, Y. Ochiai, S. Manako, J. S. Tsai, and S. Matsui, Appl. Phys. Lett. **66**, 3065 (1995).
[5]J. Fujita, Y. Ohnishi, Y. Ochiai, and S. Matsui, Appl. Phys. Lett. **68**, 1297 (1996).
[6]M. Isaacson and A. Murray, J. Vac. Sci. Technol. **19**, 1117 (1981).

[7]D. R. Allee and A. N. Broers, Appl. Phys. Lett. **57**, 2271 (1990).
[8]R. L. Kubena, J. W. Ward, F. P. Stratton, R. J. Joyce, and G. M. Atkinson, J. Vac. Sci. Technol. B **9**, 3079 (1991).
[9]K. Ogai, S. Matsui, Y. Kimura, and R. Shimizu, Jpn. J. Appl. Phys., Part 1 **32**, 5988 (1993).
[10]K. Ogai, Y. Kimura, R. Shimizu, J. Fujita, and S. Matsui, Appl. Phys. Lett. **66**, 1560 (1995).
[11]S. Fujita, S. Maruno, H. Watanabe, Y. Kusumi, and M. Ichikawa, Appl. Phys. Lett. **66**, 2754 (1995).
[12]J. Fujita, M. Morinaga, T. Kishimoto, M. Yasuda, S. Matsui, and F. Shimizu, Nature (London) **380**, 691 (1996).
[13]F. Shimizu, K. Shimizu, and H. Takuma, Opt. Lett. **16**, 339 (1991).
[14]M. Morinaga, M. Yasuda, T. Kishimoto, F. Shimizu, J. Fijita, and S. Matsui, Phys. Rev. Lett. **77**, 802 (1966).

Economic and technical case for ion projection lithography

I. L. Berry[a)]
Microelectronics Research Laboratory, Columbia, Maryland 21045

(Received 17 April 1998; accepted 1 June 1998)

For more than a decade, technical arguments have been used to predict the demise of optical lithography. Recently economic arguments are being used, dramatically increasing the likelihood of the predictions coming true. Presently four technologies are competing to replace optical lithography. This article will describe the current state-of-the-art in ion projection lithography, and contrast its advantages and disadvantages with respect to the other nonoptical technologies. An argument will be made that ion projection lithography, uniquely, offers the possibility of reducing the dramatically escalating costs of lithography, resulting in overall lower integrated circuit production costs. This potential for reduced production costs is the prime motivator to consider ion projection lithography as a production lithography candidate. [S0734-211X(98)13804-0]

I. INTRODUCTION

Efforts have begun in the United States and elsewhere to identify several candidate technologies capable of extending lithographic resolution to 0.1 μm and smaller, and to attempt to generate consensus on what is the "status, risk and challenges of each of the advanced lithographic alternatives."[1] Ion projection lithography (IPL) is one of the candidate technologies. This article will describe the IPL program, its cost advantages and contrast its capabilities with respect to the other competing technologies.

II. ECONOMICS OF IPL

Since the early 1980s, IMS in Vienna Austria has been developing ion projection lithography technology,[2–5] but has until recently received little interest by semiconductor manufacturers. Earlier this year, Siemens, ASM Lithography, Leica and the Institute for Microelectronics–Stuttgart, among others, have joined with IMS-Vienna under a European MEDEA[6] program to develop IPL as a lithography candidate technology for 0.13 μm generation devices and beyond. Under Siemens leadership, IPL technology is becoming more accepted worldwide. Several other major semiconductor manufacturers have expressed interest in the technology and in joining the program. Supporters of the technology agree that there are technological risks, but for dynamic random access memory (DRAM) production especially, IPL has a major appeal. It can reduce the production costs of integrated circuits (ICs), owing to IPL's low system cost and high throughput capability. IMS has been making this claim for many years, but recently SEMATECH has conducted an independent cost analysis which confirms the IMS claims:[7] IPL has the lowest, potential, cost of ownership (COO) of all of the competing lithographic technologies.

Figure 1 shows a cost analysis performed by the Microelectronics Research Laboratory which closely matches the SEMATECH COO[7] study. This figure gives the estimated costs of performing a complete lithographic step (coat, bake, expose, bake, develop, inspect, etc.) on a 300 mm diameter wafer versus the number of wafers produced per mask. Some of the input parameters are given in Table I. Since lithography represent 30%–50% of the wafer fabrication costs, lithography COO has a large impact on the economics of IC production. Figure 2 estimates the cost of producing a wafer of 4 Gbit DRAM devices utilizing each of the contending technologies relative to 193 nm lithography costs. Given the tight margins in DRAM manufacture, IPL technology could give significant competitive advantage.

IPL achieves this low COO by virtue of its high throughput capability and low overall system cost. IPL offers the possibility of raw throughputs of up to 92 wafers per hour, which is more than twice the raw throughput of the next highest competitor. See Fig. 3. (A detailed sensitivity analysis[7] by SEMATECH indicates that wafer throughput is the single largest contributor to COO.) Since the introduction of the stepper in the late 1970s, lithography costs have averaged between $10 and $20 per wafer level, but the lithography costs of the next century will be dramatically higher. Only IPL offers the possibility of keeping lithography production costs within historic trends as seen in Fig. 4.

III. LIMITS TO HIGH THROUGHPUT SYSTEMS

All lithographic technologies strive for high throughputs, but each faces fundamental limits. Ultimately all technologies must deal with power loads to masks, wafers and the optical system, as throughput is directly related to the total energy that can be delivered to the wafer. Scanning systems introduce mechanical limits on stage velocities and accelerations while maintaining high accuracy. Charged particle systems also exhibit space charge effects. Table II summarizes the technologies and their associated limits to throughput. Each of these limits will be compared in the following sections.

A. Heating issues

The exposure time of a lithography tool is a function of the total power delivered to the wafer divided by the resist

[a)]Electronic mail: ilberry@romulus.ncsc.mil

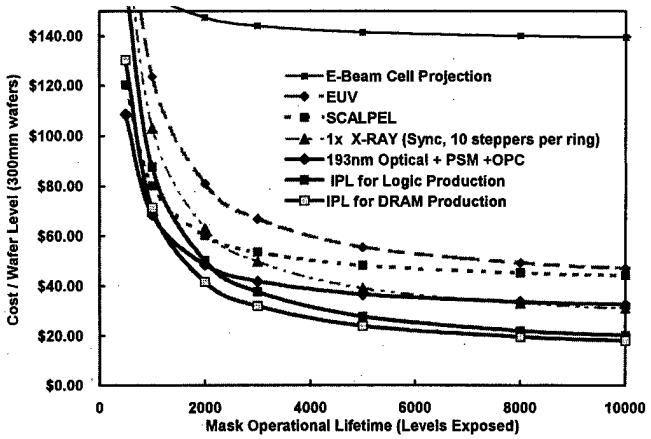


FIG. 1. Comparative cost of lithography per wafer level vs the operational mask lifetime in exposed levels for the 0.13 μm generation of devices. As mask production costs increase dramatically for all of the advanced technologies, the mask lifetime expressed in number of wafers produced per mask has a large affect on the overall IC production costs. SEMATECH estimates that microprocessor products average about 3000 wafers per mask, while DRAM averages about 5000. Input data is taken from the SEMATECH cost of ownership analysis (Ref. 7). IPL for DRAM production assumes 30% of critical levels require complementary masks, while IPL for Logic assumes 50%.

sensitivity. (Generally resist sensitivity ranges between about 10–100 mJ/cm^2 for all technologies. While slight improvements in resist sensitivity may be achieved, one generally pays a price in resolution as sensitivity is increased.) Scanning systems for example require very high power densities to achieve reasonable throughputs due to the small exposure areas. These high power densities lead to wafer and mask heating which limit the ability to overlay patterns onto the wafer. Additionally excessive heating of resists leads to variation in critical dimension control and promotes resist mass loss and outgassing.

SCALPEL has dramatically reduced mask heating effects over that which limited earlier scanning e-beam projection systems by using a scattering mask where only about 0.1% of the incident energy is absorbed. It, however, maintains as a high wafer power load, causing peak temperatures to rise as high as about 50 °C and mean temperature rise of about 5 °C.

EUV has optical elements which absorb about 30% of the power, causing elements near the source to heat.

In x-ray lithography, the scanning x-ray beam heats the mask as it sweeps by, causing an estimated 20–30 nm of mask distortion. Attempts to reduce this effect, like helium cooling, are employed. Membrane heating, however, remains a component of the overlay error that "Product Specific Emulation" [8] attempts to remove.

IPL being a full field exposure system, has dramatically reduced power loading density on both the mask and wafer. At full throughput, IPL delivers about 0.75 W of power into 1530 mm^2 of wafer area. In contrast SCAPLEL delivers about 3 W into a 0.063 mm^2 area. Figure 5 compares the power loading of the mask and wafer for the competing technologies.

TABLE I. (a) Input parameters for the cost of ownership calculation. (b) Throughput limits to the advanced lithography technologies running 300 mm diameter wafers.

(a)

Constants	
Capital depreciation (yrs.)	6
Cost of money (% of investment)	8%
Operator cost per man-year	$45,000
Maintenance cost per man-year	$50,000
Tool utilization (exposure field utilization, up-time, etc.)	72%

Technology	Unit cost (M)	Raw throughput (lvl/hr)	Mask cost (K)	Resist and other processing costs
IPL for Logic	$8	74[a]	$50	$6.00
IPL for DRAM	$8	80[b]	$50	$6.00
1× x ray	$6[c]	35	$80[d]	$6.50
SCALPEL	$6	12	$40	$6.50
EUV	$20	40	$85	$11.00
E-beam CP	$12	5	$20[e]	$6.50

(b)

Technology	Max. (est.) throughput (wafers/h)	Limits: Heating	Stage	Space charge	Other
IPL	92/57[f]			X	Shot noise
SCALPEL	12	X	X	X	
EUV	40	X			Optics reflectivity, source power, shot noise
X-Ray	30	X	X		No. of die per mask
E-beam cell projection	2	X		X	

[a]Assumes 50% of levels use single masks at 92 wafers/h and 50% complementary masks at 57 wafers/h.
[b]Assumes 70% of levels use single masks at 92 wafers/h and 30% complementary masks at 57 wafers/h.
[c]Cost for stepper and beam line. Does not include storage ring and injector charge. $35M assumed. Calculation assumes ten steppers per ring. The basic assumption is that a minimum of two rings will be needed for backup. A 20 000 wafer start per month capacity requires about 20 x-ray steppers.
[d]Assumes four chips per mask to achieve the 35 wafers per hour throughput.
[e]Assumes $5K mask cost, plus $15K data/machine prep per design.
[f]Note 92 wafer/h using a single mask and 57 wafers/h using complementary masks.

B. Space charge

Ultimately all charged particle beam systems are limited in total beam current by the space charge interactions.[9] The majority of the space charge effects occur where the current density is the highest; at the source and at the crossovers. The space charge effects can, however, be mitigated by increasing the crossover energy, the crossover angle, and the beam waist.[10] In IPL the crossover energy is 300 keV, and the crossover angle is about 100 mrad. In the case of SCALPEL, for example, the crossover is at 100 keV and the angle is a few milliradians. Additionally the IMS IPL design inten-

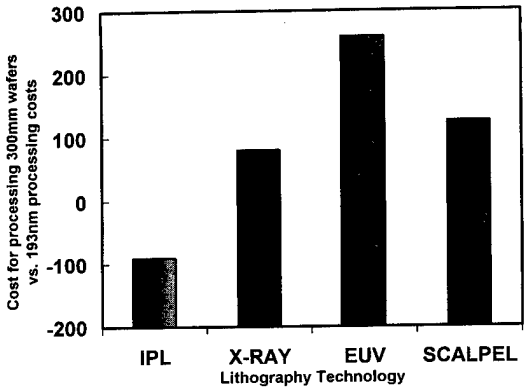


FIG. 2. Cost comparison for manufacture of 4 Gbit DRAMs using each of the competing technologies relative to the costs of fabrication using 193 nm with phase shift masks and optical proximity correction. It assumes that the advanced technology is used for ten critical layers.

tionally aberrates the crossover to increase the beam waist. Increasing the angle and beam waist to the extent done is only possible because extremely small numerical apertures achievable in IPL. Diffraction limits e-beam projection systems to numerical apertures of about 10^{-3}. Figure 6 graphically depicts the field size limits to e-beam projectors, such as SCALPEL, versus the comparatively larger fields capable in ion beam projectors.

C. Mechanical issues; stages

All wafer exposure systems rely on precise mechanical stages for accurately positioning wafers and masks. In reduction scanning systems, such as DUV, 193 nm and SCALPEL, both the mask and wafer are in a constant scanning motion, with the mask moving at four times the wafer velocity to match the 4× image reduction. In SCALPEL, to achieve the estimated 15 wafers/h. throughput the mask stage needs to accelerate at about 4 g, and slew at about 0.5 m/s. This puts severe demands on the stage design and causes concern about the stability of the thin membrane mask.

In 1× x-ray, the wafer stage needs to move beneath the membrane mask and is therefore limited to speed of about 10 cm/s. Given the gap between fast moving wafer and mem-

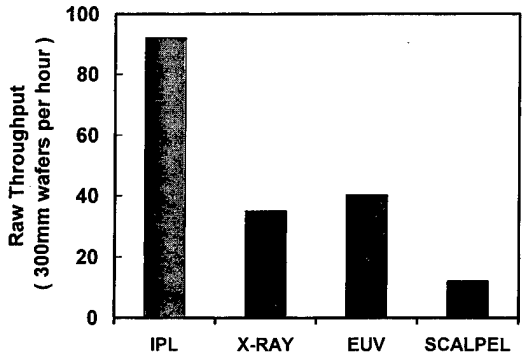


FIG. 3. System throughput capabilities for advanced lithography technologies from SEMATECH cost of ownership study and input from SEMATECH technology champions.

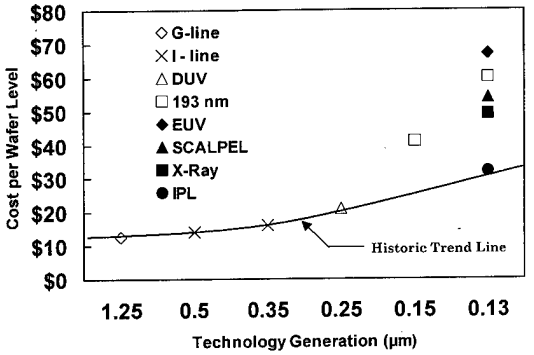


FIG. 4. Historic cost trends in lithography. Lithography of the next century will change the economics of wafer production. Only IPL offers the possibility of remaining on or close to the historic trend line.

TABLE II. (a) Input parameters for the cost of ownership calculation. (b) Throughput limits to the advanced lithography technologies running 300 mm diameter wafers.

(a)

Constants	
Capital depreciation (yrs.)	6
Cost of money (% of investment)	8%
Operator cost per man-year	$45,000
Maintenance cost per man-year	$50,000
Tool utilization (exposure field utilization, up-time, etc.)	72%

Technology	Unit cost (M)	Raw throughput (lvl/hr)	Mask cost (K)	Resist and other processing costs
IPL for Logic	$8	74	$50	$6.00
IPL for DRAM	$8	80	$50	$6.00
1× x ray	$6	35	$80	$6.50
SCALPEL	$6	12	$40	$6.50
EUV	$20	40	$85	$11.00
E-beam CP	$12	5	$20	$6.50

(b)

		Limits			
Technology	Max. (est.) throughput (wafers/h)	Heating	Stage	Space charge	Other
IPL	92/57			X	Shot noise
SCALPEL	12	X	X	X	
EUV	40	X			Optics reflectivity, source power, shot noise
X-ray	35	X	X		No. of die per mask
E-beam cell projection	2	X		X	Shot noise

[a]Does not include storage ring and injector charge; $35M assumed; calculation assumes ten steppers per ring. The basic assumption is that a minimum of two rings will be needed for backup. A 20 000 wafer start per month capacity requires about 20 x-ray steppers.
[b]Assumes four chips per mask to achieve the 35 wafers per hour throughput.
[c]Assumes $5K mask cost, plus $15K data/machine prep per design.
[d]Note 92 wafer/h using a single mask and 57 wafers/h using complementary masks.

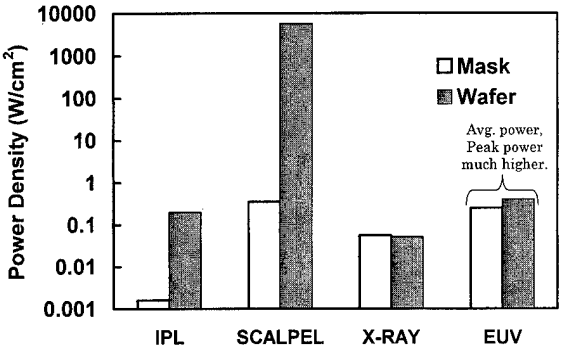


FIG. 5. Comparisons of the power density (absorbed) on the mask and wafer for each of the technologies. IPL is the only full field, nonscanning, exposure system. This offers IPL an advantage in reducing wafer and mask heating effects.

brane mask is on the order of only 10 μm, concern arises over possible pressure differentials between mask and wafer that might occur at higher stage velocities.

In IPL, the mask stage is stationary, except for mask exchange, and the wafer stage moves at about 20 cm/s. Higher wafer stage speeds should be easily achievable, increasing throughput even further.

D. Other throughput limiting issues

1. Shot noise

Owing to the high sensitivity of resist to ions (0.1–2.0 μC/cm^2), concern arises that dimension control may be lost at small feature sizes due to the small number of ions used in each pixel exposure. This effect has been searched for, but has not been noticed even for feature sizes well below 0.1 μm. As an example, well formed lines have been observed with the number of ions per pixel as low as 25–35.[11,12] One explanation is that the emission of ions from a plasma source is not random due to space charge shielding.[10] This would in principle reduce statistical fluctuations beyond what Poisson's statistics would predict. Additionally, recognize that when an ion interacts with resist, its mean-free path for electronic interaction (the mechanism

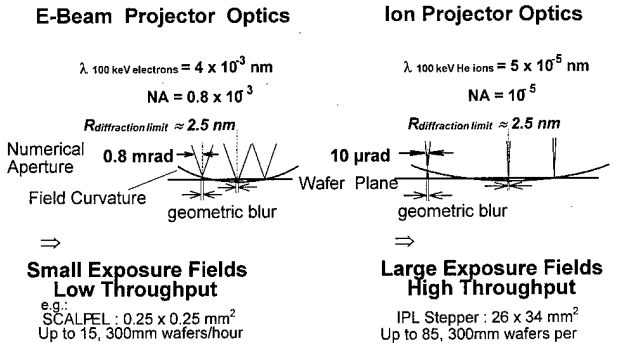


FIG. 6. Comparison between the optics of a full field IPL exposure tool and a small field e-beam projector. The relatively larger numerical aperture needed for e-beam systems increases lens aberrations and field curvature. This reduces the useable exposure field size resulting in a lower system throughput.

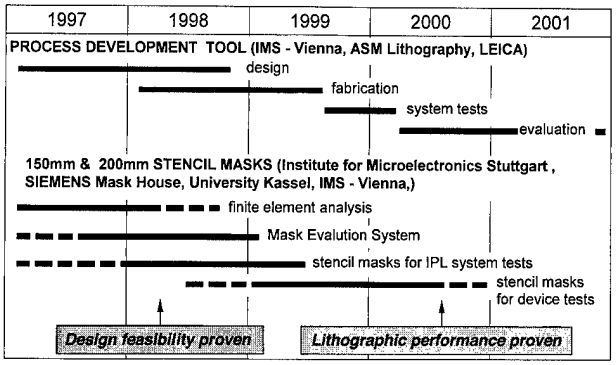


FIG. 7. MEDEA IPL program plan.

which exposes resist[13]) is on the order of the atomic lattice spacing, while for 100 keV electrons it is of the order of a few thousands of angstroms. So at a 10 nm pixel volume, every ion that enters will expose resist, while less than 1 in 10 electrons will contribute to exposure. Shot noise (and resist resolution) may also be an issue in other technologies requiring extremely high sensitivity resists to meet throughput expectations.

2. Number of die per exposure field

One of the components for increasing x-ray throughput up to the 40, 300 mm wafers/hour is to expose more than one die field at a time, possibly as high as 16. This requires writing multiple chips per mask. It is not clear that this is a workable strategy. Current x-ray mask production experiences significant yield losses in fabricating single die masks. There is concern that fabricating four or more die per mask will result in extremely high mask costs, low yield and high turn-around time.

IV. MEDEA IPL EFFORT

The MEDEA effort began in September 1997, as a ~$36M, 160 man-year effort to assess the viability of IPL technology. A goal is to construct a process development tool and mask manufacturing capability commensurate with a pilot production demonstration of 1 Gbit DRAMs by 2001, and production capable machines by 2003. See Fig. 7. The program consists of companies, universities and Institutes in Germany, Austria, and the Netherlands. Running in parallel is a collaborative U.S. effort in IPL stencil mask making. The program organization and members are shown in Fig. 8. This program has chosen to scrap the previous Advanced Lithography Group design[14] for a 3:1 reduction vertical machine in favor of a 4:1 reduction horizontal tool. In addition the design will utilize multielectrode element stacks that have been successfully demonstrated by IMS in 1996. A drawing depicting the tool is shown in Fig. 9.

V. RISKS

One cannot talk about the benefits without addressing the risks. The most significant risk, as it is for all technologies, in achieving the needed image placement and overlay. The

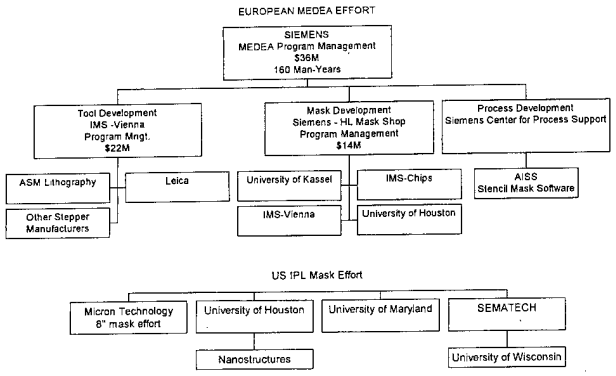


FIG. 8. Organization of the MEDEA IPL effort and the collaborative U.S. mask effort.

largest factor of which are the mask contributions. Of prime concern is the stress relief that occurs when the stencil patterns are formed in the membrane. The IPL strategy is to utilize extremely low stress (1–10 MPa) membranes combined with stress relief structures external to the chip fields. Studies are underway in Europe and the United States to assess the magnitude of this issue. Preliminary results look very encouraging but additional data is needed and is planned to be collected during 1998. The Institute for Microelectronics-Stuttgart and the Siemens mask house have begun the manufacture of test stencils on 6 in. wafers. See Fig. 10.

Other risks include: the ability or inability to maintain mask cleanliness without pellicles, ion optical column performance, and space charge/shot noise limits to throughput.

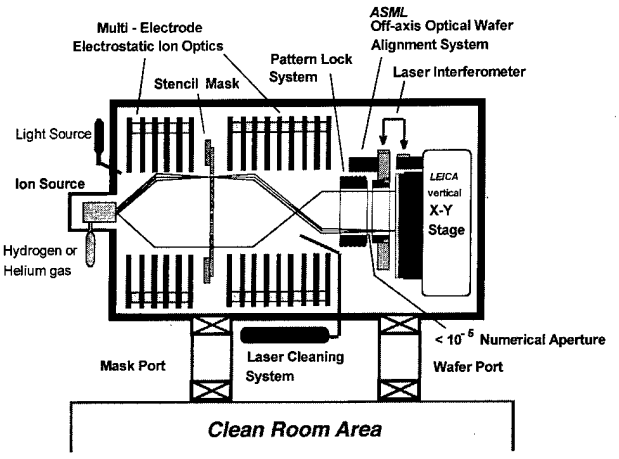


FIG. 9. MEDEA IPL process development tool. This tool will feature a 22 × 22 mm exposure field, with an exposing beam voltage of 75 kV. It will include active mask temperature stabilization and an *in situ* mask laser cleaning system. It will be installed at the Institute for Microelectronics-Stuttgart.

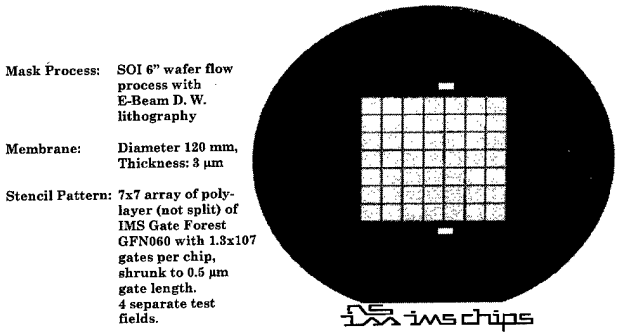


FIG. 10. First 6 in. IPL stencil mask fabricated at the Institute for Microelectronics-Stuttgart in cooperation with the Siemens Mask House (1997).

VI. CONCLUSIONS

All competing next generation lithography technologies strive to improve their user cost of ownership, but are ultimately constrained by fundamental limits. IPL is unique among them in that it is capable of maintaining lithography costs consistent with historic cost trends. While it is difficult to accurately assess the degree of difficulty in achieving the requisite system performance. It is clear that the potential economic advantages justify continued IPL technology development.

[1]K. Brown, Next Generation Lithography Meeting Press Release, 7 November 1997.
[2]G. Stengl, H. Löschner, W. Maurer, and P. Wolf, J. Vac. Sci. Technol. B **4**, 194 (1986).
[3]W. Fallmann, F. Paschke, G. Stangl, L.-M. Buchmann, A. Heuberger, A. Chalupka, J. Fegerl, R. Fischer, H. Löschner, L. Malek, R. Nowak, G. Stengl, C. Traher, and P. Wolf, AEU Electron. Commun. **44**, 208 (1990).
[4]G. Stengl, G. Bosch, A. Chalupka, J. Fegerl, R. Fischer, G. Lammer, H. Löschner, L. Malek, R. Nowak, C. Traher, and P. Wolf, J. Vac. Sci. Technol. B **10**, 2838 (1992).
[5]G. Stengl, G. Bosch, A. Chalupka, J. Fegerl, R. Fischer, G. Lammer, H. Löschner, L. Malek, R. Nowak, C. Traher, and P. Wolf, Microelectron. Eng. **21**, 187 (1993).
[6]MEDEA—Microelectronics Development for European Applications.
[7]J. Canning, S. Mackay, P. Seidel, and W. Trybula, 2nd International High Throughput, Charged Particle Lithography Workshop, Kauai Hawaii, August 1997 (unpublished).
[8]Product specific emulation is a process employed by IBM that corrects for systematic overlay errors such as mask/wafer heating or systematic components of e-beam writing errors by writing a mask, measuring the resultant overlay error, then writing another mask with corrections to offset the previously measured errors.
[9]G. H. Jansen, *Coulomb Interactions in Particle Beams* (Academic, New York, 1990).
[10]J. Melngailis, A. A. Mondelli, I. L. Berry, and R. Mohondro, J. Vac. Sci. Technol. B (to be published).
[11]S. Matsui, K. Mori, K. Sago, T. Shiokawa, K. Toyoda, and S. Namba, J. Vac. Sci. Technol. B **4**, 845 (1986).
[12]R. L. Kubena, F. P. Stratton, J. Ward, G. M. Atkinson, and R. J. Joyce, J. Vac. Sci. Technol. B **7**, 1798 (1989).
[13]G. M. Mladenov and B. Emmoth, Appl. Phys. Lett. **38**, 1000 (1981).
[14]H. Löschner, G. Stengl, I. L. Berry, J. N. Randall, J. C. Wolfe, W. Finkelstein, R. W. Hill, J. Melngailis, L. R. Harriott, W. Brünger, and L.-M. Buchmann, Microlithography World **3**, 6 (1994).

Microelectromechanical tunneling sensor fabrication and post-processing characterization using focused ion beams

F. P. Stratton,[a] R. L. Kubena, H. H. McNulty, R. J. Joyce, and J. Vajo
Hughes Research Laboratories (HRL), Malibu, California 90265

(Received 17 April 1998; accepted 28 May 1998)

Focused ion beams (FIBs) have been previously used as tools for such diverse tasks as high-resolution lithography, where their sub-10 nm spot sizes enable the patterning of diverse nanostructures, and surface compositional analysis, where their ability to sputter material in a localized area allows discrete components of a device or circuit to be characterized. Recently, the authors have capitalized on the FIB's versatility by using it for microelectromechanical sensor fabrication as well as post-processing device characterization. The HRL FIB nanoprobe system has been used in the fabrication of high-performance surface-micromachined accelerometers operating on the principle of tunneling between a cantilevered beam and a sub-0.1-μm-diam tip lying beneath it on a Si substrate. The 8-nm-diam FIB has been used to pattern small dots in a bilevel negative-positive resist layer which are then transferred into a Au layer to form pyramid-shaped tunneling tip structures whose narrow dimensions are essential to high device performance and stability. High-purity, contamination-free Au on both the tunneling tip and cantilever underside is also critical to high-sensitivity tunneling devices. Because the undersides of the beams cannot be viewed with a scanning electron microscopy, even at high mechanical tilt angles, the cantilevers must be physically peeled back in order to expose their bottom surfaces and analyze them for cleanliness. Depending on the material used for fabricating the cantilevers, the rigidity of the structures can render them difficult to bend. We have used a commercial FIB milling system to cut through a portion of the cantilever width, thus creating a hinge, which facilitates the subsequent peeling back of the structure. Comparison of Auger spectroscopy data on peeled-back beams with and without a FIB-milled hinge shows similar surface contamination levels, indicating that redeposited material due to ion milling is localized enough to not affect the compositional analysis of the tunneling region. © *1998 American Vacuum Society.* [S0734-211X(98)11604-9]

I. INTRODUCTION

Focused ion beams have shown promise as high-resolution lithographic tools in device prototyping applications where throughput is not a high-priority issue. In particular, their fast writing times and low proximity effects have been exploited for tasks such as subtractive patterning of high atomic mass materials for potential use as x-ray lithography masks[1] and self-aligned high electron mobility transistor (HEMT) gate patterning.[2] We now demonstrate an application involving the fabrication of sub-0.1 μm tunneling tip structures contributing to greater device performance and stability.

In our accelerometer design, a free-standing cantilever, biased at roughly 100 mV relative to a tunneling structure lying beneath it, is electrostatically deflected downward in the vertical direction by applying a larger (30 V) bias voltage to a control electrode located underneath the beam closer to its anchor point, as shown in Fig. 1. When the gap, denoted by h, between it and the tunneling tip is small enough, i.e., about 1 nm, a tunneling current begins to flow. The current is highly sensitive to changes in the gap and is kept constant by means of external servo electronics. Any acceleration sensed in the vertical direction is compensated for by a corresponding change in control voltage in order to maintain the constant tunneling current. By measuring the change in control voltage due to a physical perturbation in the vertical direction, the acceleration can be calculated. However, the performance of the device depends on the stability of the tunneling characteristic, which is compromised if the force of attraction between the cantilever and the tip is large enough to cause the two surfaces to contact one another.

This force can be modeled as that between a large surface (cantilever underside) and a flat-topped conical structure (tunneling tip), whose roughness is assumed to be no greater than 1–2 nm. The critical radius of the tip is defined as that above which the cantilever will collapse onto it due to the attractive force between the two structures. Figure 2 shows the estimated critical radii of conical, flat-topped tunneling tips versus bias voltage for different cantilever lengths and widths.[3] For our longer cantilever structures with higher displacement sensitivities, the critical radius decreases, since less applied voltage will be required to deflect the beams into tunneling position. Less attractive force between the tip and the cantilever will be necessary to induce the two surfaces into contact. For these devices, assuming a cantilever bias voltage of 100 mV, the fabrication of tips with radii in the 10–50 nm range is essential.

Optical stepper lithography provides a high-throughput, reproducible means of fabricating tunneling structures, and the pattern placement accuracy is very high. However, using

[a]Electronic mail: fp.stratton@hrl.com

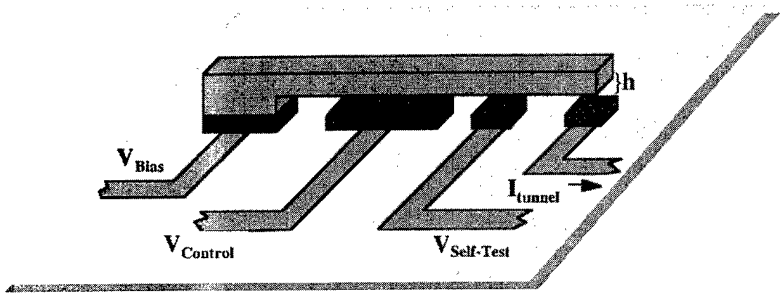


FIG. 1. Surface-micromachined tunneling sensor.

our Ultratech 110 Optical Stepper, the pattern resolution is limited to about 1.0 μm due to limitations in the mask hole size that can be created, and when combined with a metallization and liftoff process, conical structures with flat tops of roughly 0.5 μm radius are produced. Contact aligner optical lithographic techniques, due to the reduced working distance between the optical source, contact mask, and wafer, yield higher-resolution structures with radii of roughly 0.1–0.2 μm (see Fig. 3). Tunneling tips with radii in this range are suitable for our shorter (100-μm-long) cantilevers, but the placement accuracy of the tip is compromised due to mask run-out and misalignment issues. In addition, the larger aspect ratio required to attain the 0.1 μm radius at the top yields structures that are more susceptible to damage during post-process handling and packaging. For sub-0.1 μm radius tunneling tips, a nanolithographic process is required.

II. FIB IN DEVICE FABRICATION

Using our 8 nm, 50 keV focused liquid metal source Ga^+ ion Nanoprobe beam as the exposure tool, we have devised a reproducible, high-placement accuracy lithographic process to fabricate pyramid-shaped tunneling structures with radii of between 10 and 20 nm at their peaks. The fabrication sequence is shown in Fig. 4. The FIB is used to expose a bi-level photoresist made up of a thin 40 nm imaging layer of PPSQ (polyphenylsilsesquioxane) and a thicker 150 nm positive AZ 1350B resist layer. The exposure consists of a single pixel and the dose is roughly 6×10^{12} ions/cm^2. The negative resist layer is wet developed so only the exposed resist dot remains. Using the PPSQ cap as a mask, a highly selective, anisotropic O_2 reactive ion etch of the underlying positive resist layer transfers the tip pattern down to the metal (Au) region. Next, a flood Ar^+ ion mill of a portion of the Au layer forms the tunneling tip in a pyramid-shaped profile. A final isotropic O_2 plasma etch removes the resist cap and reveals a pyramid with a height of roughly 0.2 μm, as shown in Fig. 5.[4] The width of the structure at its base is also 0.2 μm. However, the radius at the apex of the pyramid is estimated at about 10 nm, a roughly tenfold decrease in radial size over the typical dimension achieved using standard contact mask optical lithography. The geometry of the FIB-defined structure results in a more resilient structure, and the reproducibility of the dimensions from site to site on a given wafer is superior to that of the contact mask optical lithographic process. Based upon scanning electron microscopy (SEM) observation of the tips on tens of adjacent sites on a given wafer, the radial size variation of tips fabricated by FIB lithography is roughly ±5 nm, as compared with ±50 nm for contact aligner optical lithography tips.

III. FIB IN POST-FABRICATION DEVICE CHARACTERIZATION

Another issue critical to the performance of our tunneling sensors is the metal purity and cleanliness of the tunneling surfaces. A high-performance tunneling device must have atomically pure metals free of organic contamination. Two or three monolayers of surface contamination are sufficient to degrade the device performance. It is important to assess

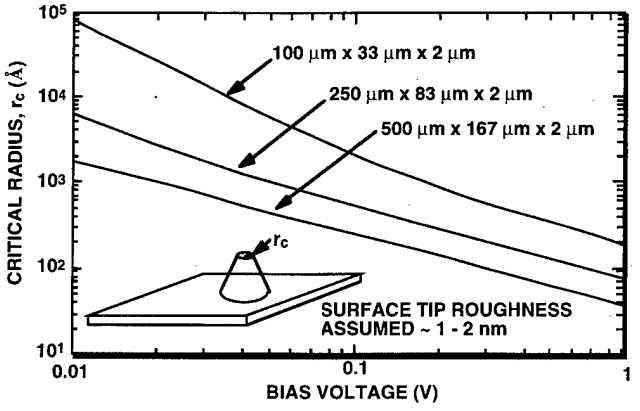


FIG. 2. Critical radius vs bias voltage for different cantilever lengths.

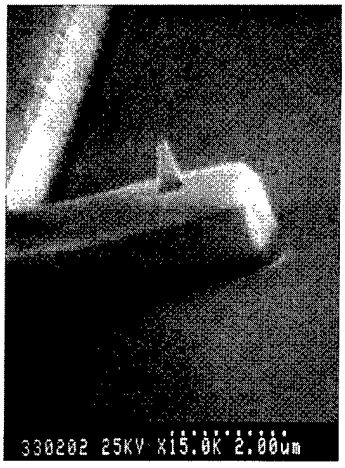


FIG. 3. Tunneling tip fabrication sequence using FIB nanolithograph.

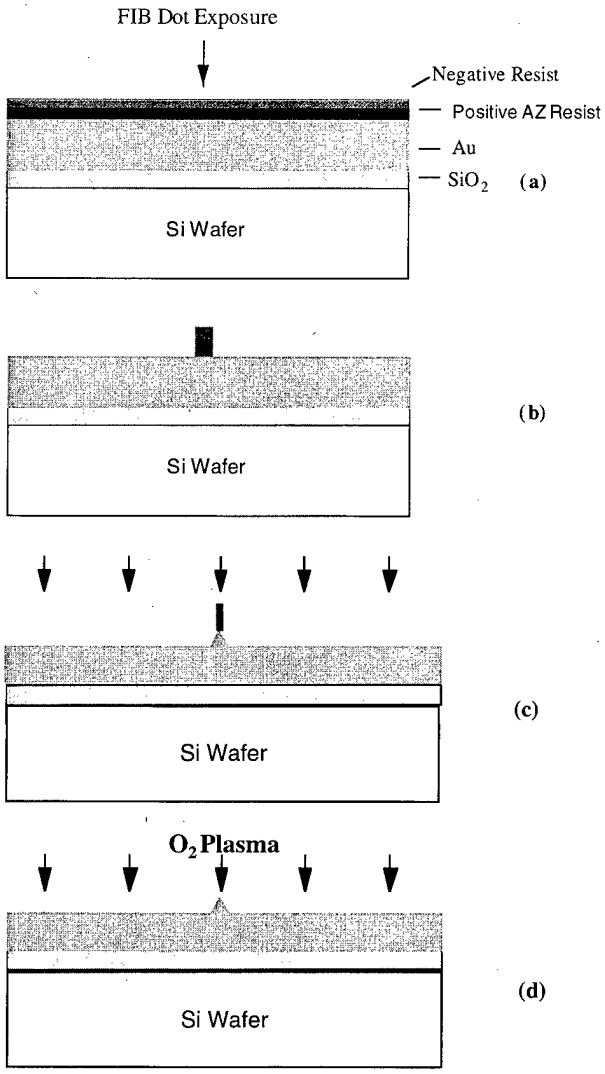


FIG. 4. Tunneling tip fabricated with optical lithography.

the carbon and oxygen levels at the surface of not only the tip region, but also of the underside of our metal cantilever structures. Both areas are concealed from view as seen through an optical microscope, and although the tip region can be accessed by mechanical tilting of an SEM stage, thus

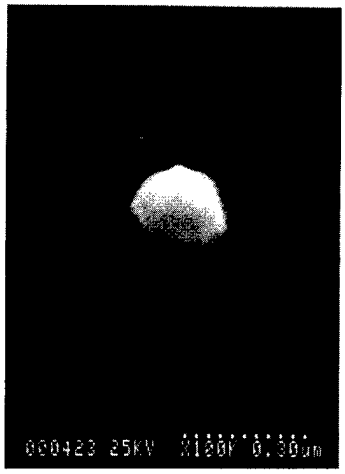


FIG. 5. Tunneling tip fabricated with FIB nanolithography.

permitting Auger spectroscopy or energy dispersive x-ray (EDX) characterization of its surface, analysis of the underside of the cantilever requires additional preparation. To carry out our surface analysis, we must gain access to the undersides of the cantilevers. In order to mechanically peel back the cantilever, thus exposing its underside, the FIB's localized sputtering capability is utilized to create a hinge near the anchor point of the beam, thus rendering the structure more flexible. We are currently constructing our cantilevers out of Au or Ni, and results for Au cantilevers will be presented. For Ni cantilevers, stiffer than their Au counterparts, the FIB milling process is even more desirable.

Figure 6 shows a sequence of SEM micrographs of a Au cantilever in various stages of preparation for surface compositional analysis. The beams for this experiment are 100 μm long and 33 μm wide. The thickness of the Au is roughly 2 μm. A 25 keV Ga^+ FIB is used to mill a slot through the cantilever metal covering roughly two-thirds of the width of the cantilever structure, as shown in Fig. 6(b) (for a 100-μm-long and 33-μm-wide beam, a 4 μm$\times$23 μm area is FIB milled). Roughly 700 scans of the area are necessary to mill through the cantilever into the underlying substrate. The beam current is 1 nA, and the total milling time is

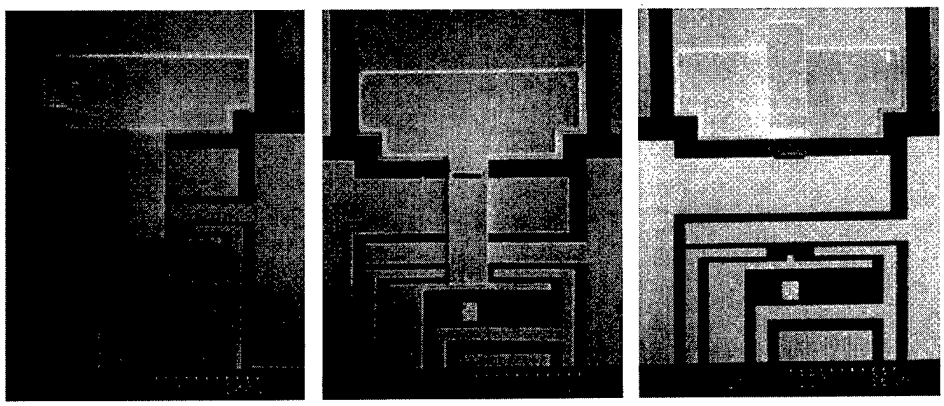
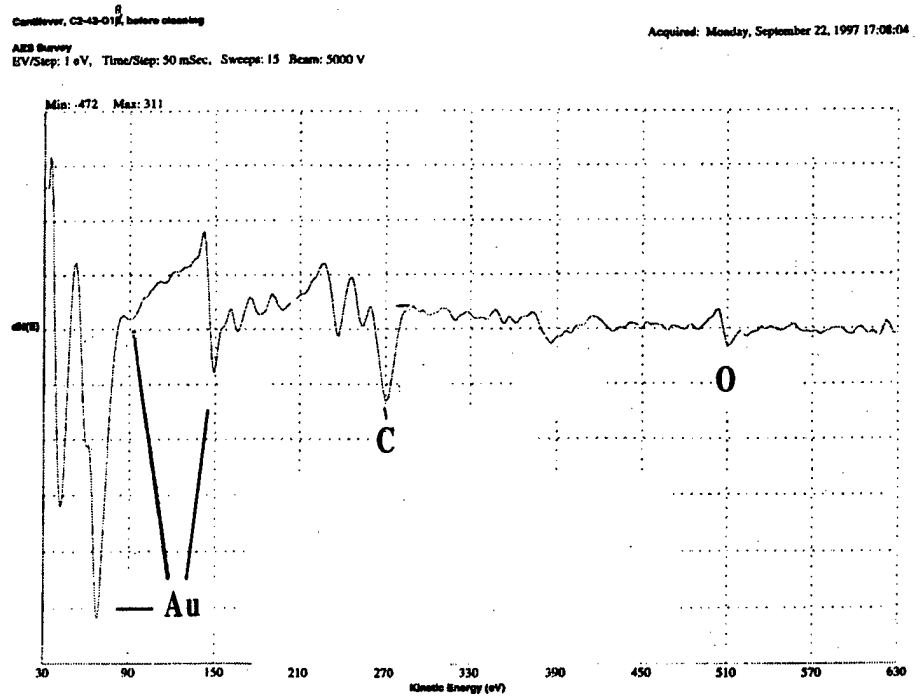
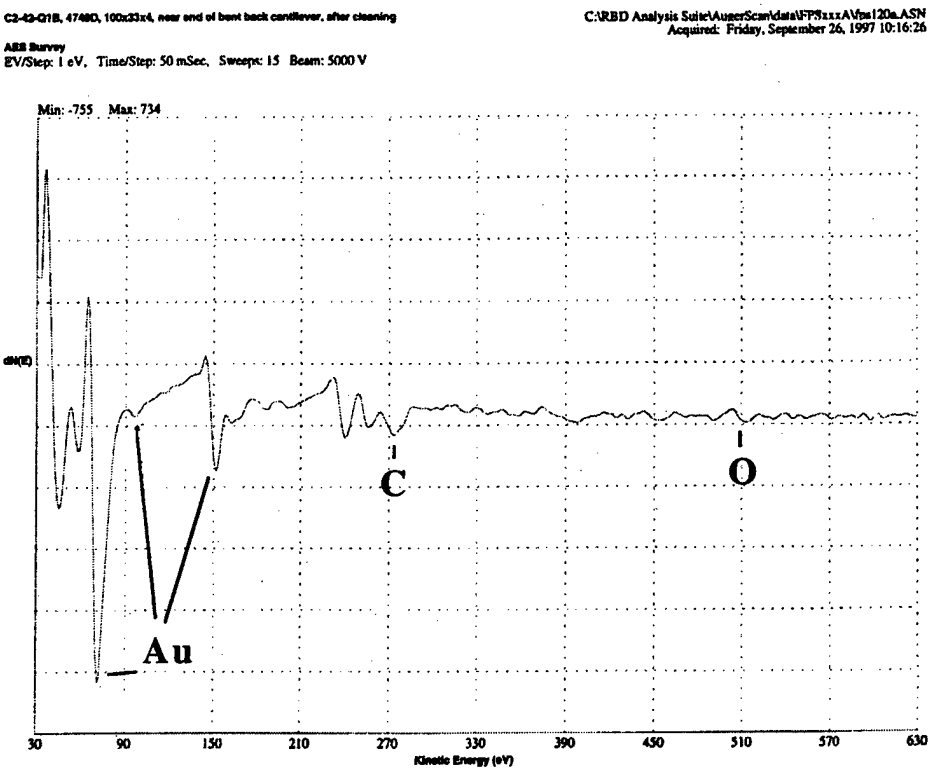


FIG. 6. (a) Au cantilever tunneling sensor as processed. (b) Au cantilever with FIB-milled hinge. (c) Au cantilever mechanically peeled back to expose underside.

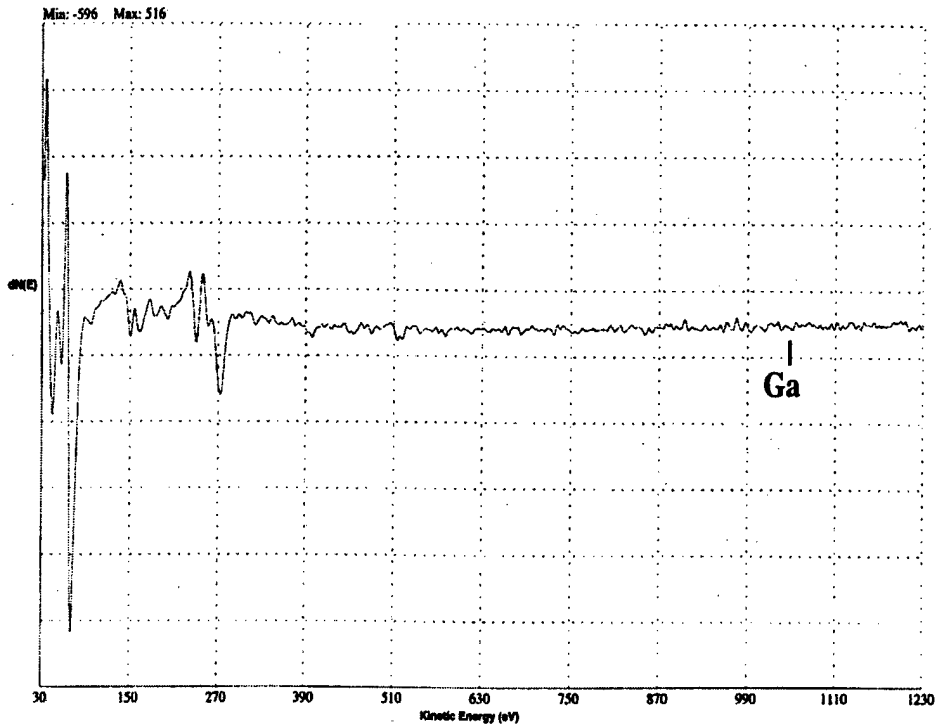


(a)

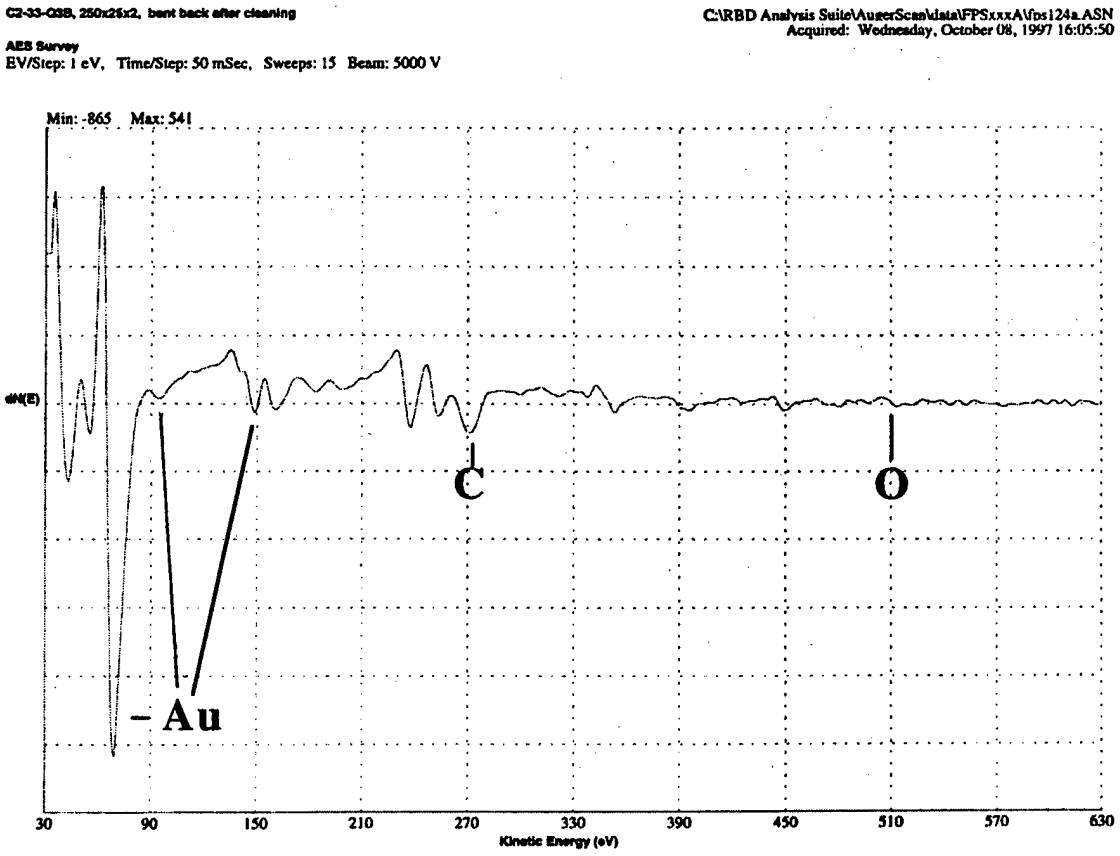


(b)

FIG. 7. (a) Auger spectroscopy data of cantilever underside before cleaning (with FIB-milled hinge). (b) Auger spectroscopy data of cantilever underside after cleaning (with FIB-milled hinge). (c) Auger spectroscopy data (out to 1100 eV) of cantilever underside after cleaning (with FIB-milled hinge). (d) Auger spectroscopy data of cantilever underside after cleaning (no FIB-milled hinge).



(c)



(d)

FIG. 7. (*Continued.*)

10 min. The cantilever can then be mechanically bent back such that it lays flat on its "back" [see Fig. 6(c)], thus permitting its underside to be analyzed.

A key concern of using the FIB milling technique is the effect of any redeposition of sputtered material such as Si or SiO_2, located in the substrate upon which the cantilever is built, as well as Ga from the ion source deposited on the beam. Because Ga oxidizes readily, even a relatively small quantity could invalidate the measured oxygen level. Figure 7(a) shows the Auger spectroscopy data of the composition of a cantilever underside surface before cleaning of the device wafer. The cantilever has been physically peeled back after having been "hinged" with the FIB. The contamination levels are typical of those observed on a processed device which has been exposed to air for a prolonged period. There is a significant carbon level as well as smaller traces of oxygen and nitrogen, the three of which comprise roughly 2.5 monolayers of contamination. Figure 7(b) shows Auger data of another beam, FIB milled and bent back after cleaning of the same device wafer. The carbon and oxygen levels on the surface have been significantly reduced, and the nitrogen peak is no longer visible (a device tested in air immediately after the cleaning step will exhibit improved performance). Also, no major Si peak at 90 eV is observed (the small peak in the vicinity of 90 eV is a secondary Au peak), indicating that any redeposited substrate material due to sputtering has not reached the region underneath the beam where the analysis is being conducted. Figure 7(c) shows Auger data of the same cantilever extending to 1100 eV to show the Ga level, which is in the noise level of the measurement, indicating that the FIB itself is not depositing any contamination in the region under analysis. The noise levels of the measurement correspond to roughly one-tenth of a monolayer of surface film, which is indicative of an almost atomically clean metal surface. The very low oxygen level in the previous graph further substantiates the absence of Ga, since the latter element readily forms an oxide film at room temperature when exposed to air. Figure 7(d) shows Auger data of a longer (250 μm) Au beam which was physically peeled back with much difficulty without the benefit of a FIB hinge, once again after a device cleaning step. The contamination levels are virtually identical to those in Fig. 7(b), further demonstrating that the FIB facilitates the device preparation for surface compositional analysis without compromising the integrity of the data.

IV. SUMMARY

The FIB has found a place in two different aspects of the fabrication and analysis of surface-micromachined tunneling sensors. Its high resolution has been utilized to fabricate sub-0.1-μm-diam tunneling tips, whose small size is essential to the stable operation of high-displacement sensitivity accelerometers. The combination of the FIB's sub-10 nm spot size and a highly directional reactive ion etch and ion milling procedure result in tunneling structures whose dimensions are unattainable using more conventional optical lithography combined with metallization/liftoff strategies.

The FIB's localized sputtering capability makes it a useful tool for the post-process evaluation of tunneling surface cleanliness, which is essential to optimizing the performance of our devices. By enabling the mechanical deformation of the cantilever to access its underside, contamination levels can be measured, and the effectiveness of tunneling tip and cantilever cleaning strategies can be assessed. Furthermore, the FIB activity does not perturb the analysis of the surface's composition.

ACKNOWLEDGMENTS

The authors would like to thank Norm Goodwin of the Hughes Research Laboratories for his assistance in sample preparation for the Auger analysis of our tunneling devices.

[1]G. M. Atkinson, F. P. Stratton, and R. L. Kubena, J. Vac. Sci. Technol. B **10**, 3104 (1992).
[2]G. M. Atkinson, R. L. Kubena, L. E. Larson, L. D. Nguyen, F. P. Stratton, L. M. Jelloian, M. V. Le, and H. Mcnulty, J. Vac. Sci. Technol. B **9**, 3506 (1991).
[3]R. L. Kubena, G. M. Atkinson, W. P. Robinson, and F. P. Stratton, J. Vac. Sci. Technol. B **14**, 4029 (1996).
[4]R. L. Kubena, G. M. Atkinson, W. P. Robinson, and F. P. Stratton, J. Vac. Sci. Technol. B **14**, 4030 (1996).

Multipurpose nanobeam source with supertip emitter

S. Kalbitzer[a] and A. Knoblauch
Max-Planck-Institut für Kernphysik, Postfach 103980, D-69029 Heidelberg, Germany

(Received 19 November 1997; accepted 9 March 1998)

Protrusions of 1 nm dimensions have been grown on single-crystal tips of tungsten and iridium of about 100 nm size. With positive bias the tip can be operated as an intense ion source and with reversed bias intense electron beams can be obtained. In both cases, specific brightness values of up to 1 TA/cm^2 srd eV have been recorded under optimized conditions. As for gases, tungsten can be used only for less reactive species, such as hydrogen and rare gases, whereas iridium is also useful for aggressive gasses, such as oxygen. Both metals are excellent electron emitters. Thus, a powerful multipurpose source is realized with supertips of tungsten and iridium, as will be demonstrated by a variety of experimental results on their relevant properties. © *1998 American Vacuum Society.* [S0734-211X(98)06104-6]

I. INTRODUCTION

With the advent of nanotechnology, higher demands on intense currents of finely focused charged particle beams have come up. Depending on their charge, mass, and velocity these projectiles are useful for quite a variety of applications for physicochemical modification and analysis of submicron solid structures.

It is obvious that with shrinking patterns all serial processing steps take increasingly longer times to carry out the intended operations. Thus, a reduction of feature size by a factor of 10 means exposure periods increasing by two orders of magnitude. Therefore, attempts of substantially improving the available target currents have accompanied the miniaturization efforts in semiconductor fabrication technology.

Recent developments of gas field-ionization sources (GFIS) have demonstrated that ion beams can be emitted with a spectral brightness of up to 1 TA/cm^2 srd eV.[1] To this end, a special emitter geometry, in shape of a "supertip," has been employed. It is a tiny protrusion of the order of 1 nm on top of a single-crystal base tip.[2–4] In principle, this supertip GFIS, or in short GFIS*, is capable of delivering ions of all gaseous elements and compounds. When tungsten as tip material is used, however, aggressive gases quickly destroy the supertip, whereas ion beams of hydrogen or rare gases can be produced over extended periods of time without deteriorating source properties.[5] The latest development has been to produce similar emitter superstructures from the chemically more inert material iridium. In this way, oxygen beams have successfully been produced without any apparent damage to the iridium supertip even after prolonged operation.[6]

These supertip structures can be used as an efficient field electron emission source, FEES*, as well, just by inverting the operation voltage. This is an important point, as finely focused electron beams have been the standard tool for serial writing of photoresist films. In this way, photomasks for chip production have been fabricated and many other devices for microtechnology applications. Again, with diminishing dimensions electron exposure times will increase at the same pace as ion beam irradiations, so that higher electron current densities are desirable as well.

II. BASIC EXPERIMENTAL PROCEDURES AND THEORETICAL ASPECTS

After electrochemical etching single-crystal metal wire to form a base tip of some 100 nm radius helium ion bombardment is used to generate a supertip on the very top of this structure as described in detail previously.[2] Figure 1 is a schematic presentation of the emission geometry of a "regular" tip and supertip. By using a supertip structure, the total beam current can be concentrated into a half-angle of about 1° instead of 30°, so that the angular current density is enhanced by about three orders of magnitude.[2,3] These properties are quite similar for ion and electron beams, since the electrical field distribution of a tip remains the same for the same geometrical conditions.[7] Our source system furthermore contains a retractable current detector combined with a fluorescent screen and a multichannel plate.[6,15] In this way, reliable estimates of the emission angles of ion and electron beams are facilitated.

For both ion and electron beam generation the basic physical mechanism is electron tunneling. Field ionization, or electron field emission, occurs by electron tunneling from the occupied ground-state energy level of an outside gas atom into an empty energy level of the conduction band of the tip metal [Fig. 2(a)]. Conversely, electron field emission takes place by tunneling of an electron from the inside Fermi sea of the tip metal into the outside continuum of free electron states [Fig. 2(b)]. The corresponding electrical field strengths are about 10 V/nm for gas field ionization, and about 1 V/nm for electron field emission, respectively, reflecting the corresponding differences in tunnel barrier height and width. Note the "tunneling shadow" for the gas ionization process: at too close approach the gas atom can no longer be ionized, because no unoccupied conduction band state is available to match the electron energy level of the atom.

[a]Corresponding author; electronic mail: S.Kalbitzer@mpi-hd.mpg.de

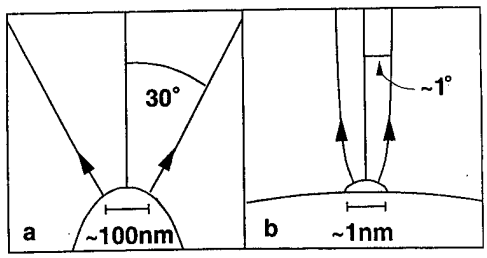


FIG. 1. Emitter geometry of (a) a regular tip of radius $r \sim 100$ nm, and (b) a supertip of radius $r \sim 1$ nm on top of a regular tip. Note the concomitant reduction of solid angle by about three orders of magnitude.

III. RESULTS AND DISCUSSION

Atomic pictures of typical supertips on tungsten and iridium have been shown in recent publications by us.[6,8] Although the crystal structures of these metals are different, bcc and fcc, respectively, both protrusions are nevertheless of similar geometrical size: lateral and vertical dimensions are about 2–3 nm. Thus, a typical supertip comprises a total of the order of 100 atoms assembled in the shape of a pyramid with a few atoms at the apex position. In this way, a considerable field enhancement is established at this structure with the most important consequence that the emission of charged particles can be restricted to this tiny area. The large residual body of the base tip can be kept inactive; it only serves for gas supply along the surface in the field-ionization mode and as a bulk electron reservoir in the field-emission mode. In both modes, however, the tip has to be kept at cryogenic temperatures for optimum performance. With ion beams, this temperature depends on the physical properties of the respective gas species,[6,9] whereas with electron beams, it is imperative to immobilize impurity adsorbates at the surface of the base tip, as will be shown below.[6,10]

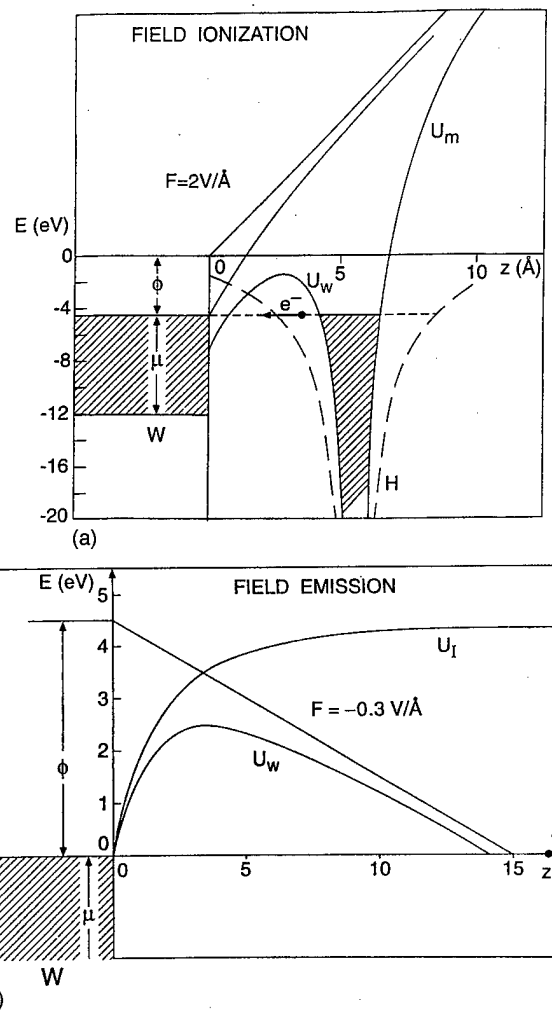


FIG. 2. (a) Field ionization of a H atom near the W surface: the electron tunnels into an empty level in the conduction band of the metal. (b) Field emission of an electron from the conduction band of the metal into an energy level of the vacuum continuum. E = particle energy, z = particle distance from metal surface, F = electrical field, ϕ = metal work function, μ = metal conduction band width, U_I = image potential, U_m = modified atomic potential, U_W = resulting potential wall [adapted from Gomer (see Ref. 19)].

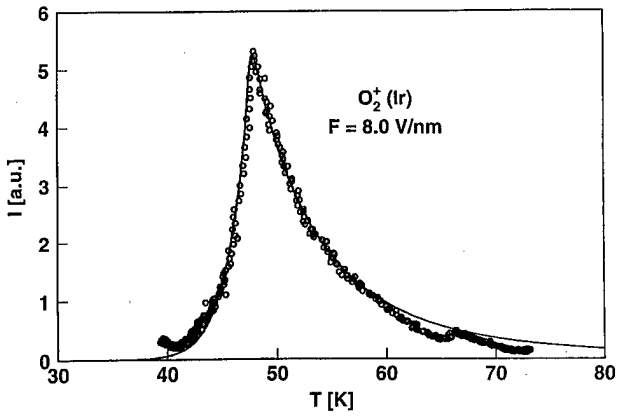


FIG. 3. Oxygen ion beam intensity from an Ir GFIS* as a function of tip temperature at constant electrical field and gas pressure.

Figure 3 demonstrates by the example of oxygen that maximum ion currents are obtained in the vicinity of its condensation temperature, of $T \sim 50$ K. Assuming thermodynamical equilibrium conditions for a gas in a force field, the enhancement of the gas density at the tip reads:

$$\eta = n/n_0 = \exp \Phi, \quad \Phi = \tfrac{1}{2}\alpha F^2/kT, \tag{1}$$

where the enhancement parameter Φ, the ratio of polarization energy and thermal energy, amplifies the actual gas density n over the value n_0 corresponding to regular gas pressure in the remote parts of the source, where the electrical field F is negligible. For a given gas of polarizability α under optimum electrical field conditions, the gas density can further be augmented by reducing the system temperature.

Figure 4 demonstrates the failure of the attempt to produce a useful oxygen beam with a tungsten supertip. Violent fluctuations in emission intensity indicate destructive processes which eventually are seen to lead to a vanishing current level within a few minutes.

An iridium supertip, however, withstands the exposure to oxygen without deterioration, as Fig. 5 shows. The current fluctuations are due to problems of controlling the oxygen pressure, but not to destructive processes as indicated by the remaining average current level. We emphasize that current levels of the order of 0.1 nA O_2^+ into a beam spot of about

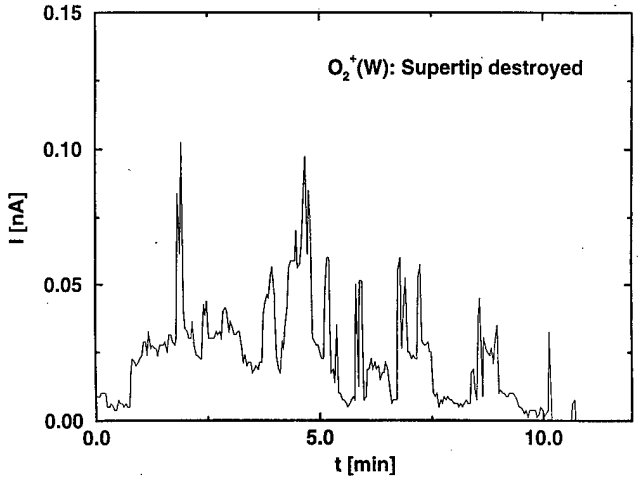


FIG. 4. Unstable oxygen ion beam emission from a W GFIS*.

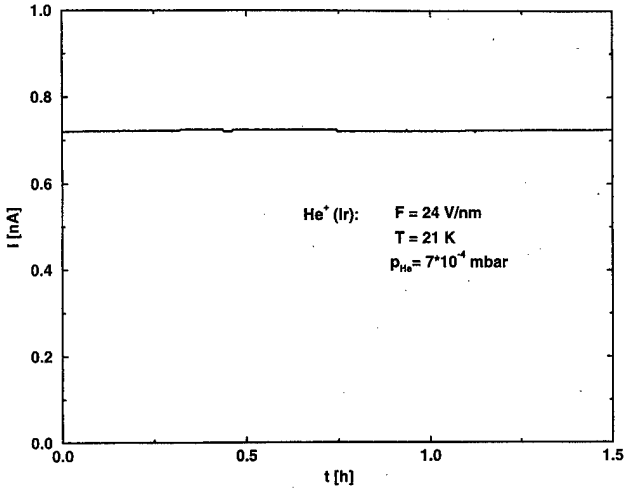


FIG. 6. Stability of a helium ion beam from an Ir GFIS*.

100 nm would be an outstanding improvement to the presently encountered conditions for secondary ion mass spectroscopy (SIMS). Here either O_2^+ beams of low brightness are produced with plasma sources or Ga^+ beams from liquid metal sources with low ionization yield of secondaries are in use.

A supertip source has to be operated at optimum electrical field strengths for maximum current stability, as Fig. 6 demonstrates by the case of a helium beam from an Ir GFIS* with fluctuations in the few percent range. Also, a minimum energy spread of ion beams is achieved by the same measure as shown in Fig. 7, where we see that low fields can cause sizable energy widths leading to inferior imaging properties. Figure 8 indicates that a substantially increased pressure, producing higher current levels in proportion, has to be traded in for some additional beam energy spread, e.g., from $\Delta E \sim 1$ eV at $p \sim 10^{-5}$ mbar to $\Delta E \sim 1.5$ eV at $p \sim 10^{-3}$ mbar.[11] Figure 9 is a more detailed correlation plot of ion current stability and operation point in the current-field characteristic.[6] At the optimum field strength, the ionization probability approaches $P \approx 1$, so that all atoms delivered to the supertip are emitted as ions "instantaneously," i.e., the emission is supply limited. Furthermore, the existence of an optimum gas pressure for a given electrical field, as shown by Fig. 8, points into the same direction: gas pile-up problems have to be avoided, as they cause both beam current fluctuations and energy broadening. In Fig. 9, at the best-current field, located near to the onset of the plateau and below the best-image field position,[4] we observe a very stable current, whereas at lower fields considerable beam intensity fluctuations can be seen. As discussed elsewhere,[12] at the best-current field the supply-limited regime prevails, where the neutral current of atoms drifting/diffusing along the tip shaft to the supertip equals the charged current of gas ions emitted from the supertip, so that the continuity condition is fulfilled. The rear base-tip region, where most of the gas collecting area is located, is supplied with a flow of atoms from the gas phase enhanced by the local electrical field. Supply-limitation theories[13,14] predict an enhancement of gas particle flux to the tip by a linear to sublinear dependence on the enhancement parameter Φ. The gas kinetic flux towards a hyperboloidal tip is enhanced by the function:[14]

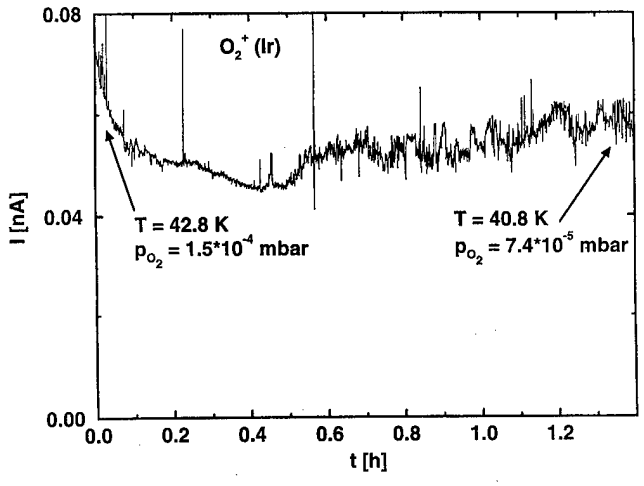


FIG. 5. Stable oxygen ion beam emission from an Ir GFIS*.

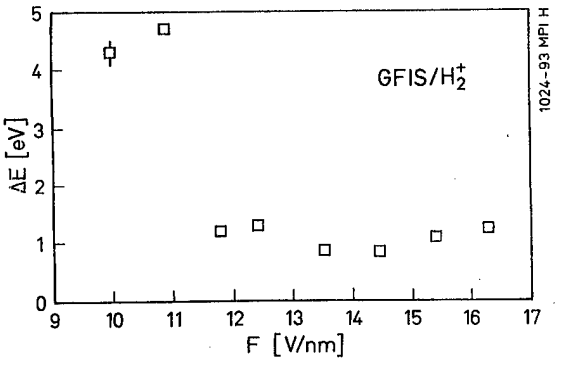


FIG. 7. Energy width of a hydrogen molecular ion beam from a W GFIS* as a function of electrical field strength.

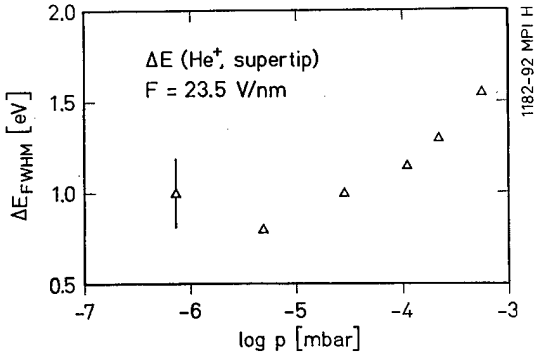


FIG. 8. Energy width of a helium ion beam from a W GFIS* as a function of gas pressure.

$$\sigma(\Phi)=\tfrac{1}{4}(\Phi+2.7\Phi^{2/3}+2.7\Phi^{1/3}+1). \quad (2)$$

All our experimental results, e.g., the data of Fig. 5 obtained for varying temperature at constant electrical field and gas pressure, however, point to an exponential density enhancement according to Eq. (1), so that a current dependence of $I_D \propto n = \eta n_0$ exists similar to the ionization limited regime, in which the disturbance of the gas density is negligible. A detailed analysis of these current–temperature characteristics for different gases agrees with their thermodynamic properties of gaseous and liquid phases, so that also the corresponding peak positions can be explained.[6] Thus, the conclusion is that the supply mechanism is different for supertips than for regular tips, probably due to other conditions for accommodation and transport of the gas kinetic flux.

Finally, Fig. 10 is a demonstration of long-time stability of a neon current emitted from a W GFIS*. Since neon, much heavier than helium, is an excellent choice for applications in materials modification by sputter action, it is essential to know its behavior under practical conditions of extended operational periods. Current stability is seen to be sufficient over a period of 2 days. Moreover, field ion micrographs at the indicated times have shown little or no structural changes in the atomic arrangements of several supertips. This positive result, may be not so surprising, is nevertheless of utmost importance for applied purposes.

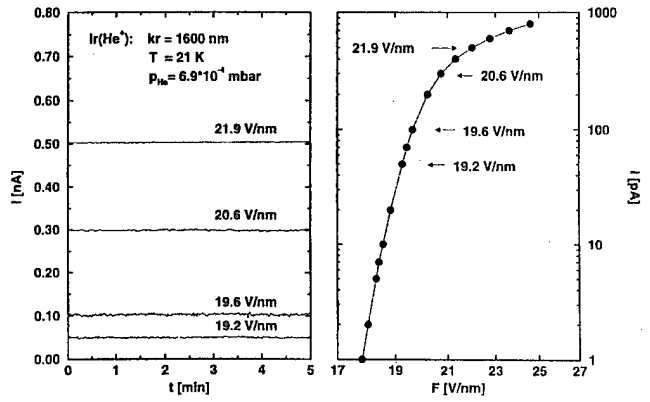


FIG. 9. Helium ion beam from an Ir GFIS*. Left: Beam fluctuations with time at different electrical field strengths. Right: Current-field characteristic.

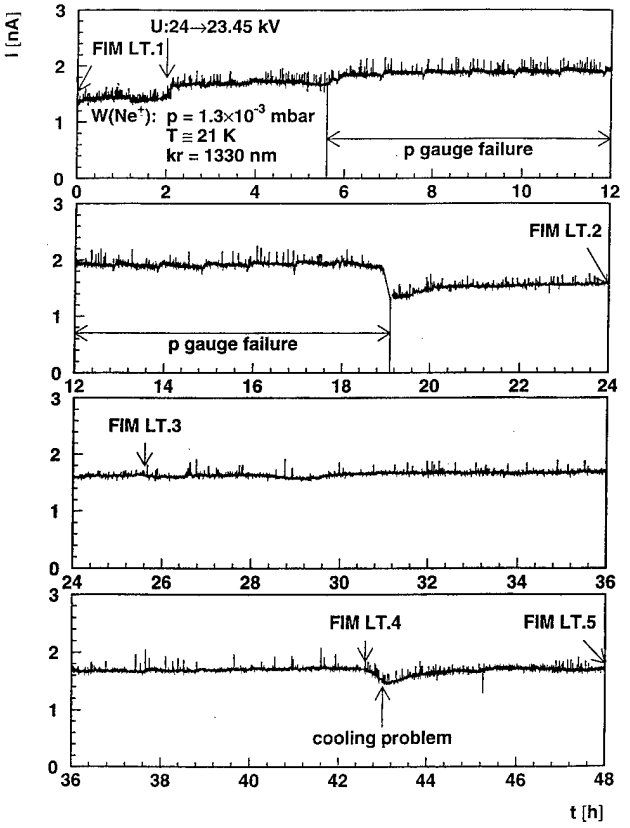


FIG. 10. Long-time stability of a neon beam emitted from a W GFIS*: FIM=field ion micrograph of the long-time run taken at the indicated points $LT.n$, $n=1-5$.

Figure 11 refers to the same system as Fig. 10, but operating at a considerably higher emission current due to different conditions in geometry and gas pressure. We note that the stability of the emitted current is excellent and seems not to depend on its level, if optimum operational parameters are being met.

We now turn to the results obtained for electron beams. Figure 12 demonstrates the excellent stability of electron currents emitted by an Ir EFES*. It has to be cooled to a temperature, where most impurity species adsorbed to the tip are immobilized. This is a very essential point, because mobile impurities can drift to the supertip and cause considerable emission instabilities, as observed, e.g., at room temperature. So far, maximum current levels of about 100 nA have been observed without any optimization of the emitter geometry. Although the requirement of cooling may somewhat impede the incorporation of the EFES* into other equipment, the advantages provided by the excellent stability and high brilliance of the supertip emitter have to be considered. Applications for both material modification, e.g., photoresist exposure, and analysis, e.g., microscopic imaging, appear quite attractive.

Figure 13 is the density profile of an electron beam from a W GFIS*.[10] We see that there is an approximately Gaussian intensity distribution with a half width of about $\frac{1}{2}\alpha \sim 2°$ in lateral directions. This feature has to be accounted for when angular current density figures are quoted.

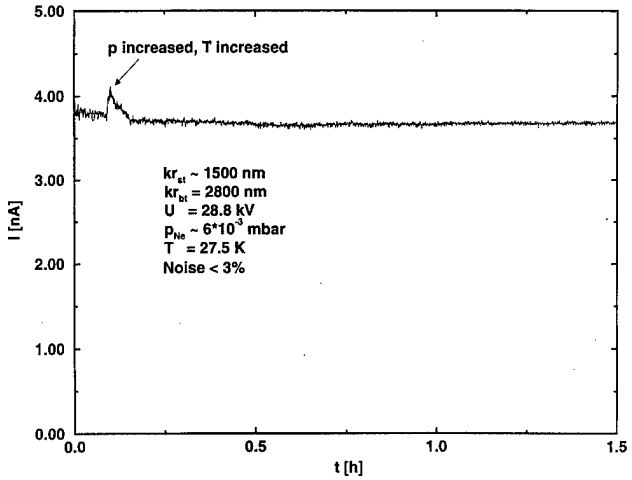


FIG. 11. Neon ion current stability at high intensity from a W GFIS*.

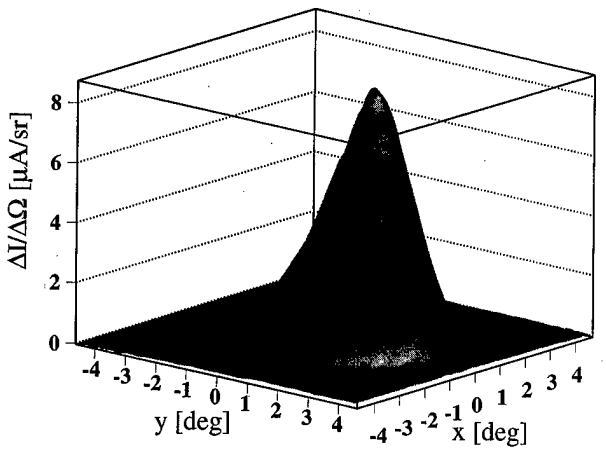


FIG. 13. Electron current-density profile obtained from a W EFES* at 20 K.

The energy spread of an EFES*, as quoted in Tables I and II, can be as low as $\Delta E \sim 0.3$ eV at cryogenic temperatures of $T \sim 20$ K, although with increasing emission intensity values of up to $\Delta E \sim 0.5$ eV were observed but still considerably lower than about $\Delta E \sim 1-2$ eV of a regular EFES.[10] Apparently, the potential energy broadening depends on geometrical factors determining the critical zone of Coulomb interaction between the emitted particles.

IV. PERSPECTIVES

It is of considerable interest to estimate the ultimate image size of both ion and electron beams from a supertip. The main contributions to the total image size are:

$$d_t^2 = (M d_v)^2 + d_c^2 + d_s^2, \tag{3}$$

where the right-hand terms denote virtual source size times magnification, chromatic, and spherical aberrations. In a GFIS* the virtual source size can be neglected, so that the chromatic and spherical terms remain:

$$d_c = C_c \alpha \Delta E / E, \tag{3a}$$

$$d_s = \tfrac{1}{2} \alpha^3 C_s . \tag{3b}$$

While the source energy spread ΔE is important for the chromatic error, the respective aberration coefficients depend on the design of the optical system. The acceptance half angle α of the optical systems can be controlled by apertures, although only at considerable losses of beam intensity, so that the angular current density of the source is one of the most important parameters.

An estimate for a specific lens system to focus 10 nA of ions into a 100 nm image spot was given previously.[15] As in general, the chromatic aberrations become dominant at smaller beam spots. If a nanoscopic application of the GFIS* should be intended, a different optical system must be used. Recent calculations indicate that an optimized lens system can focus charged particles into a target spot of less than 1 nm.[16] If so, imaging applications would become very attractive using both primary electron or ion beams and secondary electrons or ions depending on the actual case to be studied.

Figure 14 is an overview over the present state of the art of ion beam sources, where the image current density is cor-

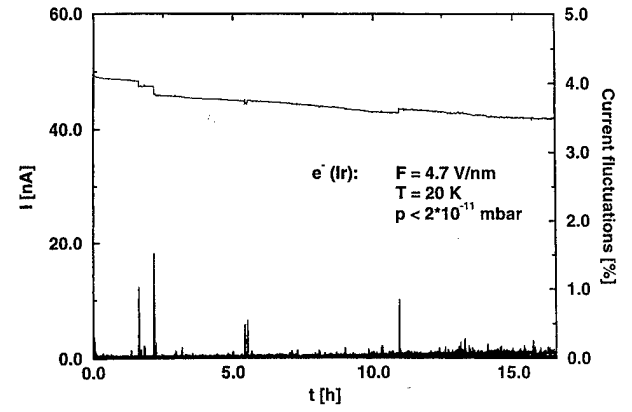


FIG. 12. Stability of an electron beam from an Ir EFES*. Note that the current fluctuations (right ordinate) are about 1% in the local spikes and about 0.1% on the average.

TABLE I. GFIS* performance for charged particles.

Particle species	e^-	H_2^+	He^+	Ne^+	O_2^+
Current (nA) into ±0.5°	100	8	5	5	0.2[a]
Angular current density (μA/sr)	30	35	20	20	1
Energy spread (eV)	0.3	1	1	1	1[b]
Brightness (A/cm² sr)			$10^{10}-10^{12}$[c]		
Spectral brightness (A/cm² sr eV)			$10^{10}-10^{12}$[d]		

[a]Ir supertip, maximum particle current: ~1 nA O^+ (estimate).
[b]Ir supertip, extrapolated value.
[c]Virtual source size: $d_v \sim 0.1$ nm.
[d]Energy spread: $\Delta E \sim 1$ eV (FWHM).
Note: The electron emission half angle is 2°.

TABLE II. Characteristic of electron guns. Thermal emission | Field emission.

	W	LaB_6	ZrO/W	TFE W	CFE W	EFES*
log B (A/cm^2 sr)	~6	~7	~9	~9	~9	~12
d_{tip} (μm)	~50	~10	0.1–1	0.01–0.1	0.01–0.1	~0.003
ΔE (eV)	2.3	1.5	0.7	0.7	0.6	0.3
log p (Pa)	−3	−5	−7	−7	−8	−9
T (K)	2800	1800	1800	1600	300	20
I (μA)[a]	~100	~20	~100	20–100	20–100	>0.1
$\Delta I/I$ (%)	1	1	1	7	5	1
$\Delta I/I$ (%/h)	1	3	1	6	20	<1
Features	···	···	start-up time	build-up process	frequent flashing	initial formation
Operation	simple	simple	easy	easy	complex	easy
Price	low	low	high	high	high	??

[a]Note: The emission half angles for the conventional sources amount to 10–30°, whereas the EFES* typically exhibits about 1°, so that angular current densities scale down by a factor of 100–1000 as referred to EFES*. TFE=thermal field emission, CFE=cold field emission. [Table adopted from JEOL information sheet on electron guns.]

related with the spectral brightness.[1] Plasma sources, gas field ion sources without supertip, liquid metal ion source, and gas field ion sources with supertip are compared with the clear result that at present the GFIS* is the most powerful device for producing high current densities at target position. Note that we have replaced the former conservative estimate of the GFIS* virtual source size of $d_v \sim 1$ nm by a recently determined value of $d_v \sim 0.1$ nm,[1] so that the corresponding spectral brightness figures have increased by a factor of 100.

As outlined previously, target current intensity is decisive for applications in nanostructure technology.[17] Estimates of processing times for ion beam modification of solid materials have been published recently.[18] The current level of 1 nA, however, is a conservative assumption, since maximum currents of about 10 nA have already been observed with a GFIS*. Hopefully, by optimizing the emitter structure, i.e., the geometry of base tip and supertip, and the operating conditions, i.e., system temperature, gas pressure, and electrical field, even further improvements may be possible.

Table I is an overview over our experimental results for beams of various gaseous ions and electrons obtained with supertips of both tungsten and iridium. Table II compares conventional electron sources operating in thermal and field emission mode with the characteristics of our EFES*.

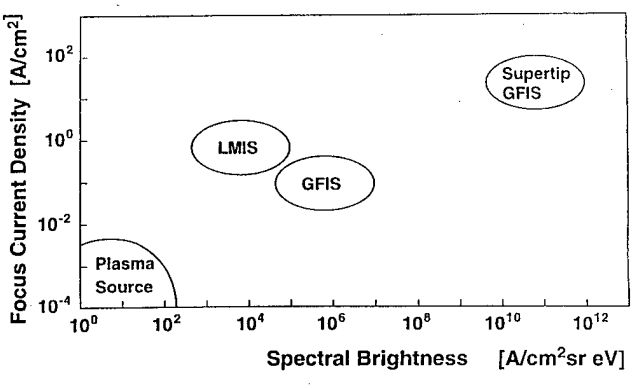


FIG. 14. Focused current density of different types of ion sources vs spectral brightness.

V. CONCLUSIONS

While in the pioneering period of some decades ago, typical target currents had amounted to a few pA, sufficient to render the atomic structure of surfaces visible, liquid metal ion sources have already provided currents of the order of 100 pA Ga ions, applicable to materials modification and analysis in the microstructure regime. Now, in a further development step, gas field ion source have been rendered capable of delivering even 10–100 times higher beam intensities to the image spot, which may become as small as 1 nm in diameter, so that novel applications in the field of nanotechnology become feasible. The future of gas field ion sources looks bright.

[1]Th. Miller, A. Knoblauch, and S. Kalbitzer, Mater. Sci. Forum **248/249**, 433 (1997).
[2]K. Jousten, K. Böhringer, R. Börret, and S. Kalbitzer, Ultramicroscopy **26**, 301 (1988).
[3]K. Böhringer, K. Jousten, and S. Kalbitzer, Nucl. Instrum. Methods Phys. Res. B **39**, 289 (1988).
[4]K. Jousten, K. Böhringer, and S. Kalbitzer, Appl. Phys. B: Photophys. Laser Chem. B **46**, 313 (1988).
[5]R. Börret, K. Jousten, K. Böhringer, and S. Kalbitzer, J. Phys. D **21**, 1835 (1988).
[6]A. Knoblauch, Th. Miller, Ch. Klatt, and S. Kalbitzer, Nucl. Instrum. Methods Phys. Res. B **139**, 20 (1998).
[7]S. Kalbitzer, Ch. Wilbertz, and Th. Miller, in *Nanolithography: A Borderland between STM, EB, IB, and X-Ray Lithographies*, edited by M. Gentili *et al.* (Kluwer, Dordrecht, 1994), p. 137.
[8]Th. Miller, A. Knoblauch, and S. Kalbitzer, Appl. Phys. A: Mater. Sci. Process A **61**, 99 (1995).
[9]R. Börret, K. Jousten, K. Böhringer, and S. Kalbitzer, J. Phys. D **21**, 1835 (1988).
[10]A. Knoblauch, Ch. Wilbertz, Th. Miller, and S. Kalbitzer, J. Phys. D **29**, 470 (1996).
[11]Th. Maisch, Ch. Wilbertz, Th. Miller, and S. Kalbitzer, Nucl. Instrum. Methods Phys. Res. B **80/81**, 1288 (1993).

[12]S. Kalbitzer and A. Knoblauch, Rev. Sci. Instrum. **69**, 1026 (1998).
[13]M. J. Southon, thesis, University of Cambridge, 1963.
[14]H. A. M. van Eekelen, Surf. Sci. **21**, 21 (1970).
[15]Ch. Wilbertz, Th. Miller, and S. Kalbitzer, Proc. SPIE **2194**, 407 (1994).
[16]B. Thompson, Micrion (private communication).
[17]S. Kalbitzer, Ch. Wilbertz, Th. Miller, and A. Knoblauch, Nucl. Instrum. Methods Phys. Res. B **113**, 154 (1996).
[18]K. Gamo, Nucl. Instrum. Methods Phys. Res. B **121**, 464 (1997).
[19]R. Gomer, *Field Emission and Field Ionization* (Harvard University Press, Cambridge, 1961).

Focused ion beam optical column design and consideration on minimum attainable beam size

Kiyoshi Sakaguchi[a] and Tetsu Sekine[b]
JEOL Ltd., 1-2 Musashino 3-chome Akishima, Tokyo 196, Japan

(Received 18 November 1997; accepted 14 April 1998)

We discussed a two lens optical system focused ion beam (FIB) employing a Ga-emitter (LMIS), whose acceleration voltage V_{acc} is ranged 30–100 kV, which will be used for milling and secondary ion microscope observation purposes. On such an optical system, we investigated the relation between minimum obtainable beam diameter and chromatic aberration coefficient against V_{acc}, and studied the possibility of improving FIB resolution by increasing the V_{acc}. The beam diameter is mainly determined by the size of a Gaussian image and the chromatic aberration, especially that of objective lens, if the beam current is very low. From this fact, we showed that the magnification optimization method which is one of the optical optimization methods can be greatly simplified at a lower beam current region. Using this simplified method, we summarized a guide line for evaluating an V_{acc} value from the standpoint of realizing a finer beam. Also given is information useful for designing a FIB column in consideration of optimized optical column design for realizing the optimized magnification. Finally, calculating the attainable minimum beam diameter and optimum column length against V_{acc} values, we discussed the limitation of beam diameter by increasing the V_{acc}. © *1998 American Vacuum Society.* [S0734-211X(98)04204-8]

I. INTRODUCTION

The Ga metal ion source is used quite widely as an ion source in focused ion beam (FIB) tools, with acceleration voltages V_{acc} typically in the range from 10 to 100 kV and which provide functions for ion milling and image observation as a scanning ion microscope (SIM). The Ga ion source provides excellent features such as ease in handling, high beam stability, good focusing property due to the small energy spread, and enough mass for high milling rates. Many FIB instruments, which are oriented for the above applications, employ the Ga metal ion source and use V_{acc} between 30 and 50 kV. The typical performance of such instruments is 5–8 nm as a SIM resolution with the beam currents of 0.5–1.5 pA.

Many FIB tools today employ a two-lens optical system except for a few cases.[1] A lot of efforts on optics investigations toward higher resolution have been made. Among them there seems to be two main research directions. The first one is to study the best beam operation modes (Fig. 1) for forming a smaller beam.[2,3] The recent study revealed that the diverging mode [Fig. 1(b)] has an advantage at any beam currents.[4] The second is to study how to optimize an ion optical system.[5,6] The typical examples would be "optimization of beam half-angle at the image plane" [7] and the "magnification optimization." [8] However, one of the problems in the second direction is one has to decide initially the acceleration voltage.

In designing an ion optics column, we must examine the V_{acc} and determine the most suitable value for milling and SIM observation. This must be done by taking into account the following factors: (1) milling rate, (2) specimen damage, and (3) required minimum beam diameter. The first purpose of this study is to establish a method for optimizing the V_{acc} for a two-lens FIB system providing Ga ion source, and thereby minimizing a beam diameter. Here, the V_{acc} was limited in the range 30–100 kV where the milling rate and the specimen damage do not change significantly. Thus, it can be determined only by taking into account the third factor, namely the minimum diameter requirement. Furthermore, we show that the magnification optimization method becomes a very simplified form in the low current range. Using the simplified form, the magnification is easily optimized for given V_{acc} value, extraction voltages, beam currents, working distances, and other lens parameters. We describe here the details of the proposed column design method and discuss the possible minimum beam diameter based on the calculated results derived from the above-mentioned method.

II. OPTICAL OPTIMIZATION FOR MINIMIZING BEAM SIZE

In this paragraph, we discuss the magnification optimization method in the low current region and show its validity in designing FIB column which uses Ga emitter and whose acceleration voltages are of 30–100 kV.

A. Magnification optimization at low current

It is well known that diffraction can be neglected in an ion optical system and is reported[8] that the quadratic addition of chromatic disk size d_c and spherical disk size d_s is a good approximation to calculate a current density integral d_{NFC} for the chromatic and spherical aberrations exactly. It is also assumed there[8] that the diameter of the finite source image d_G can also be added quadratically to give the probe size d_p to comparable accuracy. It is reported by other authors[9] that the above method does not give an accurate enough solution

[a]Electronic mail: sakaguch@jeol.co.jp
[b]Electronic mail: sekine@jeol.co.jp

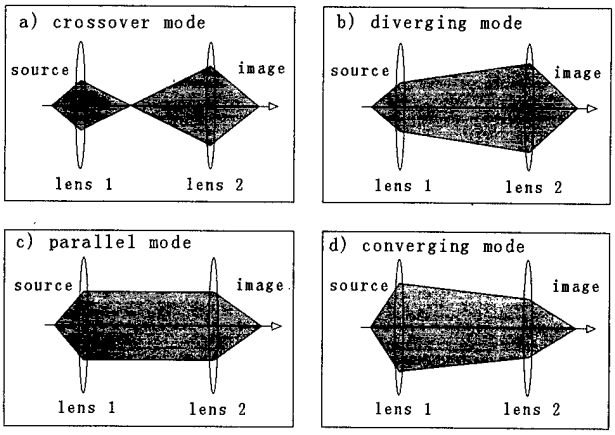


FIG. 1. Operation mode for two lenses optical system.

for the beam size calculation. Instead of this method they introduced "a new root power sum (RPS)" approximation. We compared these two methods in the case that the probe size is determined mainly by d_G and d_c. Two beam sizes obtained from quadratic addition and from RPS (FW50) differ from each other by under 1%. Therefore, we choose the quadratic addition method in this report because of its easiness to handle. The beam size is then given by the formulas below if all aberrations refer to the image side:

$$d_p^2=(m_1m_2d_0)^2+\left(C_c\alpha_i\frac{\Delta E}{V_i}\right)^2+\left(\frac{1}{2}C_s\alpha_i^3\right)^2, \tag{1}$$

$$C_s=C_{s,2}+m_2^4\ C_{s,1}, \tag{2}$$

$$C_c=C_{c,2}+m_2^2\ C_{c,1}, \tag{3}$$

where m_1 and m_2 and magnifications, $C_{s,1}$ and $C_{s,2}$ spherical aberrations, $C_{c,1}$ and $C_{c,2}$ chromatic aberration of lenses 1 and 2, C_s and C_c spherical and chromatic aberration of the total optical system, d_0 virtual source diameter, α_1 beam half-angle at image plane, V_0 and V_i are the potentials of object and image sides (i.e., V_0) is ion extraction voltage V_{ext} and V_i acceleration voltage V_{acc} and ΔE energy spread of beam. α_i can be replaced with beam acceptance half-angle α_0 using the following relationships:

$$\alpha_0=\sqrt{\frac{I_p}{J_0\cdot\pi}},\quad \alpha_{0,i}=\frac{\alpha_0}{m_1}\sqrt{\frac{V_0}{V_i}}, \tag{4}$$

$$\alpha_i=\frac{\alpha_{0,i}}{m_2}=\frac{1}{m_1\cdot m_2}\sqrt{\frac{V_0}{V_i}\cdot\frac{I_p}{J_0\cdot\pi}}, \tag{5}$$

where I_p is ion beam current, J_0 angular emission density, α_{qi} half-angle of the beam after passing through the lens 1 (this is equal to the half-angle of incident beam to lens 2). We assume here the following:

$$d_c^2=C_c^2\cdot\left(\frac{\Delta E}{V_i}\right)^2\cdot\alpha_i^2\gg\frac{C_s^2\cdot\alpha_i^6}{4}=d_s^2,$$
$$\frac{\partial C_c}{\partial m}\approx 0, \tag{6}$$

where

$$m=m_1\cdot m_2. \tag{7}$$

Equation (6) indicates that d_s is very small compared with d_c, thus it can be neglected. Equation (7), as stated in the reference,[8] indicates that the chromatic aberration weakly depends on magnification. In other words, it means that $dC_c/dm\ll C_c/m$, $dC_s/dm\ll C_s/m$. We discuss the validity for introducing Eqs. (6) and (7) in the next paragraph. Using Eqs. (4), (5), and (6), Eq. (1) is rewritten as follows:

$$d_p^2\approx(m_1m_2d_0)^2+\left(C_c\cdot\alpha_i\cdot\frac{\Delta E}{V_i}\right)^2$$
$$=(m_1m_2d_0)^2+\frac{C_c^2}{(m_1m_2)^2}\cdot\frac{(\Delta E)^2V_0I_p}{V_i^3\cdot\pi J_0}. \tag{8}$$

Based on Eq. (7), differentiating Eq. (8) with m, we can easily obtain the magnification m^* which minimize the beam diameter and the minimum beam diameter $d_{p,\min}$ at $m=m^*$:

$$m^*=\frac{d_{p,\min}}{\sqrt{2}\cdot d_0}, \tag{9}$$

$$d_{p,\min}=(2\cdot d_0\cdot C_c\cdot\Delta E)^{1/2}\cdot\left(\frac{V_0\cdot I_p}{\pi\cdot V_i^3\cdot J_0}\right)^{1/4}. \tag{10}$$

We must give an unknown value of $C_c(m)$ to calculate Eqs. (9) and (10), therefore, we will use an iteration method. We first give the initial value of $d_{p,\min}$ and m^* as follows:

$$d_{p,\min}(C_c)\approx d_{p,\min}(C_{c,2,\text{inf}}),\quad m^*(C_c)\approx m^*(C_{c,2,\text{inf}}), \tag{11}$$

where $C_{c,2,\text{inf}}=C_{c,2}$ at $m_2=-\infty$. Then we get an initial value of $d_{p,\min}$ and m^*, but these values are not exact. We then use the conclusion discussed in Sec. III A that the optimum column length will be shortest, if lens 1 is so operated, to minimize the value of m_1 ($m_1>0$: diverging mode). We put $m_{1,\min}$ as the minimum value of m_1 at $m_1>0$. Then we can calculate the first iteration value of C_c by using the formula below:

$$C_c=C_{c,2}(m_2)+m_2^2C_{c,1}(m_1),$$

where

$$m_1=m_{1,\min}\quad m_2=m^*/m_{1,\min}. \tag{12}$$

By substituting C_c into Eqs. (9) and (10), we can get the first iteration value of $d_{p,\min}$ and m^*. These are exact enough to our purpose. Equation (10) is the approximate expressions and valid as far as Eqs. (6) and (7) are satisfied. Also, it was derived by neglecting spherical aberration, thus always holding $d_p>d_{p,\min}$. Therefore, we regard that the $d_{p,\min}$ is the minimum beam diameter which the system of interest can attain. Equation (10) will be very effective for evaluating the possible minimum diameter in an optimization. For example, when emitter characteristic: J_0, ΔE and d_0 are given, we can easily extract the $d_{p,\min}$ vs. C_c relation from the parameters V_i, V_0, and beam currents I_p. An example is shown in Fig.

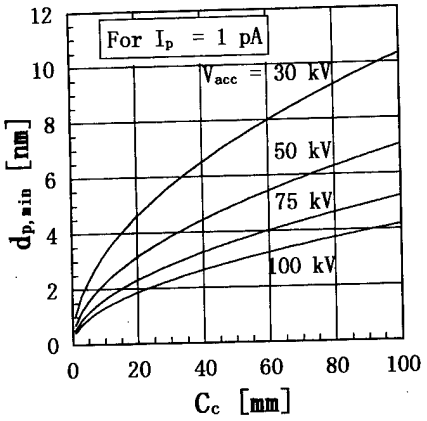


FIG. 2. C_c–$d_{H\,\min}$ characteristic.

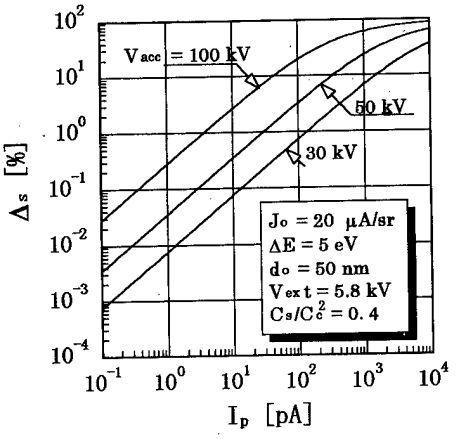


FIG. 3. Calculation error by ignoring spherical aberration.

2. In the calculation, the emitter characteristics were supposed, from the Refs. 10–13, as $J_0=20\ \mu$A/sr, $\Delta E=5$ eV, and $d_0=50$ nm, $V_0=5.8$ kV and $I_p=1$ pA were also supposed. Now, from Eq. (3), $C_c>C_{c,2}$ ($C_c\approx C_{c,2}$ for $V_i<50$ kV) always remains, so that we may regard $C_{c,2}$ instead of C_c, as the best possible total chromatic aberration for the two lens system.

B. Validity of the calculation

First, we examine the validity of introducing the Eq. (6). To compare the sizes of spherical disk size d_s to chromatic disc size d_c, we take their ratio as d_s/d_c and, from Eq. (6), it is expressed as:

$$\frac{d_s}{d_c}=\frac{1}{2}\cdot\frac{C_s}{C_c}\cdot\frac{V_i}{\Delta E}\cdot\alpha_i^2$$

$$=\frac{1}{2\cdot(m_1\cdot m_2)^2}\cdot\frac{C_s}{C_c}\cdot\frac{V_0}{\Delta E}\cdot\frac{I_p}{\pi\cdot J_0}. \tag{13}$$

If the value of m_1m_2 were limited in the values close to m^*, it can be approximated as:

$$\frac{d_s}{d_c}\approx\frac{1}{2}\cdot\frac{C_s}{C_c^2}\cdot\frac{d_0}{\Delta^2 E}\cdot\left(\frac{V_0\cdot V_i^3\cdot I_p}{\pi\cdot J_0}\right)^{1/2}. \tag{14}$$

It is effective when Eq. (7) holds. Using Eq. (14), we discuss here the error Δ_s (%) included in the calculation for the beam diameter, when neglecting the spherical disk size d_s. The validity of this is discussed later. The magnification m^* given by Eq. (9), at which the minimum beam diameter is attained, also meets the condition that the Gaussian image diameter and the chromatic disk size d_c are equal each other. Taking this into account, Δ_s (%) is given by the equation:

$$\Delta_s(\%)\approx 100\cdot\left(1-\sqrt{\frac{2}{2+f^2}}\right),$$

where

$$f=\frac{d_s}{d_c}. \tag{15}$$

An example for I_p–Δ_s characteristic at different V_i values is shown in Fig. 3. The conditions $J_0=15\ \mu$A/sr, $\Delta E=5$ eV, $d_0=50$ nm, and $V_0=5.8$ kV are supposed there. In addition, as described in the later paragraph on primary lens design, we used the relation $C_s/C_{c,2}^2=0.4\ \text{mm}^{-1}$, which is easily obtained in the range $30\leqslant V_i\leqslant 100$ kV. From the result, we have concluded that, in the calculation of beam diameter, the term for spherical disk size d_s can be neglected in the beam current range of below several pA.

Generally speaking, it is not possible to express the aberration coefficient in an analytical form with magnification as a parameter. Therefore, it is also impossible to consider the validity of Eq. (7) by an analytical manner. Instead, we have compared the results of two calculations, one based on full calculation and the other based on approximation by Eqs. (7) and (12), for the optical model discussed in Sec. IV B. The comparison was made in the ranges $30\leqslant V_I\leqslant 100$ kV and $I_p<3$ pA ($0<m<0.1$) of our interest. We have found from the comparison that the beam diameter given by Eq. (10) is always less than that given by the full calculation, and the error included in the former calculation varies depending on V_i value and gives the maximum value (<5%) at 100 kV.

Until now, we have neglected the Coulomb interaction effects in the discussion. The reasons are as follows:

(1) It is well known that the space charge does not much affect the beam diameter in the beam current range below several hundred pA.[14,15] Specifically, the beam size degradation at the beam current of $I_p\approx$ several pA will be <1 nm[4]. The regions of interest here are $I_p<1.5$ pA and $3.2<d_p<6$ nm (see Fig. 8). In this case it seems that the quadratic value of degradation is negligible compared with that of other contributions to beam size.
(2) We investigate only diverging mode which is better for forming a smaller beam. The Coulomb effects will be less in the diverging mode than in the cross over mode.[16]
(3) Coulomb interaction between the emitter (LMIS) and the first beam acceptance aperture is included in the virtual source size (~50 nm).

From the considerations shown above, Eqs. (9), (10), and

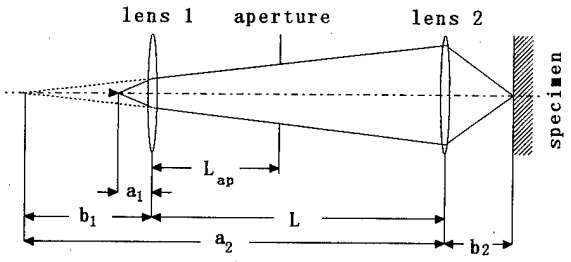


FIG. 4. Definition of optical system parameters.

(12) derived by neglecting the Coulomb effects and the estimated error value seem for our purpose, to be valid.

III. OPTIMIZED OPTICAL COLUMN DESIGN FOR FIB

In the previous section we discussed the optimum magnification by which the minimum beam size is realized. However, the optimum magnification is not always fulfilled in arbitrary optical systems. Here we investigate the optical system which realizes the optimum magnification, discuss the items that should be considered as the factors related to minimum beam diameter, and submit useful information for optical column design.

A. Optimized column length

The value of magnification m is bound to form an image at a specimen surface. In the optical system shown in Fig. 4, the relations between the magnification and the object point and also between the magnification and image point are given as follows:

$$|m_1|=\left(\frac{V_0}{V_i}\right)^{1/2}\cdot\frac{b_1}{a_1},\quad |m_2|=\frac{b_2}{a_2}=\frac{b_2}{L+b_1}. \tag{16}$$

From these, the focusing condition which m_1 and m_2 must satisfy is determined by:

$$|m_2|=\frac{b_2}{L+a_1\cdot|m_1|/g},$$

where

$$g\equiv\left(\frac{V_0}{V_i}\right)^{1/2}. \tag{17}$$

From Eqs. (9), (10), and (17), the magnification that meets the optimum magnification condition and the focusing condition simultaneously is derived as:

$$|m_1\cdot m_2|=\frac{|m_1|\cdot b_2}{L+a_1\cdot|m_1|/g}=\frac{d_{p,\min}}{\sqrt{2}\cdot d_0}\equiv m^*. \tag{18}$$

Solving it for the distance L between lenses, we obtain

$$L=|m_1|\cdot\left(\frac{b_2}{m^*}-\frac{a_1}{g}\right)=\frac{1}{|m_2|}\cdot\left(b_2-\frac{a_1}{g}\cdot m^*\right), \tag{19}$$

where a_1 is the distance between the emitter tip and the principal plane of the lens 1 and b_2 the distance between the principal plane of objective lens and the specimen surface.

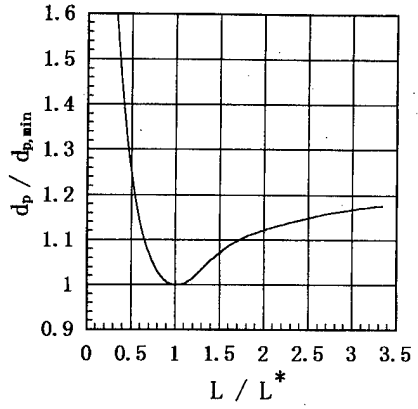


FIG. 5. L–d_p characteristic.

When a_1 and b_2 are bounded to satisfy $b_2/a_1>m^*/g$, the optimum magnification m^* exists, by which the focusing condition is simultaneously satisfied, because then $L>0$.

We then put L^* as the optimal column length L which satisfies Eq. (19) and $m_1=m_{1,\min}$ for the shortest column. Then, we discuss the degradation of resolution against the shift of L from L^*. An example is given in Fig. 5. It reflects the characteristic of the optical system described in the next section (Fig. 8). The data are for the case where $V_i=50$ kV and the lens 2 works in deceleration mode. It is noticed that the resolution degrades slightly where $L>L^*$, while it degrades rapidly where $L<L^*$. From our experience, it is valid for general cases.

B. Consideration on lens design

From the theoretical point of view, the optimum column length L can be determined uniquely by Eq. (19). However, for column length, shorter is preferable in forming a smaller beam practically, because it is good for (1) reducing mechanical vibration, (2) minimizing space charge effects. Meanwhile, in order to form a small beam in the low current region, a lens of small chromatic aberration is required. We do not comment here about it because many authors have already stated it in Refs. 17 and 18. We consider the requirements imposed on the lenses for reducing L^*.

Among the parameters relating to m^*, d_0, J_0, and ΔE, which are the parameters concerning emitter characteristics, may be regarded as fixed parameters, because one would use the generally available best values. The I_p may be treated similarly as well, because it must be set in a limited narrow range due to the tradeoff of the requirement that lower current is good for finer beam and the requirement that higher current is good for higher image quality.

Then, under a given acceleration voltage V_i, the column length L^* must be determined by the extraction voltage V_0, the object position toward lens $1a_1$, the magnification m, the chromatic aberration of lens $2C_{c,2}$, and the image position b_1. Through the analysis of Eqs. (9), (10), and (19), we can summarize them as follows:

(1) Extraction voltage V_0: A larger V_0 allows shorter L^*, however, it would make $d_{p,\min}$ larger as seen in Eq. (10). For

example, at $V_i=50$ kV, the increment for V_0 from 6 to 12 kV, will induce a beam broadening of about 1.0–1.5 nm. On the other hand, the L^* will be shorter by 50–100 mm. Thus, V_0 cannot be determined only by the emitter characteristics.

(2) Image plane b_1 and magnification m_1 of lens 1: In designing lens 1, one must set b_1 as small as possible in order to reduce L^*, namely making $|m_1|$ as small as possible in the diverging mode ($m_1>0$). Moreover, it is very important to operate lens 1 holding $|m_1|$ to be minimum.

(3) Image plane b_2 and chromatic aberration $C_{c,2}$ of lens 2: The chromatic aberration and b_2 are the values which must be derived as the design parameters of a fundamental optical component, namely the objective lens. We only give a comment here that, regarding the nature of electrostatic lens used as an objective lens, shorter L^* can be attained in the deceleration mode than in the acceleration mode (see Sec. V).

C. Beam defining aperture

We discuss the position L_{ap} and the diameter D_{ap} of the beam defining aperture which meets the minimum beam diameter condition at the minimum current. Referring to Fig. 4, we obtain the beam current by substituting Eqs. (5) and (16) into Eq. (4):

$$I_p=\pi J_0 m_1^2 \frac{V_i}{V_0}\cdot \arctan^2 \frac{D_{ap}}{2\cdot(L_{ap}+a_1 m_1\sqrt{V_i/V_0})}. \quad (20)$$

Solving it for L_{ap}, the following equation is derived:

$$L_{ap}=-a_1 m_1\sqrt{\frac{V_i}{V_0}}+\frac{1}{2}D_{ap}\left[\tan\left(\frac{I_p\cdot V_0}{\pi J_0 m_1^2 V_i}\right)^{1/2}\right]^{-1}. \quad (21)$$

As for the minimum aperture diameter, it is empirically recommended that $D_{ap}/T_{ap}>1$ will be satisfied (T_{ap}: the thickness of aperture) for reducing generation of species at the edge of the aperture hole, which degrades the resolution.

Once the aperture size corresponding to the minimum beam diameter is determined, then its position will be determined by Eq. (21). In a practical column design, the condition $L^*>L_{ap}$-must be held. However, if $L^*<L_{ap}$, we must try to make L^* longer according to the guide shown in Sec. III A unless the minimum beam condition described in Sec. II cannot be attained. Since the L^* depends on acceleration voltages, such a case will be possible when the system works under the lower kV. Because the ion optical system is usually designed in a way that L^* is determined at the maximum acceleration voltage, the degradation of beam diameter will not be so significant since it is included in the abovementioned condition.

IV. OPTIMIZATION OF ACCELERATING VOLTAGE

In this section we study the optimization method of acceleration voltage V_i to get the smallest beam diameter and show one example of optimizing to guess a practical value of attainable beam diameter and the most suitable V_i value when the lens properties and working distance are given after the primary lens design. We assume the conditions as described below:

(1) usage of the characteristic of available Ga emitter and available electrostatic lenses today;

(2) diverging mode as an operation mode because of its advantage[4] for the smallest beam diameter;

(3) accelerating voltage: $30\leqslant V_i\leqslant 100$ kV so that the milling rate and damage on the specimen does not change significantly.

A. Optimizing procedure of acceleration voltage

Through the study of the magnification optimization method, we summarize the recommended procedure for determining the best acceleration voltage to get the minimum beam diameter as shown below.

(1) Consideration of the minimally applicable beam current: The beam current I_p for probing the minimum beam diameter is determined based on the required S/N ratio for a SIM image (it depends on the secondary electron yield, detector efficiency kind specimen and accelerating voltage), the required time for image accumulation and the level of specimen drift, which are mutually related. Currently it is thought that $0.5<I_p<1$ (pA) will be reasonable.

(2) Primary design for lens 1 and lens 2: For a given maximum voltage to the lenses, the primary design is performed and the shapes for the lenses 1 and 2 are determined, where they meet the request for the upper limit for the applicable electric field strength and beam stability for lens 1. Also, the optical properties at each acceleration voltage are calculated. Since the beam broadening at the minimum size is dominated by the chromatic aberration, which is not so sensitive to the lens shape, the precise shape optimization may not be necessary at this stage.

(3) Magnification optimization and optical design: From the effective beam current range [examined in (1)] and the derived lens properties [calculated in (2)], we can perform the "magnification optimization" (refer to Sec. II) and get the minimum attainable beam diameter at a given accelerating voltage. Through the optimized optical design (refer to Sec. III) we get the information about the optical system parameters, i.e., optimized column length, position, and diameter of beam defining aperture, etc., which can realize the optimized magnification. Working distance and extraction voltage must be decided from other considerations.

(4) Determination of optimum acceleration voltage: By performing process (3) at each accelerating voltage, the obtainable minimum beam diameter is derived together with the conditions: the best column length, aperture size, acceleration voltage, etc. The final acceleration voltage is decided here by taking into account the cost as well.

(5) Reoptimization of the total optical system with the derived parameters (lens shape optimization is included).

B. Example for optimizing

We show here an example for acceleration voltage optimization and estimate the possible minimum beam diameters. We discuss it by limiting the beam current at 0.5 pA only. In the primary design of lenses, the following conditions were imposed on them:

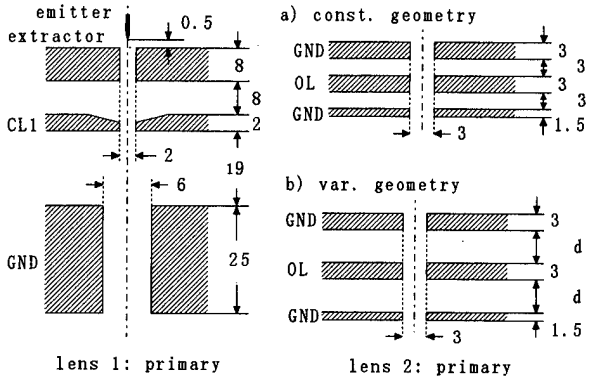


FIG. 6. Primary lens design of lens 1 (CL_1) and lens 2 (OL).

TABLE I. Lens 2 properties of primary design: "Mode" means the operation mode of OL, "geometry" describe the OL lens shape, which is decided independently on V_1 (const.) or dependently (var.). WD means working distance.

V_i	Mode	Acceleration		Deceleration	
(kV)	geometry	Const.	Var.	Const.	Var.
30	WD (mm)	17.71	17.71	10.00	10.00
	C_{cinf} (mm)	30.66	30.66	49.09	42.69
50	WD (mm)	42.38	31.35	10.00	10.82
	C_{cinf} (mm)	72.38	53.61	48.86	48.73
75	WD (mm)	85.26	84.45	28.74	13.37
	C_{cinf} (mm)	148.88	144.65	92.26	60.34
100	WD (mm)	141.56	176.00	66.03	30.98
	C_{cinf} (mm)	252.36	305.67	175.37	107.35

(1) Condition for lens 1: maximum electrical field strength: 20 kV/mm, extractor voltage V_0:6.0 kV, acceleration voltage V_i:30–100 kV, maximum CL1 voltage V_{CL1}:20 kV (refer to V_i), lens geometry, which realizes the beam stability. The schematic of the lens geometry which will satisfy the above conditions is shown in Fig. 6 (left-hand side) from the calculation. For simplification, it is supposed that the same lens geometry is used in $30 \leqq V_i \leqq 100$ kV. Since the beam diameter at low beam current depends weakly on the lens 1 form (discussed Sec. II B), it will be correct, in estimating a beam diameter. The lens characteristic is shown in Fig. 7.
(2) Condition for lens 2: maximum electrical field strength: 20 kV/mm, maximum OL voltage V_{CL}:50 kV, minimum working distance: 10 mm.

The lens geometry which will satisfy the above conditions is shown in Fig. 6 (right-hand side). As the operation modes for the objective lens, we discuss both the acceleration and deceleration modes. The calculation is made for the shapes shown in (a) and (b) in Fig. 6. The reasonable space length d (defined in Fig. 6) is given as follows: acceleration mode: $d=0.10\,V_i$ (mm) (V_i in kV), deceleration mode: $d=0.05V_i$ (mm) (V_i in kV).

These rules have been drawn from the experience. The main lens characteristic at the acceleration voltages: 30, 50, 75, and 100 kV are given in Table I. The working distance there indicates the distance between the surface of the lower lens electrode and the specimen. With the characteristic data, applying the optimum magnification method mentioned in Sec. II, the dependency of optical parameters shown in Fig. 8 are obtained. Since the chromatic coefficient does not much depend on the lens form (substantially <6%–7%) in these voltages, the calculation for the acceleration voltage dependency of beam diameters based on available emitter characteristics and practically possible lens forms will be acceptable (error may be <10% including that of space charge effect neglect). The calculation was made under the constant lens 1 form. If lens form optimization were also made at each acceleration voltage and the absolute value of magnification in the diverging mode could be close to 1, the column length shown in Fig. 8 may become shorter.

V. CONCLUSION

For a FIB comprised of a two lens optical system with acceleration voltages ranging from 30 to 100 kV, which is used mainly for milling and SIM observation purposes, we have investigated the minimum obtainable beam diameter

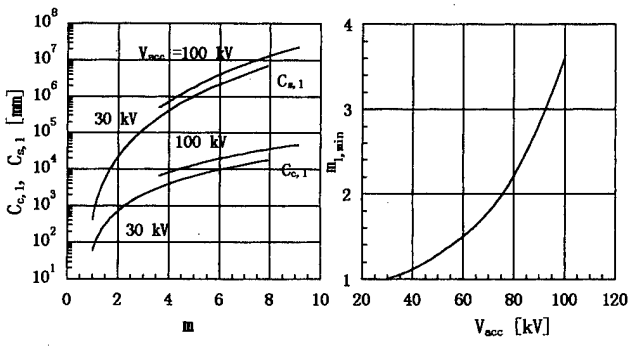


FIG. 7. Lens 1 properties of primary design.

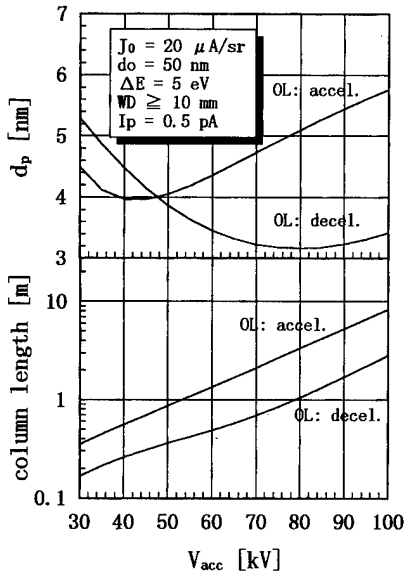


FIG. 8. Beam size and optimum column length depending on V_{acc}.

and chromatic aberration of the objective lens against acceleration voltage and we have shown that the magnification optimization method, which is effective in the low beam current region as one of the optical system optimization methods, can be simplified very much and be made a guideline for evaluating the optimum acceleration voltage that produces the highest resolution under given conditions. Also, in discussing the optimized optical column design to realize the magnification optimization, we have summarized useful information for designing a FIB column. Finally, we have calculated the minimum obtainable beam diameter and optimum column length dependence on acceleration voltage where optical parameters have been prepared. We found from these a conclusion that the best lens mode for lens 2 is the acceleration mode for $V_{acc}<47$ kV, and the deceleration mode for $V_{acc}>47$ kV. Also, the beam diameter is minimum at around $V_{acc}\approx 75$ kV. It is noticed that higher acceleration voltage is not necessary for higher spatial resolution if the characteristics of the Ga emitter and optical system are limited within that available today. Moreover, the column length can essentially be shorter in the deceleration mode operation than in the acceleration mode. Finally, it should be noticed that although Coulomb effects were not taken into account in this calculation, the accuracy of our calculation seems to be within 10% of the result if they were taken into account.

[1]K. Sakaguchi, K. Hoshino, T. Ishimoto, D. Kamuki, T. Goto, K. Kumagai, and T. Sekine, Proceedings of LSI Testing Symposium, Japan (1996), p. 152.
[2]J. Orloff, J. Vac. Sci. Technol. B **5**, 175 (1987).
[3]P. Kruit and X. R. Jiang, J. Vac. Sci. Technol. B **4**, 1635 (1996).
[4]L. Wang, J. Vac. Sci. Technol. B **15**, 833 (1997).
[5]R. Hill, Proc. SPIE **2014**, 112 (1993).
[6]T. Ishitani and T. Kawanami, J. Vac. Sci. Technol. B **13**, 372 (1995).
[7]S. Takashima, JEOL News **31E**, 33 (1994).
[8]Y. L. Wang and Z. Shao, Adv. Electron. Electron Phys. **81**, 177 (1991).
[9]J. E. Barth and P. Kruit, Optik (Stuttgart) **101**, 101 (1996).
[10]S. Rao, A. E. Bell, G. A. Schwind, and L. W. Swanson, J. Vac. Sci. Technol. B **7**, 1787 (1989).
[11]L. Zhou *et al.*, J. Vac. Sci. Technol. B **9**, 2953 (1991).
[12]L. W. Swanson, G. A. Schwind, and A. E. Bell, J. Appl. Phys. **51**, 3453 (1980).
[13]M. Komuro, H. Arimoto, and T. Kato, J. Vac. Sci. Technol. B **6**, 923 (1988).
[14]P. Kruit and X. R. Jiang, J. Vac. Sci. Technol. B **13**, 1635 (1996).
[15]X. R. Jiang, J. E. Barth, and P. Kruit, J. Vac. Sci. Technol. B **14**, 3747 (1996).
[16]Y. Kawanami and T. Ishitani, J. Vac. Sci. Technol. B **8**, 1673 (1990).
[17]X. R. Jiang and P. Kruit, Microelectron. Eng. **30**, 249 (1996).
[18]K. Kurihara, J. Vac. Sci. Technol. B **3**, 41 (1985).

Overlay accuracy tests for direct write implantation

S. Mogren and I. L. Berry
Microelectronics Research Laboratory, 9221 Rumsey Road, Columbia, Maryland 21045

(Received 4 March 1998; accepted 14 April 1998)

Direct write implants are usually made in registration with pre-existing features on the wafer surface. The accuracy of the registration is a critical component of the error budget for a given device, just as it is for optical lithography layers. We present a scheme for creating alignment marks in field oxide on a silicon wafer. These alignment marks can be applied near the beginning of the process, are robust in a silicon processing environment, and can be reused for multiple layers of maskless implantation. We also present our method of measuring overlay accuracy, and data showing that an accuracy of 62 nm one sigma can be achieved by our system.
[S0734-211X(98)01604-7]

I. INTRODUCTION

Overlay accuracy is one of the critical factors limiting the smallest critical dimension that can be drawn in any multilayered lithographic process. Direct write implants using a focused ion beam (FIB) system are no exception. The ability to write fine features must be accompanied by the ability to align these features accurately with respect to previous lithographic steps.

To align to different features, each step uses the measured position of reference marks previously placed on the wafer. As in previous work,[1] the ultimate alignment precision largely depends on five factors. First, the reference mark must be written with good fidelity and accuracy. Second, subsequent processing must not significantly degrade the reference mark appearance. This includes damage to the mark due to the sputtering from the ion beam, and loss of image quality due to surface charging effects. Third, the stage and laser systems must be stable and accurate. Fourth, the beam deflection system must be calibrated to the stage. Fifth, stage drift must be small for the time it takes to implant a feature. In this paper we describe an alignment system for silicon wafers with durable marks, and a robust capture scheme that allows mechanical leeway in mounting the wafer to the pallet. We also describe the tuning requirements for our FIB system necessary to achieve good alignment. Finally, we describe our approach to measuring the overlay accuracy of direct write implants, and present our results.

II. ALIGNMENT SCHEME

In a standard silicon process, many ion implants are performed after the field oxide has been grown on the wafer. We place the alignment marks by patterning the field oxide in the same way active area is patterned. This leaves marks that are compatible with subsequent processing, and are made of relatively thick and durable silicon dioxide. The marks have good contrast in a secondary electron image, as shown in Fig. 1, giving bright regions where the oxide is thin. The contrast is probably due to a buildup of surface charge on the thick oxide preventing secondary electrons from escaping, while the passage of an ion all the way through the thin oxide into the substrate allows the carriers generated in the oxide to discharge the surface, producing a large secondary electron signal. The presence of surface charge does not seem to affect the apparent mark position. These marks have been reused more than twenty times, and averaging the positions of opposing edges of the mark produced a position that did not change noticeably.

The system uses the alignment marks in a two step process. First, four large global marks are used to determine the shift and magnification difference between the computer database and the wafer surface. Second, as each die is started, four local marks on the die are located to remove the effects of long term drift. The strategy of finding these marks is described below.

A global alignment mark is shown in Fig. 2. The system starts the alignment process by assuming there is a one-to-one correspondence between the wafer surface and its database, and moving to the position of the middle of the first

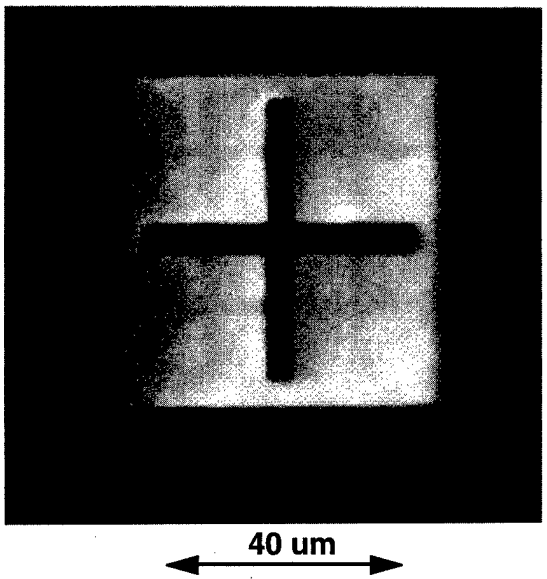


FIG. 1. Secondary electron FIB image of an alignment cross. The silicon dioxide is 40 nm thick in the bright regions, and 700 nm thick in the dark areas. The high contrast reduces the effects of noise on the measured cross center position.

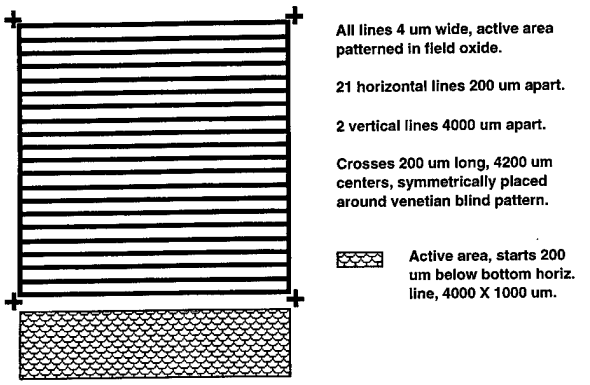


FIG. 2. Global alignment mark. The 4 mm × 4 mm structure allows the system to capture the wafer position even when the wafer has been placed on the pallet as much as 1 mm off of optimum. The 4 mm × 1 mm area of thin oxide at the bottom is useful for sputtering patterns used to calibrate the system.

global mark. The system then scans along a 200 μm vertical line looking for one of the horizontal lines. Once a horizontal line has been located, the system moves up and checks for more lines until it reaches the top. Next it drops down to the bottom line, and steps to the left until it finds the lower left corner of the 4 mm square. Using the position of the lower left corner, and the slope of the bottom horizontal line, the system estimates the shift and rotation of the wafer. The locations of the lower left, upper left, and upper right crosses around the global are then measured, and these locations are used to calculate a transform between the database and wafer coordinates. Next, the lower left cross positions on three other global marks are measured, allowing a more accurate transform to be calculated. Finally, the lower left crosses are all measured again, positioning the stage so that the minimum of electronic deflection is used, reducing the component of this error that contributes to the transform calculation. We have found that this alignment scheme allows us to accurately capture the position of a wafer that is placed more than 1 mm off of its ideal position. This entire process requires the electronic deflection to be carefully tuned to produce a minimum of errors in cross position measurement and pattern placement.

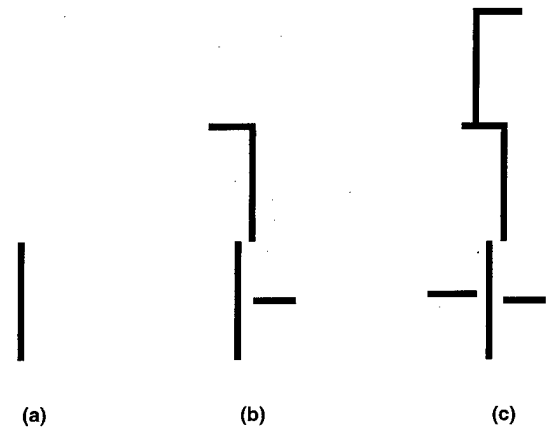


FIG. 3. After the system has been focused, the deflection step size and rotation are calibrated by sputtering in thin oxide (a) a line written in the center of the writing field, (b) an additional pattern written with the stage moved to the right 100 μm and the beam deflected right to match, and (c) a final pattern written with the stage moved 100 μm to the left of its original position and the beam deflected 100 μm to the left to match. The distance the stage moves is measured by an interferometer to an accuracy of 4 nm. When the step size (shown here as slightly too big) and rotation are correct, the line segments will join to form a continuous line.

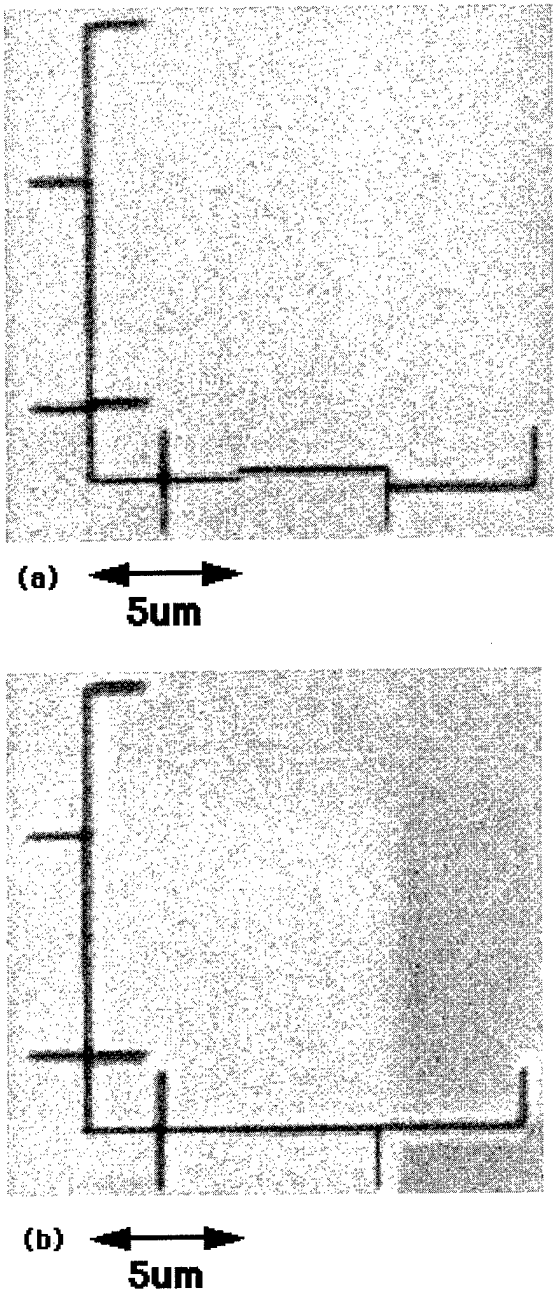


FIG. 4. Secondary electron FIB images of a sputtered deflection calibration pattern made using 240 keV As^{++}. They show (a) proper X step size, a small X rotation error, a need for a decrease in the Y step size, and proper Y rotation; and (b) the pattern for a well tuned system.

III. TUNING THE SYSTEM

The goals of tuning the system are to obtain a linear mapping between the binary deflection numbers and the physical

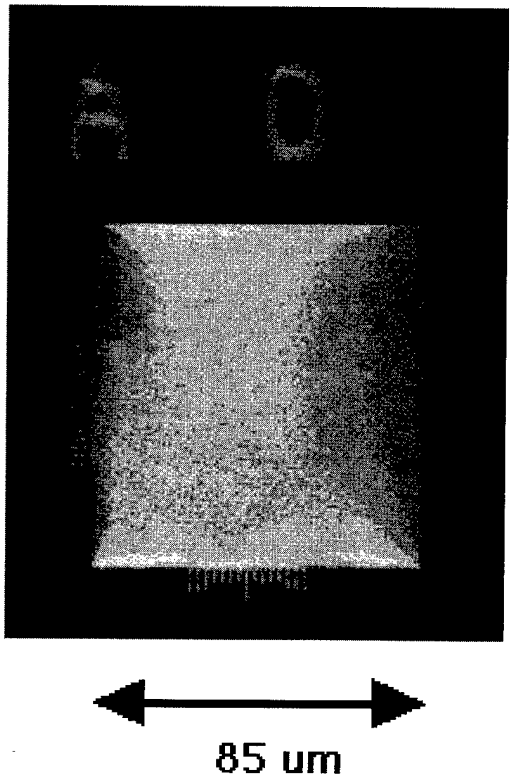


FIG. 5. Secondary electron FIB image of the overlay structure patterned into the field oxide. Each pattern in the 10×10 array has a unique alphanumeric label. Overlay errors are measured by observing the shift between the central mark in the X or Y vernier and a matching 2 μm sputtered mark.

position the beam is deflected to on the wafer, to produce the smallest beam size, and to calibrate the beam step size to an accuracy of 0.05%. Nonlinear mapping in our system [a NanoFab 150 manufactured by Finkelstein Associates, Inc. $(FAI)^2$] is a result of the lower lens being shifted from center. By imaging at low energy (15 kV), the field of view becomes large enough that a silhouette of the lower lens becomes visible. This can be used to precisely align the lower lens. Next, the energy is ramped up to the value needed for the implant, and the focus and stigmator values are adjusted to optimum. Finally, the beam deflection (step size) is calibrated in three steps. First, a reference line is sputtered into the surface in the center of the writing field. The resulting line is shown in Fig. 3(a). Second, the stage is moved to the right 100 μm, the beam is deflected to the right 100 μm, and an "L" shape is sputtered into the sample surface. This produces the pattern in Fig. 3(b). The motion of the stage is measured by the laser interferometer to an accuracy of a few nanometers. If the X deflection is correctly calibrated, the long side of the L forms a continuous line with the previously sputtered segment. If the value the system uses for the step size is too big, the beam will not be deflected enough, and there will be a shift between the two line segments. Third, the stage is moved 200 μm to the left, and another L shape is sputtered into the sample surface as shown in Fig. 3(c). If the X deflection is correctly calibrated, the long side of this L forms a continuous line with the previously sputtered segments. The Y deflection is calibrated in a similar way. In addition, two short line segments near the bottom of Fig. 3(c) form a continuous line when the rotation adjustment is set properly. In practice, steps one through three are accomplished for both axes at the same time by running a single file on the machine, producing a pattern like that shown in Fig. 4(a). The operator then examines the resulting

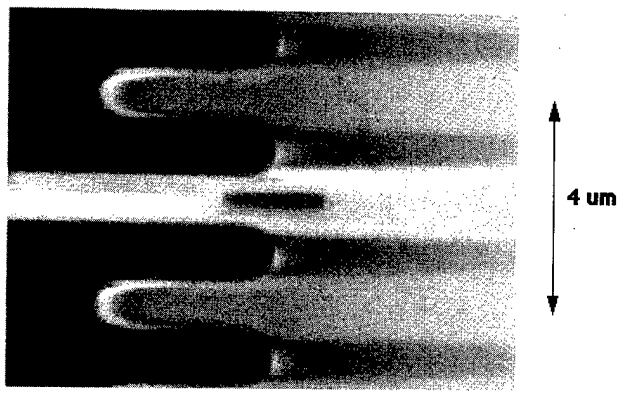
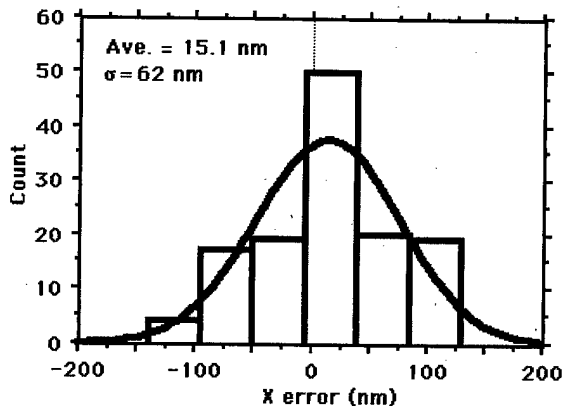
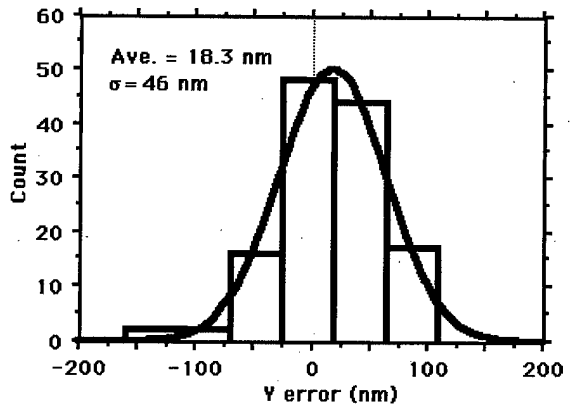


FIG. 6. Secondary electron FIB image of a sputtered mark properly placed next to its matching mark in the Y alignment vernier.



(a)



(b)

FIG. 7. Overlay error histograms for (a) the X direction, and (b) the Y direction. The average stage move is 1000 μm in X and 10 μm in Y, producing a wider distribution for the X overlay.

pattern, and makes the appropriate adjustments. After a few iterations, sputtering the deflection calibration pattern produces the properly aligned pattern shown in Fig. 4(b). The beam step size is a sensitive function of sample height, and the wafer surface must be flat to within 10 μm to maintain calibration. This procedure for calibrating the system produces the minimum error in overlaying one pattern onto a previous one.

IV. OVERLAY ACCURACY MEASUREMENTS

An overlay alignment structure is shown in Fig. 5. On a given chip, a 10×10 array of these is placed on 1 mm centers, giving 100 measurement sites. This structure was designed to have a matching vernier sputtered into the surface, and the alignment results read in an optical microscope. However, it takes several minutes to create a sputtered vernier, giving an overlay error that is dominated by the long term stage drift. The device implants we are interested in are microns in size, and only take a fraction of a second. Thus we find that a single 2 μm line sputtered next to the central line in the vernier structure in the field oxide in Fig. 5 gives a more accurate picture of machine performance. An example of this is shown in Fig. 6. In practice, it takes approximately 10 min to expose 44 sites arrayed around the periphery of a chip, and approximately 1 h to manually image each site and read the alignment error. Automating these measurements will hopefully eliminate much of this tedious work.

The aggregate data from three chips is shown in Fig. 7. These implants use As++ at 120 kV, with a 5 mil beam defining aperture. Two features are immediately apparent. First, we have difficulty manually measuring the position of the sputtered line to better than 45 nm precision. Despite this lack of resolution, the Y error average and standard deviation is 18.3±46 nm, and the X error average and standard deviation is 15.1±62 nm. Second, the normal distribution curve overlaid on the Y error histogram matches fairly well with the data, while the X error data is more spread out within the range from −100 nm to 100 nm than the normal distribution would predict. This results from the way the exposure sites are implemented, starting with the lower left most, and proceeding to the right until all of the sites on a horizontal line have been visited. Next, the stage moves to the left hand side of the chip, and proceeds to implant a row slightly higher than the last. As a result, all stage moves in the X direction are 1 mm or greater while most of the Y stage moves are of the order of 10 μm, and the stage takes longer to settle in X than it does in Y. There is an adjustable software delay time to allow the stage to settle after each move. This delay has been set to be as short as possible and still yield useable results, and this is a major factor in the increased X position error.

V. CONCLUSION

We have developed a simple scheme for direct write FIB implantation that does not add new lithography layers to a silicon process, and results in alignment marks that can be used repeatedly for implants of various energies and ion species. We have used this alignment scheme to experimentally demonstrate overlay errors of 62 nm one sigma using a mark sputtered next to a matching field oxide feature.

[1]K. Hosono, H. Morimoto, Y. Watakabe, and T. Kato, Proc. SPIE **923**, 84 (1988).
[2]N. W. Parker, W. P. Robinson, and J. M. Snyder, Proc. SPIE **632**, 76 (1986).

Development of ion and electron dual focused beam apparatus for high spatial resolution three-dimensional microanalysis of solid materials

Zhaohui Cheng,[a] Tetsuo Sakamoto, Masanori Takahashi, and Yasuyuki Kuramoto
Institute of Industrial Science, University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8558, Japan

Masanori Owari
Environmental Science Center, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Yoshimasa Nihei
Institute of Industrial Science, University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8558, Japan

(Received 16 November 1997; accepted 27 March 1998)

We constructed an ion and electron dual focused beam apparatus to develop a novel three-dimensional (3D) microanalysis technique. In this method, a Ga focused ion beam (Ga FIB) is used as a tool for successive cross sectioning of the sample in the "shave-off" mode, while an electron beam (EB) is used as a primary probe for Auger mapping of the cross sections. Application of postionization with EB to Ga-FIB secondary ion mass and two-dimensional (2D) elemental mapping with Ga-FIB-induced Auger electrons are also in the scope of the apparatus. The 3D microanalysis was applied to a bonding wire on an integrated circuit (IC). A series of EB-induced sample current images of the successive cross sections were obtained as a function of the cross-sectioning position. This result showed the capability to realize the 3D Auger microanalysis. Two-dimensional elemental mapping with Ga-FIB-induced Auger electrons was realized for the first time on the IC surface. Its applicability to surface analysis was evaluated. © *1998 American Vacuum Society.* [S0734-211X(98)04104-3]

I. INTRODUCTION

Development of microelectronic devices requires new microanalysis techniques for three-dimensional (3D) elemental distribution in the evaluation and failure analysis of products. In environmental science, 3D analysis of individual particles of suspended particulate matter (SPM) has been recognized to be effective in deducing pollution sources and formation processes of the SPM. Information on not only the chemical identity but also their spatial distribution can be provided by the 3D analysis.

Combination of layer-by-layer ion beam etching of a sample surface and two-dimensional (2D) elemental mapping of the crater bottom produces the 3D analysis is called "image depth profiling."[1] With the use of the secondary ion mass spectrometry (SIMS) in the 2D mapping of the crater bottom, the image depth profiling has been applied to the analysis of impurities with concentration in ppb level.[2] In this method, the depth from the initial surface is determined from the total sputtered depth by assuming a constant sputtering rate. Thus the image depth profiling is limited to be applied to samples of homogeneous matrix with flat initial surfaces. Furthermore, it is well known that the depth resolution gradually deteriorates as the sputtered depth increases due to the roughening effect in sputter depth profiling with ion bombardment. In case of samples like SPM particles and integrated circuits (IC) with nonflat surfaces or heterogeneous matrix, the depth profiling suffers from uneven sputtering because of the different local sputtering rates which depend on incident angles and target species even when the intensity of the ion beam is kept constant. Preferential sputtering of analyzing surface and implantation of primary ion species during etching also deteriorate quantification of these 3D analysis techniques.

In order to realize 3D analysis for samples with arbitrary shape or of heterogeneity and to improve the quantification in the SIMS analysis, we designed an ion and electron dual focused beam apparatus. The following analysis methods were considered. (1) Three-dimensional Auger mapping: In this method, a 3D elemental map is obtained by repeating the cross sectioning of a sample with a gallium focused ion beam (Ga FIB) and 2D Auger mapping of the cross section with the electron beam (EB). The high depth resolution in this method is achieved by the precise control of the coordinates of the Ga FIB. (2) Application of postionization with EB to the Ga-FIB SIMS: The EB is utilized as a continuous ionization source and focused on the spot which is bombarded by the Ga FIB simultaneously. High production of ions is then expected in the dynamic analysis and will improve its quantification. (3) Two-dimensional elemental mapping with Ga-FIB-induced Auger electrons: With the use of the Auger electrons emitted into the vacuum from the sputtered species, elemental mapping is performed. In this article, the concepts of these three analysis techniques, together with the apparatus itself, are introduced. The 3D imaging is reported as an initial stage of developing the 3D Auger mapping. The result of 2D elemental mapping with Ga-FIB-induced Auger electrons is also described.

[a]Electronic mail: chengzhh@iis.u-tokyo.ac.jp

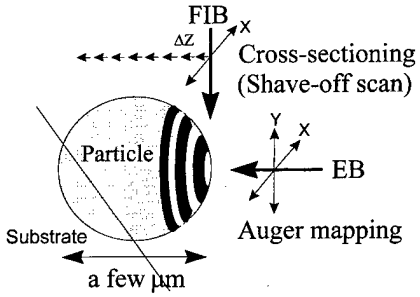


FIG. 1. Concept for the three-dimensional Auger mapping.

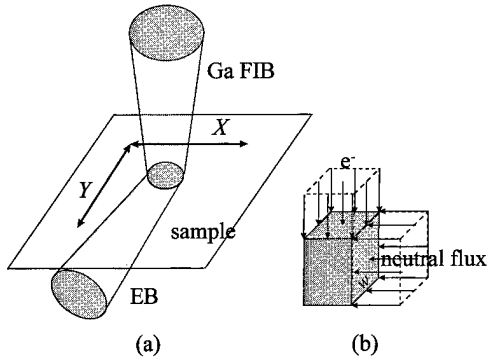


FIG. 2. Application of postionization with EB to the Ga-FIB SIMS.

II. CONCEPTS OF THE NEW MICROANALYSIS METHODS

A. Three-dimensional Auger mapping

Conventional 3D analysis or the image depth profiling is performed by the repetition of etching a surface with a raster-scanning ion beam and 2D elemental mapping with SIMS or scanning Auger microprobe. This method requires samples with flat and smooth surface and a constant erosion rate within the analyzing volume. However, most of the electronic devices and SPM particles have heterogeneous matrix and irregular facets. Therefore, the image depth profiling is only applicable to some favorable cases. Some techniques have been developed to improve the depth resolution and quantification, e.g., (i) postprocessing of 3D raw data for correcting the depth scale,[3] (ii) using a Cs^+ primary ion beam and detecting of CsM^+ ions to reduce the variation in relative ion yield as a function of material composition[4] or (iii) direct measurement of the depth with atomic force microscope.[5] But each technique encounters its own problems to realize true 3D analysis in analyzing uneven crater bottom.

It is crucial to produce flat analyzing surface at "aimed depth" in the image depth profiling. But usual raster scanning ion beam bombardment cannot achieve it, especially for heterogeneous samples or those with irregular facets. We previously developed one-dimensional Ga-FIB SIMS analysis of individual microparticles named "shave-off" analysis.[6] In the shave-off analysis, the horizontal raster by a Ga FIB sweeps very slowly in the vertical direction over the analyzing area. A flat cross section developed by the line scanning of the Ga FIB continuously moves from one edge of the sample to the other within one scanning frame. Gallium atoms implanted into the cross section during the scanning is negligible[7,8] and could not be detected in Auger electron spectroscopic (AES) analysis. We involved the shave-off cross-sectioning technique to develop a novel 3D microanalysis method which is illustrated in Fig. 1. The Ga FIB is used as a micromachining tool. First the cross section is made from the sample edge through the Ga-FIB bombardment in the shave-off mode. Then the nondestructive 2D Auger mapping over the cross section (X-Y plane in Fig. 1) from its normal is carried out with the EB arranged perpendicular to the Ga FIB. Repeating this procedure with the renewal of the cross section along its normal (Z direction in Fig. 1), we can obtain a 3D Auger map. The depth (Z) resolution in this method is higher than the spot diameter of the Ga FIB because only the steep edge of the Ga FIB is used in the micromachining. The lateral resolution is determined by the Auger mapping resolution depending on the spot size of the EB. As the coordinates in a 3D map are exactly determined by the positions of the two beams, samples with arbitrary surface and/or heterogeneous matrix can be analyzed. With the Ga-FIB current of nA order, a volume of about $30\times30\times30$ μm^3 can be cross sectioned in several hours.

B. Application of postionization with EB to the Ga-FIB SIMS

Postionization has been introduced into the SIMS analysis to improve the quantification of this technique. In sputtered neutral mass spectrometry (SNMS) positively charged ions are produced by impact of electrons or photons. SNMS using nonresonant multiphoton ionization (NRMPI) has two important advantages for quantitative analysis: its sensitivity to an element is rarely affected by the surrounding matrix and its sensitivity to different elements has less spread than SIMS.[9] But SNMS using NRMPI can only be utilized in pulse mode and requires long time for large-scale analysis. Elemental depth profiling over depths greater than 100 nm is best performed using a continuous primary ion beam and hence, DC electron beam postionization may be the most efficient method for increasing detection sensitivity.[10]

In our method, the EB is designed to work cooperatively with the Ga FIB. During the Ga FIB is raster scanning over the surface, the EB is focused on the spot which is simultaneously bombarded by the Ga FIB [Fig. 2(a)]. In this condition, the sputtered neutral species will be postionized through electron impact in very close proximity to the surface. As to the degree of ionization α^+ in this case, it is expressed as[11]

$$\alpha^+ = \sigma i_e / v_0 w, \tag{1}$$

where σ is the cross section for ionization by electron impact, i_e the flux of the EB, v_0 the velocity of the sputtered neutrals and w the width of the interaction volume which is normal to the direction of the electron beam and that of the neutral flux, respectively [Fig. 2(b)]. The ionization cross section reaches its maximum at electron energies between 30 and 100 eV except for that of elements with high ionization

potential.[12] Above about 300 eV, the cross section of all elements decreases with roughly the same dependence on the energy E the EB.[13] However, as Gersch *et al.* suggested, the difference in cross section among elements becomes smallest in this high energy region and is quite attractive for quantification in the analysis. In addition, the maximum $i_e(i_{e,\max})$ that can be fed through the interaction volume is proportional to $E^{3/2}$ (Ref. 14). Thus,[11]

$$\sigma i_{e,\max} \sim E^{1/2} \ln(E/U), \tag{2}$$

where U is the surface binding energy, i.e., the ionization probability increases with increasing electron energy. With this method, it is expected to improve the quantification and sensitivity in the surface elemental analysis.

C. Two-dimensional elemental mapping with Auger electrons induced by Ga-FIB

It is known that ion-induced Auger electron (IAE) spectra of light elements, such as Mg, Al and Si, show considerably different characters from those of conventional electron-excited Auger spectra.[15] The Auger decay of excited species induced by energetic projectiles is generally explained by the formation of instantaneous quasimolecules consisting of a projectile and a target atom and/or two target atoms. Through crossing of the molecular orbitals, innershell vacancies are produced. An IAE spectrum is composed of a continuous background related to the decay of excited atoms in the solid (bulklike peak) and sharp peaks, called atomiclike peaks on the background. These sharp peaks originate from excited atoms or ions with inner shell vacancies which are ejected into the vacuum through ion bombardment. By now, most of the work concentrates on the excitation mechanisms of inner-shell electrons in the collision cascades[16] and the related phenomena like particle-surface interactions.[17] Some studies intended to find the usefulness of IAE as detection species in surface analysis.[18–20] In IAE spectrum measurements, Si and Al are the most popular samples because of their high Auger yield compared with heavier elements. Certain identifications of the atomiclike peaks in case of Ar^+ bombardment have already been given by Whaley and Thomas.[16] Our study on Ga-FIB-induced Auger electron emission from Si and Al also shows atomiclike peaks, which exhibit much higher signal-to-background (S/B) ratios than the $L_{2,3}VV$ lines in the electron-excited ones, and narrow widths (typically a few eV) in their integral spectra. These characteristics can be applied to the 2D elemental mapping with high contrast. Combining the characteristics of the atomiclike peaks in the IAE spectra and the fine-focusing property of the Ga FIB would provide us a new elemental mapping technique with sub-μm order lateral resolution.

III. INSTRUMENTATION

In consideration of the new analysis methods mentioned above, an ion and electron dual focused beam apparatus was developed. We introduced a two-lens Ga-FIB column (Eiko Engineering FI-1000) and a quadrupole mass spectrometer (QMS, Hiden Analytical EQ-300) to a scanning Auger microprobe (PHI 590) which was equipped with an EB (LaB_6 emitter) and a coaxial cylindrical mirror analyzer (CMA). A secondary electron detector was mounted onto the apparatus so that the sample surface could be observed through scanning electron microscopy. Sample current imaging could also be performed. The Ga FIB was arranged perpendicular to the EB so that the cross section created by the Ga FIB in the shave-off mode could be analyzed from its normal by the EB. The QMS was adopted so as to set the sample potential to ground which was necessary for simultaneous analysis in different modes like SIMS and AES. The sample stage is mounted on an X-Y-Z manipulator and its normal could be tilted.

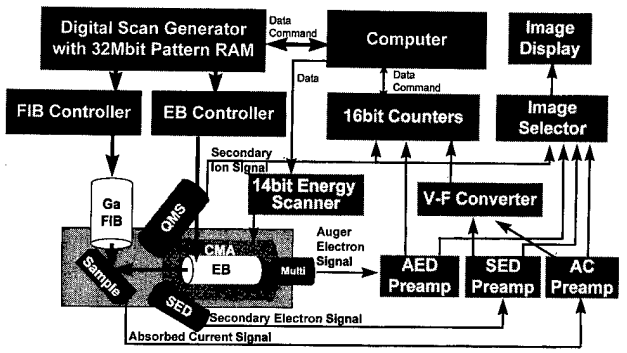


FIG. 3. Block diagram of the ion and electron dual focused beam apparatus.

The control system for the apparatus was designed for the cooperative manipulation of various parameters in the analysis. For example, in the 3D analysis, the cross sectioning with the Ga FIB and the 2D Auger mapping with the EB are carried out alternatively in different scan modes. All new analysis methods introduced in the above sections can be carried out automatically. A block diagram is shown in Fig. 3. The dual beam controller consists of a digital scan generator linked to a personal computer and two independent deflection amplifiers for the Ga FIB and the EB. The scan generator includes a ramp wave generator and a pattern generator with 32 Mbit high speed memory. The pattern memory enables arbitrary scan of the beam on the sample or serves as a frame memory for digital image acquisition. The beams can be controlled manually as well. During the postionization with the EB, for example, as the Ga FIB and the EB are required to scan on the identical point on the surface at any time, one of the two beams can be driven by the pattern generator synchronized with the ramp wave generator. All the signals in the analysis are recorded in a pulse counting mode.

IV. EXPERIMENT

The instrumentation has been completed and the individual analysis modes with the two beams and two spectrometers (QMS and CMA) have been confirmed. While the Ga FIB was able to be focused to 0.1 μm in diameter as expected, the EB could not focus to smaller than 0.3 μm (4.5 keV acceleration, 100 pA). This mainly results from the degradation of electrode insulation and noise from the control

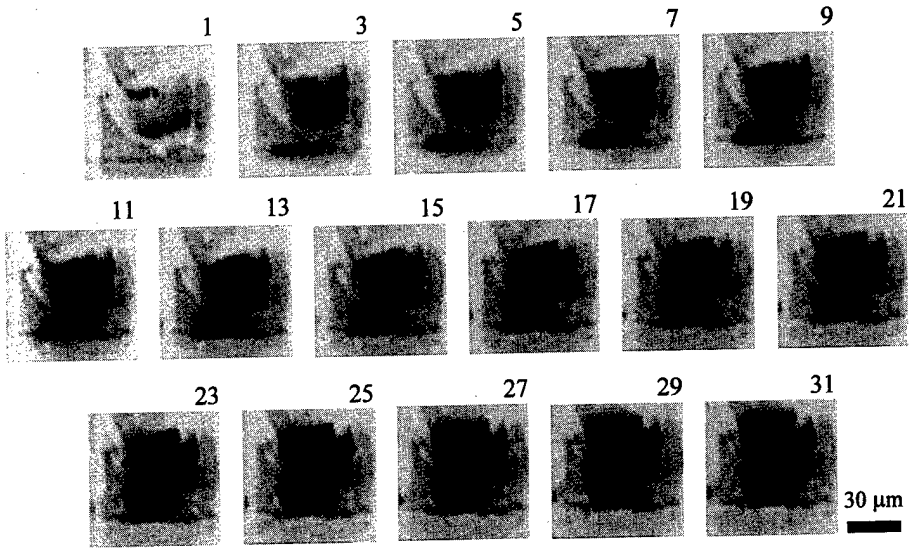


FIG. 4. EB-induced sample current images of the successive cross sections of Al wire and a bond pad; 32 photos have been taken in the experiment. Those with odd numbers are shown in the figure. EB:4.5 keV, FIB: 18 keV, 1.2 nA, analysis time: 410 min.

system which is yet to be further suppressed. Therefore, Auger mapping with sub-μm lateral resolution has not been realized yet. Experiments on the preliminary stage of the 3D analysis, i.e., 3D imaging of successive cross sections with sample current was carried out. The Ga-FIB-induced Auger spectra of Si and Al were also measured and the elemental mapping with Auger electrons induced by Ga FIB was performed.

A. 3D imaging of successive cross sections

The 3D imaging of successive cross sections was performed on an IC (EPROM, 64 Kbits) surface with the Ga FIB for micro-cross sectioning and the EB for imaging. The sample surface was vacuum deposited with Au film of about 0.3 μm thick to avoid charge-up effect from the SiO_2 protection layer on the IC surface. The EB-induced sample current was converted to pulse signals by a V-F converter for mapping. The tip of a bonding wire on the IC surface was selected for the mapping.

In this experiment, after the volume to be analyzed had been selected manually and the scan modes of the two beams and the thickness of each stage in the cross sectioning with the Ga FIB had been input to the computer as parameters, the remaining process was controlled by the computer itself: (a) the Ga FIB was blanked and the EB-induced sample current imaging on the analyzing area was performed with a scanning mode of 256×256 pixels and 150 s/frame; (b) a slice of $30(X)\times 0.94(Z)$ μm^2 (see Fig. 1 for the coordinate system) was sputtered in the shave-off scan mode in 10 min to produce the cross section for the next imaging. Operations a and b were repeated in sequence with the cross section moving in a step of 0.94 μm in Z direction for each cycle. Because of the large cross-sectioning volume, relatively high beam current was adopted. The Ga FIB was used in a beam current of 1.2 nA with 18 kV acceleration. The size of the Ga FIB was then around 1 $\mu m\phi$. The total time for the measurement was about 410 min.

The obtained sample current images of the successive cross sections are shown in Fig. 4. The cross sectioning began from the very tip of the bond wire. Then the cross section moved along its normal (Z direction). Dark areas enlarging during the primary stages are the cross sections. Besides the cross sections of the bond wire, those of the substrate (dark areas at the lower side of each image) were also produced and united to the cross section of the bond wire in later stages. The advantage of the geometric configuration of our apparatus can be seen clearly in the reconstruction of the 3D image. As mentioned above, the coordinates in a 3D map are exactly determined by the positions of the two beams: coordinates X and Y are directly determined by the pixel address within a cross section, i.e., the lateral position of the EB on the cross section which is controlled by the computer. Depth Z is the position of the cross sectioning with the Ga FIB. As a result, a 3D image is simply a serial stack of the images of the successive cross sections in our method. No extra postprocessing of raw data is necessary.

Another advantage of our 3D analysis is that the depth (Z) resolution does not deteriorate with the progress of the cross section. In this experiment, the analyzed volume was $30(X)\times 70(Y)\times 29(Z)$ μm^3, i.e., the depth was in the same order as the lateral dimension. If required, the analysis could be furthered in the Z direction without sacrificing its depth resolution. In other conventional depth profiling methods,[2,21] however, the maximum depth is no more than 10% of its lateral dimension.

B. Two-dimensional elemental mapping with Auger electrons induced by Ga-FIB

IAE spectrum measurements with Ga-FIB bombardment were first performed on Si(110) and polycrystalline Al . The

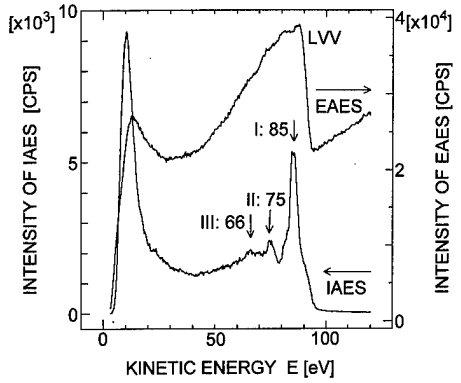


FIG. 5. Ga-FIB-induced Auger spectrum of Si. EAE spectrum is also shown for comparison. Ga-FIB: 20 keV, 4 nA, EB: 5 keV, 12 nA, CMA energy resolution: 1.2%.

samples were set at an angle of 45° both from the Ga FIB and the EB axes. The Ga FIB was used with 20 kV acceleration and about 4 nA in current. Electron-excited Auger electron (EAE) spectra were also measured on the same samples. The EB was operated under the condition of 5 kV acceleration and 12 nA in current. The energy resolution of CMA was 1.2%. The spectra were recorded with 0.35 eV energy steps. Figure 5 shows the Ga-FIB-induced Auger spectrum of Si together with EAE spectrum for comparison. Very sharp atomiclike peaks on a weak continuous background are observed at energies of 85 (peak I), 75 (peak II) and 66 eV (peak III). It is generally accepted that these atomiclike peaks originate from Auger electron emission from sputtered atoms or ions. Because the emission happens in the vacuum (a few angstroms from the surface) and these atoms and ions have discrete electron energy levels, the peak width of each atomiclike peak is narrow (typically a few eV). Although these features were quite different from those of the EAE lines, they were similar to those observed under Ar^+ bombardment.[22]

We applied these features to 2D elemental mapping on an IC (30×30 μm^2) with the Ga-FIB-induced Auger electrons. The Ga-FIB current was reduced to 100 pA and focused to around 0.15 μm in diameter. Because the atomiclike peaks of Si were not observed from protection layer (SiO_2) on the IC surface in a preliminary experiment, two craters were created by Ga-FIB bombardment to expose the Si substrate. Figure 6(a) is a sample current image of the area on which the elemental mapping was performed. Two small areas [at the upper side in Fig. 6(a)] had been heavily sputtered by the Ga FIB so that the elemental Si substrate could be exposed. Figure 6(b) is the IAE spectrum taken over the whole area shown in Fig. 6(a). The atomiclike peaks of Al (a) and Si (c) can be seen in it. In the case of Si, two secondary electron images were taken at kinetic energies c and d indicated in Fig. 6(b), respectively. Then the Si_{LMM} map [Fig. 6(c)] was extracted by subtraction of the images. We obtained the Al_{LMM} map [Fig. 6(d)] with the same procedure. Each elemental map required only 602 s for acquisition. The structure of both the Al wires and Si substrate can be observed clearly in respective elemental maps.

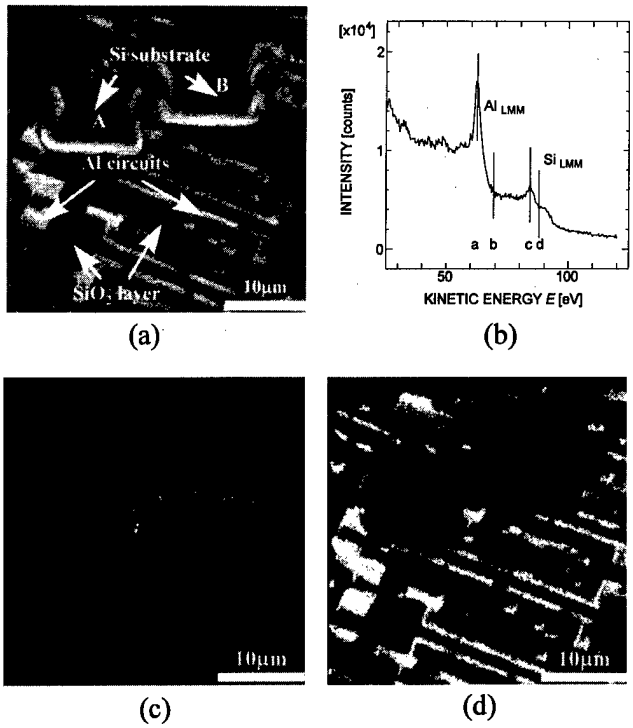


FIG. 6. Two-dimensional elemental mapping of the IC surface with Auger electrons induced by Ga FIB. (a) Sample current image where the elemental mapping was performed. (b) Ga-FIB-induced Auger spectrum of the surface in (a). (c) Si_{LMM} map. (d) Al_{LMM} map.

As shown in Fig. 5, the S/B ratio of the prominent atomiclike peaks in the integral IAE spectra is much higher than those of the LVV lines in the electron-excited ones. This enables us to perform elemental mapping in high contrast. In addition, the beam current of Ga FIB in the above elemental mapping was only 100 pA, so high lateral resolution of sub-μm could be achieved. Furthermore, every elemental map could be obtained within 602 s in our experiment. Charge-up frequently occurs in the conventional Auger analysis. However, in the Ga-FIB-induced Auger analysis, charge-up was not a serious problem. Furthermore, in our experimental conditions, species corresponding to the atomic-like peaks originate mainly from direct projectile-target collisions[23] and the Auger emission area is hardly bigger than the spot area of Ga FIB on the surface. Recently Ga FIB with a diameter of about 10 nm has been realized. With the development of the Ga FIB technique, higher lateral resolution can be expected in the Auger analysis with Ga FIB as a primary probe. In the conventional Auger mapping with electron-excited Auger electrons, intense current is necessary which prevents the decrease of the beam size. Besides, under electron bombardment, the Auger excited area is broader than the spot size.

It should be noticed that the prominent LMM Auger electrons are emitted from neutral excited atoms in the vacuum and their final states are M^+. So study on the relationship between IAE and the secondary ions under ion bombardment coincidentally may provide us a new way to investigate the formation of the secondary ions, preferential sputtering, and

other particle-surface interaction. To evaluate the applicability of IAE to surface analysis, study on its sensitivity and quantification is also an important subject.

V. CONCLUSION

We described a new 3D microanalysis method and dynamic SNMS analysis with electron impact postionization. Ga-FIB-induced Auger emission and elemental mapping with Ga-FIB-induced Auger electrons were also studied. In our 3D analysis, a 3D image is obtained through repetition of the cross sectioning in the shave-off method with the Ga FIB followed by Auger mapping of the cross section with the EB. The depth resolution throughout the analysis is determined by the position of the Ga FIB and does not deteriorate with the renewal of the cross section. This enables the 3D analysis of materials of heterogeneity or with irregular shapes. An ion and electron dual focused beam apparatus was developed on the above concept. With setting up the Auger mapping function, 3D elemental maps will be obtained. Ga-FIB-induced Auger electron emission was also studied on this apparatus and was applied to 2D elemental mapping on an IC. Elemental maps with high contrast and lateral resolution were achieved. To evaluate the applicability of IAE to surface analysis, study on its sensitivity and quantification is necessary.

[1]A. J. Patkin and G. H. Morrison, Anal. Chem. **54**, 2 (1982).
[2]M. Grasserbauer and H. Hutter, in *Secondary Ion Mass Spectrometry SIMS IX*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1993), pp. 545–550.
[3]F. J. Stadermann and H. M. Ortner, in *Secondary Ion Mass Spectrometry SIMS X*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, 1997), pp. 325–328.
[4]Y. Gao, J. Appl. Phys. **64**, 3760 (1988).
[5]G. Gillen, D. Bright, J. Lorincik, and S. Mogren, in *Secondary Ion Mass Spectrometry SIMS X*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, 1997), pp. 321–324.
[6]H. Satoh, M. Owari, and Y. Nihei, J. Vac. Sci. Technol. B **6**, 915 (1988).
[7]T. Sakamoto, M. Owari, and Y. Nihei, Jpn. J. Appl. Phys., Part 1 **36**, 1287 (1997).
[8]Y. W. Beag, M. Tarutani, K. Min, M. Kiuchi, and R. Shimizu, Jpn. J. Appl. Phys., Part 1 **33**, 2025 (1994).
[9]Y. Higashi and T. Maruo, in *Secondary Ion Mass Spectrometry SIMS IX*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1993), pp. 351–354.
[10]R. W. Odom and C. J. Hitzman, in *Secondary Ion Mass Spectrometry SIMS IX*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1993), pp. 593–598.
[11]H.-U. Gersch and K. Wittmaack, J. Vac. Sci. Technol. A **11**, 125 (1993).
[12]H. Tawara and T. Kato, At. Data Nucl. Data Tables **36**, 167 (1987).
[13]W. Lotz, Z. Phys. **232**, 101 (1970).
[14]R. Hutter, in *Focusing of Charged Particles*, edited by A. Septier (Academic, New York, 1967), Vol. III, p. 3.
[15]J. Kempf and G. Kaus, Appl. Phys. **13**, 261 (1977).
[16]R. Whaley and E. W. Thomas, J. Appl. Phys. **56**, 1505 (1984).
[17]G. E. Zampieri and R. A. Baragiola, Surf. Sci. **114**, L15 (1982).
[18]S. Valeri, Surf. Sci. Rep. **17**, 85 (1993).
[19]E. W. Thomas, Vacuum **34**, 1031 (1984).
[20]R. A. Baragiola, Radiat. Eff. **61**, 47 (1982).
[21]H. Hutter, Ch. Brunner, St. Nikolov, Ch. Mittermayer, and M. Grasserbauer, Fresenius J. Anal. Chem. **355**, 585 (1996).
[22]O. Grizzi and R. A. Baragiola, Phys. Rev. A **35**, 135 (1987).
[23]K. Wittmaack, Surf. Sci. **85**, 69 (1979).

Key technologies of a focused ion beam system for single ion implantation

T. Matsukawa,[a] T. Shinada, T. Fukai, and I. Ohdomari
School of Science and Engineering, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169, Japan and Kagami-Memorial Laboratory for Materials Science and Technology, Waseda University, 2-8-26 Nishiwaseda, Shinjuku-ku, Tokyo 169, Japan

(Received 15 December 1997; accepted 9 February 1998)

A focused ion beam (FIB) system was remodeled specially to realize single ion implantation (SII) by which we intended to implant an accurate number of ions one by one into ultrafine semiconductor regions. In the SII, single ions are extracted by chopping the ion beam, and one-by-one extraction of ions have become possible by installing an ultrahigh speed amplifier for chopping. In order to achieve accurate detection of each single ion incidence in the SII, detection of SEs emitted upon ion incidence with a high sensitivity and a high signal to noise (S/N) ratio is essential. Signals from the SE detector which synchronize to an instance of chopping are selectively counted to achieve high S/N ratio. Contaminant particles originating from neutralization and scattering of the ion beam are eliminated by sliding the ion source off the beam axis and cutting off the ion beam at the entrance of the FIB's mass separator. © *1998 American Vacuum Society.* [S0734-211X(98)01104-4]

I. INTRODUCTION

As semiconductor device features in very large scale integration (VLSI) are scaled down, a fluctuation in the dopant number in the devices affects the device characteristics significantly. It is reported that fluctuations in the threshold voltage of metal–oxide–semiconductor field-effect transistors (MOSFETs) increase with the decrease in the channel length of the MOSFETs due to the fluctuations in the dopant number contained in the channel region of the devices.[1] For the purpose of suppressing the fluctuation in the total dopant number, implantation of exact number of ions into each active region in ultrafine device structure is essential. Single ion implantation (SII)[2] is now under development by which we intend to implant dopant ions one by one into fine structures until the necessary number is reached. A similar attempt to implant single ions using a focused ion beam (FIB) was made by Woodham *et al.*[3] However, since they do not detect each single-ion incidence during the implantation, their purpose is not to implant accurate number of ions but to control the average number of implanted ions at low ion dose. In the SII, incidence of each single ion is monitored just after the implantation into the target. Previously we reported that single ions were extracted from the ion beam one by one, and their incidence was detected with 80% efficiency. We reported that the fluctuation in the implanted ion number was expected to be half of that by the conventional implantation under 80% detection efficiency.[4]

Single ions are extracted from a FIB by means of chopping as follows (Fig. 1). A stream of ions are usually blocked by deflecting the ion beam to one side of the objective aperture. Single ions can pass the aperture just when the beam is switched to the other side by applying opposite bias voltage to the deflector. The number of implanted single ions by several repetitions of chopping is estimated by the number of detection of secondary electrons (SEs) ejected upon the ion incidence. In order to control the number of implanted ions accurately in the SII, the number of extracted ions by a single chop must be one or zero. This is satisfied in the condition that the average number of the extracted ions is much less than unity.

Special modification of the FIB column has been made for the SII which utilizes the above-mentioned chopping. In this article, techniques for extracting ions one by one by chopping, counting and detecting single ion incidence, and reducing "contamination," i.e., noise particles induced by neutralized ions are described in detail as the key technologies for the SII.

II. EXTRACTION OF SINGLE IONS BY CHOPPING

Extraction of ions one by one from the ion beam is essential in the SII for accurate control of implanted ion number. In order to prevent extraction of multiple number of ions by a chop, it is necessary that the average number of the extracted ions by a chop (N_{ext}) is much less than unity.[4] The average number of the extracted ions by a single chop is proportional to the opening of the objective aperture and transient time necessary for switching the ion beam deflection, namely chopping. Thus, either the aperture opening or the transient time for chopping should be reduced sufficiently.

The ion beam current on the target is proportional to the opening of the objective aperture. In the course of optimizing the focusing condition, such as alignment of the ion beam trajectory in the FIB column and applied voltage to the condenser and objective lenses, acquisition of scanning ion microscope (SIM) image is essential. Since the ion beam current of at least several pA is necessary for acquisition of the SIM images, we use the objective aperture with the minimum opening of 20 μm in diameter.

In order to reduce the number of the extracted ions by a chop while keeping the sufficiently high beam current for the SIM imaging, an ultrahigh speed amplifier for the beam de-

[a]Electronic mail: matsu@ohdomari.comm.waseda.ac.jp

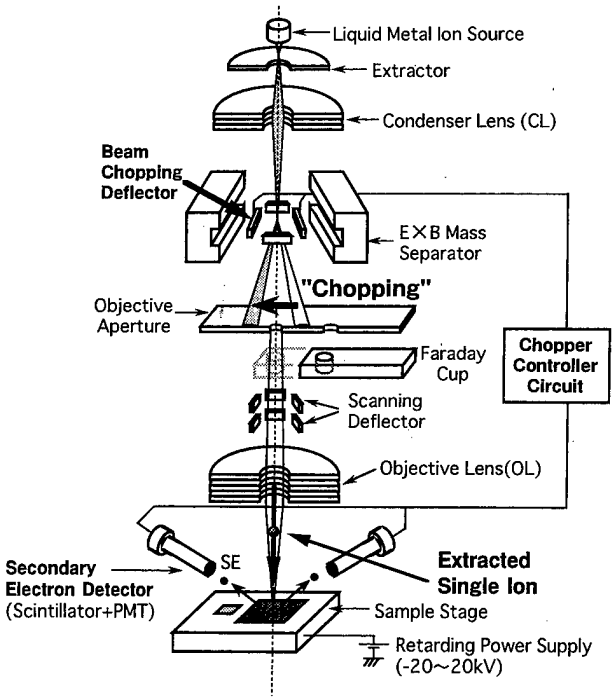


FIG. 1. Optics and principle of SII using the FIB system.

flection was installed by which the rise time of the deflection voltage was 25 ns. Actually, the moment at which the ion beam can pass the aperture is around 50% of the voltage swing, and is shorter than 25 ns. It is necessary not only to set the transient time of the deflection voltage sufficiently short, but also to decrease the flight time of the ions through the chopping deflector much shorter than the transient time of the deflection. In case the time of flight through the deflector is not negligibly shorter than the transient time of the deflection, a shift of the ion trajectory off the beam axis during the time of flight through the deflector could affect the accuracy of positioning the ion incident site significantly. The length of the deflector along the ion beam axis was set at 1 mm and the time of flight through the deflector is 1.56 ns for 60 keV Si ions, which is negligibly shorter than the transient time of chopping.

We have evaluated the number of extracted ions under a number of repetition of chopping by using a track detector CR-39 as the target and by observing the etch pits formed at each incident site of the ions. Details of counting the number of implanted ions by the SII using the CR-39 are described elsewhere.[4] Under the condition that the aperture opening is 20 μm in diameter and the transient time of chopping is 25 ns, we have confirmed that the average number of extracted ions N_{ext} can be less than 0.04 ions/chop, which is so low that the number of extracted ions by each chop is regarded as one or zero.

III. DETECTION OF SINGLE ION INCIDENCE

Accurate detection of the single-ion incidence, namely, detection of SEs emitted upon the ion incidence is essential in the SII. The accurate detection can be achieved by using SE detectors with high sensitivity and high S/N ratio. Two sets of photomultiplier tubes (PMT, Hamamatsu R268) combined with scintillator P47 (Y_2SiO_5:Ce) which are used as SE detectors are installed face to face to maximize the sensitivity against SEs. The position of the SE detectors and electric field around the detectors are optimized to maximize the sensitivity, and the evaluation of the sensitivity is described elsewhere.[5]

We concentrate our discussion on the S/N ratio of the SE detectors. It is well known that PMTs produce noise pulses which are generally referred to as "dark count." Frequency of the dark count affects the controllability of implanted ions in the SII, because we incorrectly recognize the dark count as incidence of the single ions even though there are not any incidence of ions. The single ions extracted by chopping are expected to eject SEs from the target in the certain period from the instant of chopping. In order to reduce the influence of the dark count on the count of incident ions, we have to count the signal produced by the PMTs only in the time window with a certain period after the instant of chopping. A different time window must be defined at a different condition of ion species and incident energy, since the flight time of the ions from the chopping deflector to the target depends on the acceleration energy and the ion mass.

In order to optimize the time window, we measured delay time of the PMT signals from the instant of chopping Si ion beam with energy ranging from 20 to 60 keV. Figure 2 shows the time distribution of the PMT signals obtained by a number of repetition of single ion irradiation. It is shown that the signal pulses due to the SEs emitted upon ion incidence synchronize with chopping, and the average delay of the signal increases with the decrease in the acceleration energy of the ions. In this time of flight measurement it is possible to realize the coincidence method. The flight time of the ions T_{fi} from the deflector to the target is inversely proportional to the velocity of the ions v_i, namely,

$$T_{fi}=\frac{L}{v_i}, \tag{1}$$

$$v_i=\sqrt{\frac{2E}{m}}, \tag{2}$$

where L is the distance from the deflector to the target, E is the acceleration energy, and m is the mass of the ion. In order to obtain universal relation between the acceleration energy and the appropriate time window for the SE detection, relation between the average delay and the velocity of the ions v_i at each acceleration energy is plotted in Fig. 3. Since the average delay increases linearly with the increase of $1/v_i$, the average delay is the sum of the time of flight of the ions and the other origin such as flight time of the SEs and delay in the response of the SE detectors, which correspond to the intercept in Fig. 3. Therefore, the time window can be defined from the average delay estimated from the relation in Fig. 3, and the width of the window is defined as approximately 500 ns which is sufficient to cover almost all the PMT counts due to the SE detection as shown in Fig. 2.

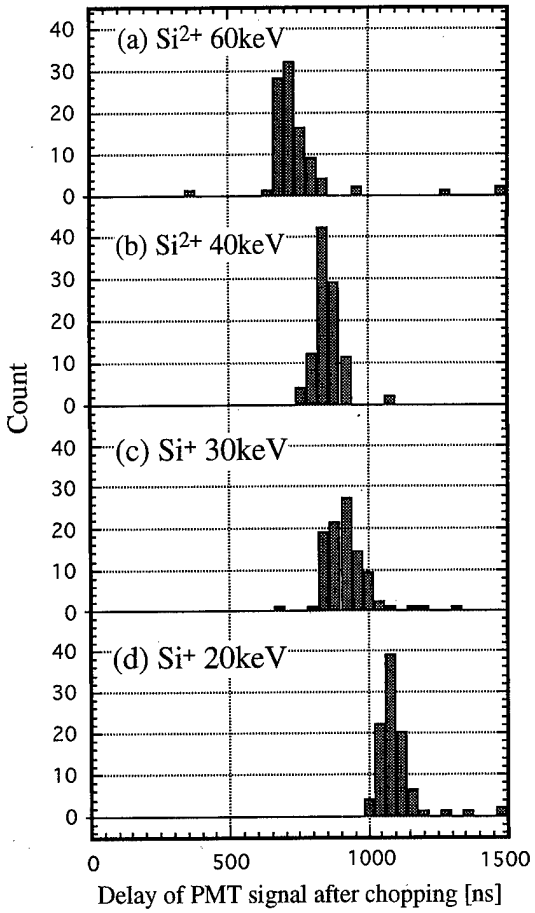


FIG. 2. Time distribution of the PMT pulses measured from the instant of chopping Si ion beam.

Under the width of the time window t_w and the interval of chopping t_{int}, frequency of counting the noise signal is reduced by t_w/t_{int}. For example, if $t_w=500$ ns and $t_{\text{int}}=10$ μs which corresponds to the chopping cycle of 100 kHz, frequency of counting the noise signal can be 1/20 of that counting all the PMT signals. In the current status, S/N ratio of 80 and detection efficiency of 90% are achieved,[5]

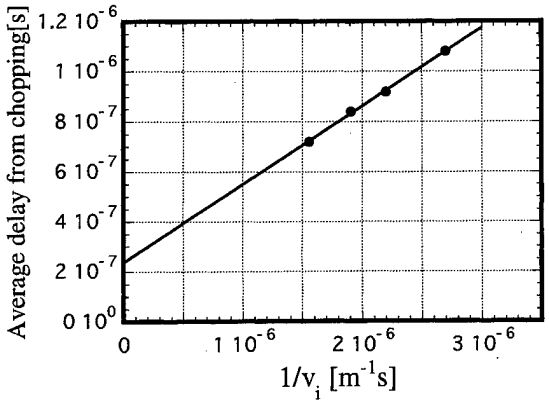


FIG. 3. Relation between average delay of the PMT signal and inverse of ion velocity v_i. The linear increase in the delay with $1/v_i$ is due to the time of flight of ions from the deflector to the target.

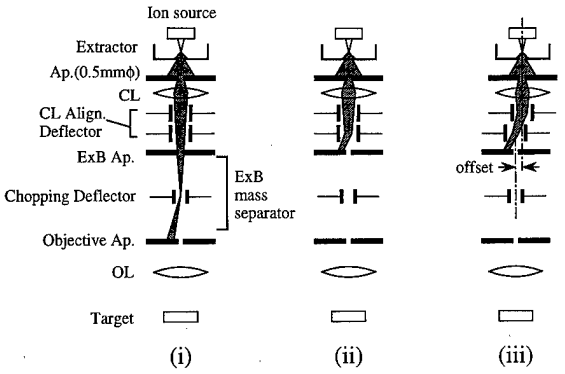


FIG. 4. Three cutoff conditions of the ion beam for evaluating flux of contamination particles.

and the fluctuation in the implanted ion number can be 30% of that by the conventional implantation under these conditions.[4]

IV. ELIMINATION OF CONTAMINATION PARTICLES

It is reported that the contamination induced during the ion implantation process could degrade device characteristics significantly in the recent ultraclean and low temperature device process.[6] In the case of applying FIB to the fabrication of VLSIs, contamination due to the neutralized and scattered particles through the transportation of ions in the FIB column must be taken into account. Especially in the SII in which accurate control of the implanted ion number is required, reduction of the contamination particles during the implantation is essential.

We have evaluated flux of the contamination particles under various conditions of cutting off the ion beam as shown in Fig. 4. In condition (i), the ion beam is deflected by the chopping deflector and cutoff by the objective aperture. In condition (ii), the ion beam is deflected before the entrance of the E×B mass separator, and cutoff by the entrance aperture of the E×B separator. In condition (iii), the ion source is mechanically slid off the beam axis and the ion beam is deflected and cutoff at the entrance of the E×B separator. The reason why the ion source is mechanically slid is that neutralized particles from the ion source is expected to be blocked at the entrance of the E×B separator.

The number of the implanted contaminants under the above-mentioned cutoff conditions is evaluated by using the CR-39 and highly oriented pyrolitic graphite (HOPG) as the target. The CR-39 is sensitive only for the particles with the same order of energy as the acceleration energy. The CR-39 were exposed under the beam axis for 6 h. After the exposure, the CR-39 were immersed in 5N–NaOH solution at 50 °C to form etch pits at each incident position of the contaminants. The number of the etch pits were counted by observing the surface of the CR-39 with atomic force microscopy (AFM). On the other hand, it is reported that "craters" are formed on the surface of HOPG by the incidence of ions with a wide variety of energies, namely, from several tens eV (Ref. 7) to MeV.[8] The HOPG was also exposed under the

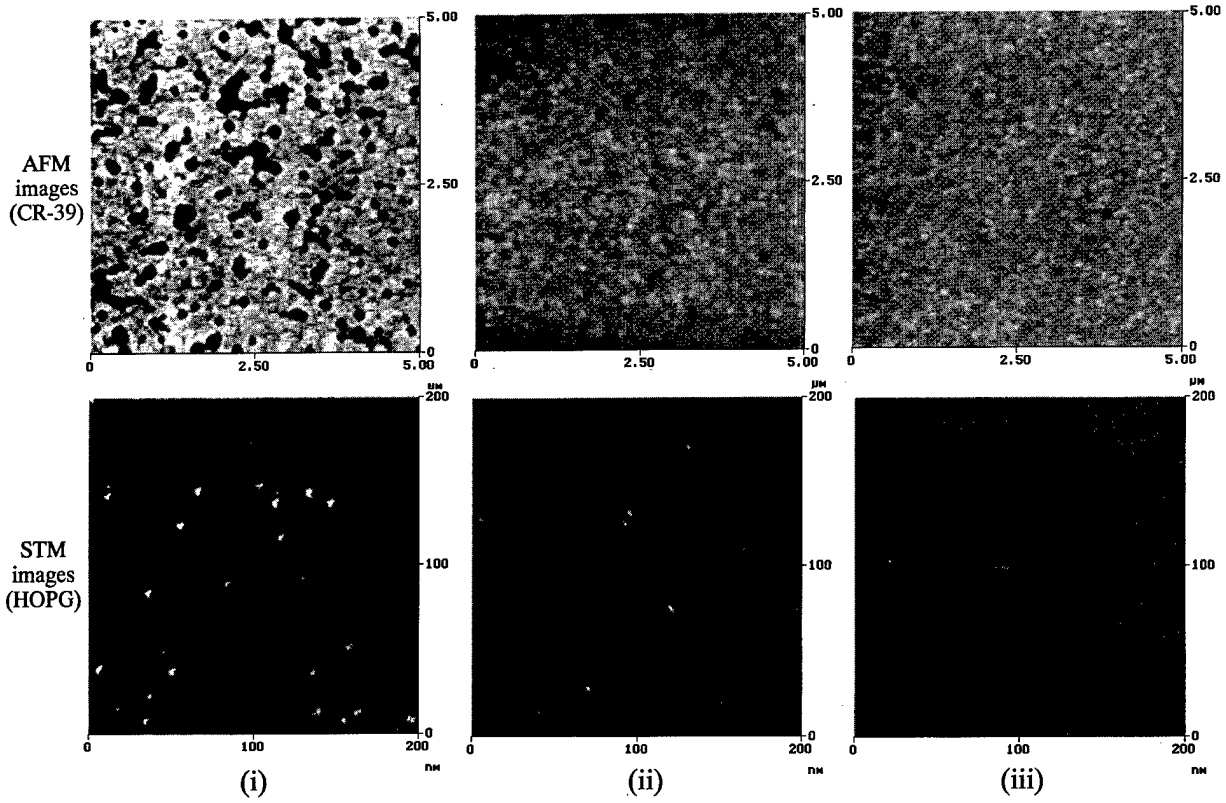


FIG. 5. AFM images of the etch pits on CR-39 and STM images of the craters of HOPG under three different cutoff conditions of ion beam. AFM images are taken with contact mode and scan area of 5×5 μm^2, and STM images are taken with constant height mode and scan area of 200×200 nm^2.

beam axis for 6 h. After the exposure, the surface of the HOPG was observed by scanning tunneling microscopy (STM), and the number of craters induced by the contamination particles was counted.

Figure 5 shows etch pits on the CR-39 observed by the AFM and craters on the HOPG observed by the STM under the three cutoff conditions. As shown clearly in the figure, density of the contamination particle is reduced in order of (i), (ii), and (iii). Average number of the etch pits and craters in an image was estimated by taking a number of AFM and STM images like Fig. 5. Areal density and flux of the contamination particles are calculated from the average number of the etch pits and the craters. The result is summarized in Table I.

Under the cutoff condition (i) in which the ion beam is deflected by the chopping deflector and cutoff by the objective aperture, the flux of the contamination particles is estimated as 4.08×10^4 cm^{-2} s^{-1} from AFM images, which corresponds to a component of the contamination particles with the same order of energy as acceleration energy. On the other hand, the flux estimated from STM images is 2 orders of magnitude higher than that from AFM images of the CR-39. While the CR-39 is sensitive only for the component with rather high energy, a HOPG is sensitive even for low energy component with energy of several tens eV. Therefore, the major component of the contamination particles are low energy ones which cannot be detected by the CR-39. Under cutoff condition (ii), the number and size of etch pits in AFM images of the CR-39 are much less than those under condition (i). Namely, irradiation of the contamination particles with rather high energy can be prevented by cutting the ion beam off before the entrance of the E×B mass separator. The flux estimated from STM images of the HOPG also get one order less than that of condition (i). Taking the difference in the ion beam pass in the FIB column into account, we consider that a major source of the contamination par-

TABLE I. Flux of contamination particles under various cutoff conditions.

Target		Condition (i)	Condition (ii)	Condition (iii)
CR-39	Average number of etch pits in an AFM image	220	6.58	<1
	Flux (cm^{-2} s^{-1})	4.08×10^4	1.22×10^3	$<1.8\times10^2$
HOPG	Average number of craters in a STM image	18.7	3.55	<1
	Flux (cm^{-2} s^{-1})	2.16×10^6	3.88×10^5	$<1.1\times10^5$

ticles is scattering of the ion beam at the entrance aperture of the E×B separator. Under the cutoff condition of (iii), we cannot recognize any etch pits on the CR-39, but sometimes there is a formation of a crater on the HOPG. Since the contrast of the image is rather weak, the contamination particles under condition (iii) seem to have low energy. By sliding the ion source off the ion beam axis, contamination particles due to neutralization at the ion source can be blocked by the entrance aperture of the E×B filter. Therefore, we consider that the contamination particles observed under condition (iii) originate from multiple scattering of the ions in the FIB column.

By sliding the ion source off the beam axis and cutoff the ion beam at the entrance of E×B mass separator, we successfully eliminated the contamination particles, especially the high energy component. The component which cannot be eliminated even under condition (iii) can be blocked by covering a target for implantation with thin oxide film.

V. SUMMARY

Details of the key technologies which are essential for the SII have been described. For the purpose of extracting single ions successfully from the ion beam, the average number of the extracted ions by a single chop must be much less than unity. By installing an ultrahigh speed amplifier for chopping the ion beam, the average number can be set less than 0.04, which is sufficiently low for extraction of ions one by one. For the purpose of detecting single ion incidence into a target accurately, detecting SEs emitted upon single ion incidence with high sensitivity and high S/N ratio is essential. We count only the PMT signals which synchronize with chopping in order to achieve high S/N ratio. In the current status, S/N ratio of 80 is achieved under the SE detection efficiency of 90%. In order to prevent contamination during implementation of the SII, scattered and neutralized particles which arise during the cutoff condition should be eliminated. By sliding the ion source off the beam axis and cutting off the ion beam at the entrance of the E×B mass separator, almost all the contamination particles can be eliminated.

ACKNOWLEDGMENTS

This work has been supported by a Grant-in-Aid for Scientific Research (B) from the Ministry of Education, Science, Sports, and Culture, Japan and by a Research for the Future from the Japan Society for the Promotion of Science (JSPS).

[1]T. Mizuno, J. Okamura, and A. Toriumi, IEEE Trans. Electron Devices **41**, 2216 (1994).
[2]I. Ohdomari, in *Proceedings of the 1st International Symposium on Control of Semiconductor Interfaces, Karuizawa, Japan* (North-Holland, Amsterdam, 1994), p. 233.
[3]R. G. Woodham and H. Ahmed, Jpn. J. Appl. Phys., Part 1 **35**, 6683 (1996).
[4]T. Matsukawa, T. Fukai, S. Suzuki, K. Hara, M. Koh, and I. Ohdomari, Appl. Surf. Sci. **117/118**, 677 (1997).
[5]T. Shinada, Y. Kumura, J. Okabe, T. Matsukawa, and I. Ohdomari, these proceedings.
[6]M. M. Oka, A. Nakada, K. Tomita, T. Shibata, T. Ohmi, and T. Nitta, Jpn. J. Appl. Phys., Part 1 **34**, 796 (1995).
[7]H. Kang, K. H. Park, C. Kim, B. S. Shim, S. Kim, and D. W. Moon, Nucl. Instrum. Methods Phys. Res. B **67**, 312 (1992).
[8]S. Bouffard, J. Cousty, Y. Pennec, and F. Thibaudau, Radiat. Eff. Defects Solids **126**, 225 (1993).

Development of wide range energy focused ion beam lithography system

M. Kinokuni,[a] H. Sawaragi, R. Mimura, and R. Aihara
Eiko Engineering Co., Ltd., 50 Yamazaki Hitachinaka Ibaragi 311-12, Japan

A. Forchel
Technische Physik, Universitaet Wuerzburg, Am Hubland, D-97074 Wuerzburg, Germany

(Received 19 November 1997; accepted 20 April 1998)

A wide range energy focused ion beam system with post objective lens retarding and accelerating field optics has been developed. This system employs two objective lens elements in tandem. The first objective lens element (OL1) is a conventional einzel lens, the second one (OL2) is a retarding and accelerating mode immersion lens. The post objective lens retarding optics is very effective to obtain a low energy fine focused beam. While the post objective lens acceleration optics is not optically attractive, it has other merits. A very high energy ion beam, more than 100 keV, can be obtained by relatively small apparatus. It is also possible to change the landing energy without adjustments of ion beam alignment. The optimized deflector system has been designed for the above mentioned optics. Stitching accuracies of the system were evaluated, and the measured stitching errors are 0.115 μm (2δ) on the X axis and 0.088 μm (2δ) on the Y axis. © *1998 American Vacuum Society.* [S0734-211X(98)06004-1]

I. INTRODUCTION

Focused ion beam (FIB) technology is useful for various kinds of semiconductor processings such as maskless implantation, ion beam lithography, maskrepair, ion beam assisted etching and deposition. For a number of applications, including maskless implantation, the availability of high ion energies, i.e., large implantation depth, is of particular interest. On the other hand, a fine focused ion beam at low energies is required for direct writing ion beam lithography in order to minimize the device damage. But the fine beam cannot be obtained by a conventional FIB column at low acceleration voltage. Because the energy spread ΔE to the total energy E_0 ratio $\Delta E/E_0$ increases, and the chromatic aberration influences the beam size.[1]

Some optical systems with an additional post objective lens retarding field have been reported.[2–4] We have developed more practical FIB system with post objective lens retarding and acceleration field optics, and investigated its performance. In this article, we report the beam characteristics of a focused Ga^+ beam with ion energies between 10 and 160 keV, which were obtained by using a 100 kV FIB system, and report deflection aberrations of this optics and their correction.

II. SYSTEM CONFIGURATION

A wide range energy FIB system with post objective lens retarding and acceleration field optics has been developed. The acceleration voltage of this column is designed to be 100 kV. The specimen stage which is insulated from the chamber can be biased with retarding and acceleration voltage of ± 60 kV. Therefore a fine beam can be obtained at wide range energies between 10 and 160 keV for single charged ions. A schematic illustration of this FIB system is shown in Fig. 1. The focusing column consists of an accelerator lens, two condenser lenses (CL1, CL2), two objective lenses (OL1, OL2), an $E\times B$ mass filter, a stigmator, pre-lens deflectors and others. The two condenser lenses are designed as a zoom condenser lens which permits the continuous change of beam current without changing the beam limiting aperture. Since the beam convergence angle at the sample does not change, the beam diameter does not greatly change.[2] The pattern delineation system (PDS) is performed by a combination of beam vector scanning and an interferometric stage with step and repeat movements. All functions of PDS, e.g., pattern data creation, stage control, dynamic deflection aberration correction, are controlled by DEC alpha station. The maximum scanning speed, registration steps, field size are 0.5 μs/pixel, 4 nm, and 200 μm$\times$200 μm, respectively.

III. POST OBJECTIVE LENS RETARDING AND ACCELERATING OPTICS

A. Objective lens design

In order to produce a low energy and high energy fine focused beam, we have employed a double stage objective lens. The first stage objective lens (OL1) is a conventional einzel lens, the second one (OL2) is a retarding and acceleration mode immersion lens. The arrangement of these lenses is shown in Fig. 2. The OL2 consists of two conical shaped electrodes. A retarding (RET) electrode is the lower electrode of the OL2. The specimen stage, which is insulated from the chamber, is biased to retarding and acceleration voltage, and the RET electrode is provided with the same voltage as the specimen. Since there is no electric field between the RET electrode and the sample, the secondary electron detectors and other hardware can be located in this area. The addition of RET gives the system much more practicality, but it worsens the optical properties. To recover some of these lost optical qualities a subretarding (SRET) electrode is located upstream of the RET electrode. The function of this lens is to control the axial potential distribution of the retard-

[a]Electronic mail: kinokuni@ma4.justnet.ne.jp

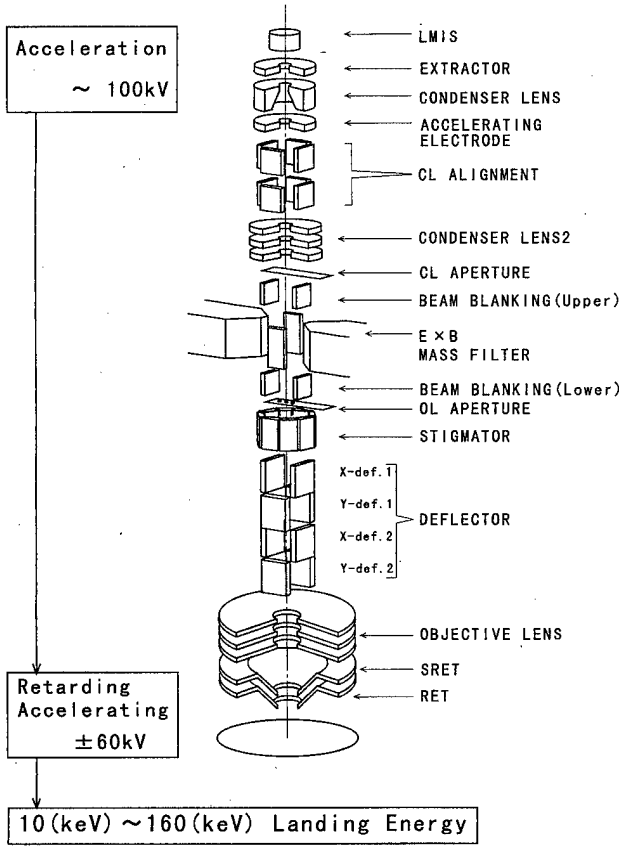


FIG. 1. Schematic illustration of wide range energy FIB system.

ing and acceleration field.[2] Since the chromatic aberration is dominant for small beam convergence angles, the variation of C_c/M is studied as a function of potential of the SRET electrode to optimize the performance, where C_c and M are the chromatic aberration coefficient and the magnification of whole objective lens (OL1 and OL2), respectively. Figure 3 shows this dependence for a series of target potentials in retarding, acceleration and normal modes. It is seen that there are different values of SRET which produce a minimum value of C_c/M at each mode, that value should be selected to obtain the finest beam at each condition. In retarding mode of $E_t=10$ keV (Acc=70 kV, RET=60 kV), the best value of SRET is 27 kV, and in the normal mode of $E_t=70$ keV (Acc=70 kV, RET=0 kV), the best is 43 kV, and in acceleration mode of $E_t=150$ keV (Acc=100 kV, RET=50 kV), the best is 62 kV, where E_t is total landing energy, and Acc is initial acceleration voltage.

B. Investigation of focusing properties

The focusing properties at retarding and acceleration mode were calculated. For the calculation, an angular intensity, a virtual source size (radius) and energy spread of the ion source were assumed to be 20 μA/sr, 30 nm and 15 eV, respectively.

1. Post objective lens retarding mode

The calculated focusing properties for retarding mode are shown in Fig. 4. They show that the post objective lens retarding optics is very effective for producing a low energy fine focused beam by the calculation, and beam diameters are primarily dominated by the initial acceleration voltage.

2. Post objective lens acceleration mode

Figure 5 shows the calculated focusing properties at acceleration, retarding and normal mode. The focusing properties in acceleration mode of 150 keV are worse than those in normal mode of 70 keV. As stated above, beam diameters are primarily dominated by the initial acceleration voltage. Therefore focusing properties are not greatly improved by post objective lens acceleration. On the contrary, the improvements of focusing property are suppressed because of the two following operating constrains. First, the OL1 of this FIB system is operated in the acceleration mode (central electrode negative) when the initial acceleration voltage is less than 70 kV. But it cannot be operated in the acceleration mode when the voltage is more than 71 kV because of insulation problems, since the objective lens needs voltage which is higher than initial acceleration voltage in the acceleration mode. So it is operated in the deceleration mode (central electrode positive), but this mode causes the worsening of properties. Second, the limit of the difference voltage between the RET electrode and the SRET electrodes is 60 kV because the gap is 3 mm. Therefore the ideal value of SRET electrode cannot be selected, and the faculties of SRET electrode cannot be exhibited in the post objective lens acceleration mode of high landing energy.

To obtain a very high energy (more than 100 keV) beam by the conventional FIB column, the scale of the apparatus becomes very large. By using the post objective lens accel-

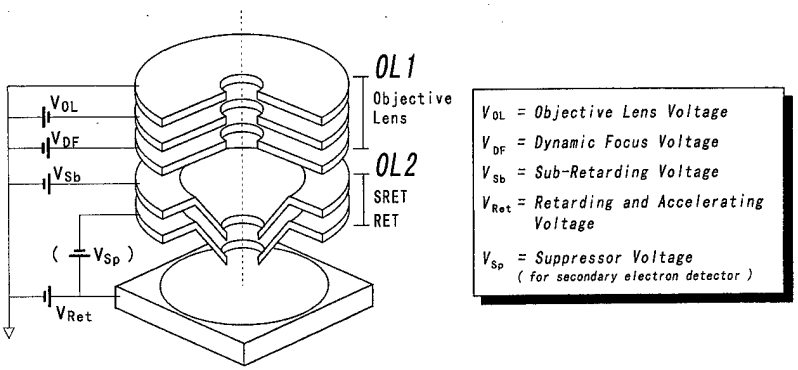


FIG. 2. Schematic illustration of two objective lens elements. The first lens element (OL1) is a conventional einzel lens. The second one (OL2) is a retarding and acceleration mode immersion lens. There is almost no potential variation between RET and the specimen, but suppressor voltage V_{sp}, which is from -260 to 50 V, is practically provided for the secondary electron detection.

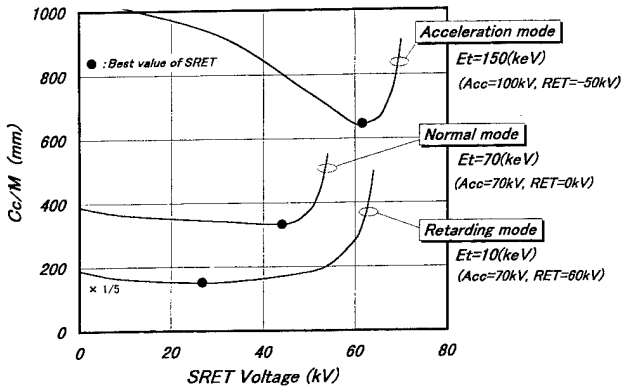


FIG. 3. Relationship between SRET voltage and C_c/M for retarding, acceleration and normal mode. (●) is the best value of SRET in each mode.

eration mode, the very high energy beam can be obtained without the large scale apparatus. It is possible to change the landing energy without adjustments of ion beam alignment.

C. Exposure experiments

In order to evaluate the focused Ga^+ beam diameter which can be achieved by the post objective retarding and acceleration optics, exposure experiments were carried out at landing energies E_t of 10, 70, and 150 keV.[5,6] Periodic lines were exposed GaAs samples coated with 30–80 nm PMMA. A beam current of 10 pA and an exposure time per pixel of 0.5 μs was used, which is equivalent to a line dose of 8×10^7 Ga^+ ions/cm. Figure 6 shows scanning electron microscope (SEM) micrographs for samples coated with PMMA and exposed with different Ga^+ ion energies.

In Fig. 5 the above experimental results are plotted in comparison to the theoretically expected beam diameter. These values are comparatively close to the calculated values. A significant reduction of beam diameter was confirmed from these results. They show the tendency that focusing properties get worse in acceleration mode which are mentioned above. The post objective lens retarding optics is very effective to obtain a low energy fine focused beam. While the post objective lens acceleration optics is not optically attractive by operation constrains, but a very high energy ion beam can be obtained by the relatively small apparatus.

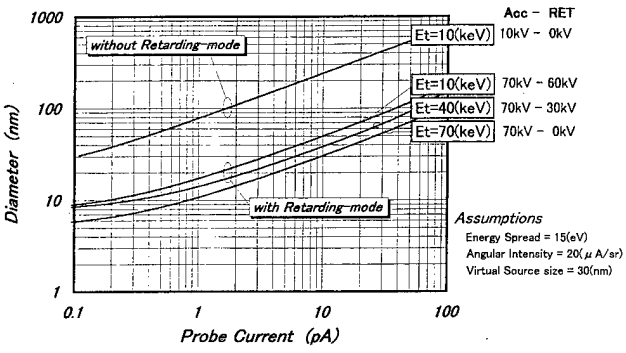


FIG. 4. Calculated focusing properties with and without retarding mode.

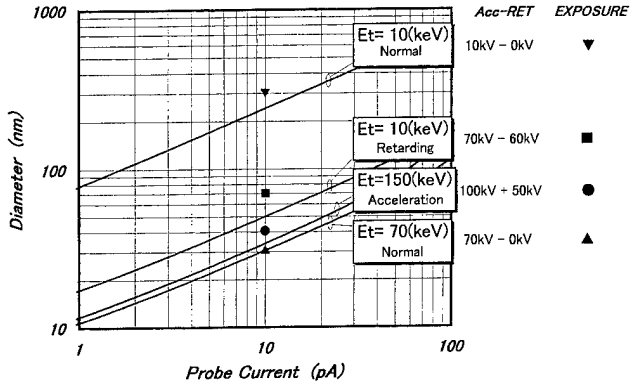


FIG. 5. Calculated focusing properties with acceleration, retarding and normal mode in comparison measured structure sizes.

IV. DEFLECTION ABERRATION AND CORRECTION

A. Pre-lens deflector design

The deflectors must be located above the objective lens by employing the above mentioned optics. In the case of the two objective lenses system, the deflection aberrations of pre-lens deflectors become relatively large compared with conventional post objective lens deflector. Because the deflection field overlaps the focusing field, the deflection aberrations are not only caused by the deflectors but also by the objective lens 1 and 2.

The deflectors which are electrostatic parallel plates were employed instead of an octupole and a quadrupole deflector. Because the octupole deflector needs many hardware to control, and the quadrupole deflector has large deflection aberrations out of beam axis. The electrostatic parallel plate deflectors consist of two stage X deflector (upper-deflector) and two stage Y deflector (lower-deflector) (see Fig. 1). The deflectors have six deflection aberrations, which are coma length, coma radius, field curvature, astigmatism, distortion, and deflection chromatism.[7] Field curvature, astigmatism and distortion can be cancelled by the dynamic correction system (which is mentioned below), but coma length, coma radius and deflection chromatism are not theoretically cancelled, therefore the deflector design needs consideration as these aberrations (coma length, coma radius, and deflection chromatism) have the least value at all operating conditions. The deflection aberrations were calculated for many deflector design proposals (length, gap, potential). The calculated results showed that the deflection design which minimizes each of the aberrations was different at each operating condition, and the influence of these aberrations becomes larger in normal mode (RET=0 kV) than other modes. The deflectors were designed to reduce the influence of these aberrations as little as possible. As will be seen later, the influence of the deflection aberrations can be almost neglected in the case of writing within 200 μm □. As an example, the deflection aberration coefficients of the final X deflector (upper) design are shown in Fig. 7. These aberration coefficients of the X deflector (upper) are almost same as the coefficients of the Y deflector (lower).

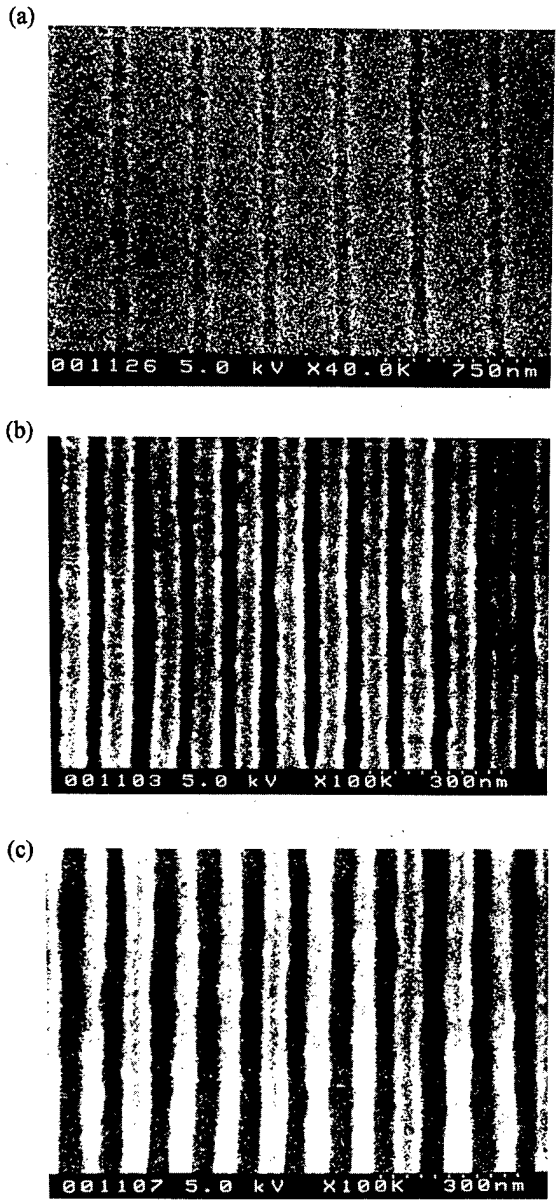


FIG. 6. Periodic line patterns exposed with a beam current 10 pA and a line dose of 8×10^7 Ga^+ ions/cm in 30 nm PMMA for (a) 10 keV, (b) in 80 nm PMMA for 70 keV, (c) in 80 nm PMMA for 150 keV.

B. Deflection aberration correction

As stated above, field curvature, astigmatism and distortion can be cancelled by the correction system. These aberrations of all scanning position are measured before pattern writing by using the ion beam scanned over the referenced mark. These measured data are saved in the memory. In the pattern writing process, these data are fed back to each correction system. The deflection distortion correction data are fed back to each deflector, the astigmatism correction data are fed back to a dynamic stigmator, and the field curvature correction data are fed back to a dynamic focus (DF) lens. As the DF lens, the last element of OL1 is used (see Fig. 2). It is not practical that the middle element of OL1 is used as the DF lens. Because it is very difficult to control rapidly that

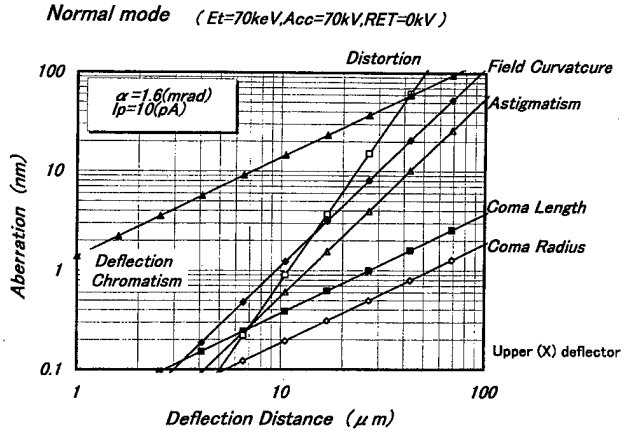


FIG. 7. Calculated deflection aberrations of the X deflector (upper) in normal mode of $E_t=70$ kV (Acc=70 kV, RET=0 kV). α is the beam convergence half angle. I_p is the probe current.

electrode which is on high potential. To control the DF lens with a speed of 0.5 μs/step, which is the maximum scanning speed, a lens voltage of less than ±500 V is desired, because the speed of dynamic focus lens is dominated by the response of the DF lens power supply. As a result of the evaluation, a focusing efficiency of 0.045 μm/V was obtained. The required DF lens voltage is about ±250 V, and this value is low enough to control with the same speed of the beam scanning. By employing the low aberration deflectors and the deflection aberration correction system, an increase of the beam diameter can be effectively suppressed. The increase of the beam diameter in the normal mode by deflection aberrations as the deflection distance is shown in Fig. 8. In the case of writing within 200 μm □, the influence of the deflection aberrations can be almost neglected. But in the case of more than 200 μm □, this influence cannot be neglected.

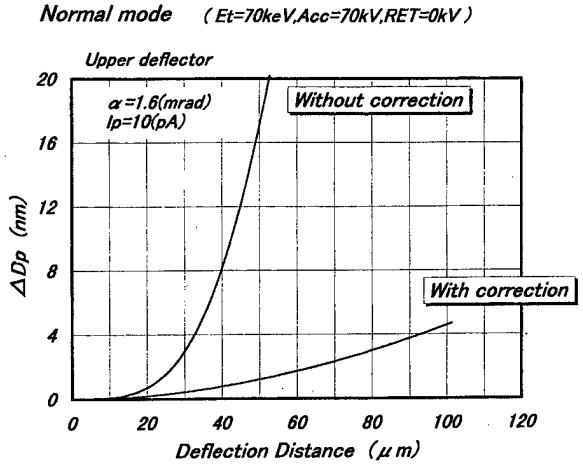


FIG. 8. Increase of the beam diameter by deflection aberrations as the deflection distance in normal mode of $E_t=70$ kV (Acc=70 kV, RET=0 kV). ΔD_p is the increase of the beam diameter. α is the beam convergence half angle. I_p is the probe current.

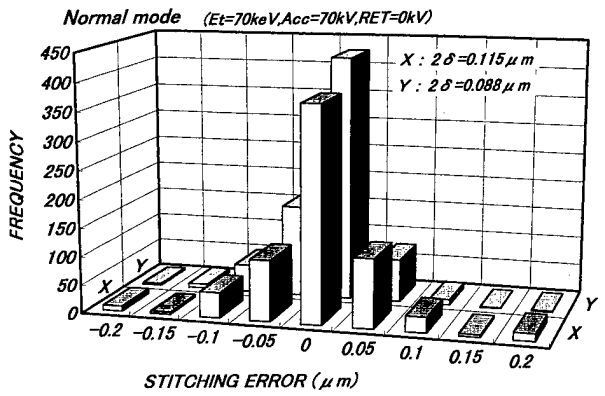


FIG. 9. Obtained stitching accuracies in normal mode of E_t=70 kV (Acc =70 kV, RET=0 kV).

C. Writing accuracies

Stitching accuracies were evaluated by measuring vernier patterns. The first and second level vernier patterns were exposed on GaAs samples coated with PMMA with Ga^+ beam in the three mode which are retarding, accelerating and normal. The measured stitching accuracies in the normal mode of E_t=70 keV (Acc=70 kV, RET=0 kV) are shown in Fig. 9. The stitching error was 0.115 μm (2δ) on the X axis and 0.088 μm (2δ) on the Y axis, respectively. The major contributions to stitching errors are considered to be due to ion beam drift, yawing and pitching motion of the stage movement, and mark detection accuracy.

V. CONCLUSION

A wide range energy FIB system with post objective lens retarding and accelerating field optics has been developed. The post objective lens retarding optics is very effective for obtaining a low energy fine focused beam. The significant reduction of theoretically expected beam diameter was confirmed by exposure experiments. While the post objective lens acceleration optics is not optically attractive, it has other merits. A very high energy ion beam which is more than 100 keV can be obtained by relatively small apparatus. It is also possible to change the landing energy without adjustments of ion beam alignment.

The optimized deflector system has been designed for the above mentioned optics. Since pre-lens deflectors are employed for the two objective lenses system, the deflection aberrations are relatively large compared with conventional post objective lens deflector. Therefore the deflectors were designed to minimize those aberrations for the design constraints (potential, gaps, etc.). In addition, a dynamic correction systems were equipped for realtime deflection aberration correction. By employing the low aberration deflector and the deflection aberration correction systems, the influence of deflection aberrations can be effectively suppressed. Stitching accuracies were evaluated. The measured stitching errors were 0.115 μm (2δ) on the X axis and 0.088 μm (2δ) on the Y axis.

[1]R. L. Seliger, J. Ward, V. Wang, and R. L. Kubena, Appl. Phys. Lett. **34**, 310 (1979).
[2]H. Sawaragi, H. Kasahara, R. Mimura, W. Thompson, and M. H. Shearer, J. Vac. Sci. Technol. B **9**, 2613 (1991).
[3]R. Aihara, H. Kasahara, H. Sawaragi, W. Thompson, and M. H. Shearer, J. Vac. Sci. Technol. B **7**, 79 (1989).
[4]D. H. Narum and R. F. W. Pease, J. Vac. Sci. Technol. B **6**, 966 (1988).
[5]J. P. Reithmaier, A. Keislich, H. Sawaragi, and A. Forchel, Microelectron. Eng. **23**, 119 (1994).
[6]A. Kieslich, J. P. Reithmaier, and A. Forchel, J. Vac. Sci. Technol. B **12**, 3518 (1994).
[7]H. Paik, G. N. Lewis, E. J. Kirkland, and B. M. Siegel, J. Vac. Sci. Technol. B **3**, 75 (1985).

Current status of single ion implantation

Takahiro Shinada,[a] Yoshinori Kumura, Jun Okabe, and Takashi Matsukawa
School of Science and Engineering, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169, Japan

Iwao Ohdomari
School of Science and Engineering, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169, Japan and Kagami Memorial Laboratory for Material Science and Technology, Waseda University, 2-8-26 Nishiwaseda, Shinjuku-ku, Tokyo 169, Japan

(Received 28 January 1998; accepted 24 February 1998)

The current status of single ion implantation (SII) which has been proposed as a novel technology to suppress the fluctuation in dopant number in fine semiconductor structures is reported. The key to control the ion number is to detect secondary electrons (SEs) emitted from a target upon ion incidence. By improving the SE detection system, we have achieved the efficiency of 90% which ensures the reduction in the fluctuation of dopant number to 30% compared to the conventional ion implantation. The improvement for the better SE detection efficiency has turned out to also be effective for the precise beam alignment. The single ion incident position can now be successfully controlled with an error of less than 0.3 μm. © *1998 American Vacuum Society.*
[S0734-211X(98)01004-X]

I. INTRODUCTION

The statistical fluctuation in the dopant number in semiconductor fine structures affects significantly the functions of devices with the minimum feature size much smaller than 0.5 μm.[1,2] So far, the effects of fluctuation on the device performance have been investigated by a theoretical consideration and an experimental study.[1,2] However, as far as we know, it has never been tried to suppress the fluctuation from the viewpoint of the dopant number. Since the conventional ion implantation is conducted by scanning a wide wafer surface with a broad beam of dopant ions, it is essentially impossible to suppress the fluctuation because of a Poisson distribution. For the precise control of the dopant number in semiconductor fine structures, we have developed single ion implantation (SII),[3,4] which is based on the focused ion beam (FIB) system.

The elemental technologies to realize the SII are extraction and aiming of single ions, detection of single ion incidence, fabrication of various liquid metal ion sources (LMIS), and method of heat treatment to electrically activate the implanted ions. Fabrication of fine test specimens is also important to confirm the implantation effect. So far we have been successful in extracting single ions, to develop the LMIS with Au–Si, Cu–P–Pt, and Cu–P–Pt–B alloy, and to fabricate silicon nm wires (6 μm long, 50 nm wide, 80 nm thick). The Si nanowires have been made reproducibly by making use of focused Si ion beam irradiation of a SiO_2 overlayer and the subsequent anisotropic etching of underlying Si with a hydrazine water solution.[5]

The problem yet to be solved is the detection efficiency of SEs, which is the key to "know" the number of incident ions. So far, we have suffered from the relatively low detection efficiency. In this report we first describe the effect of improving the SE detection system. Second, we describe the improvement of aiming precision.

II. CONTROLLABILITY OF SINGLE ION IMPLANTATION

Figure 1 shows the principle of the SII, which is explained in detail in Ref. 4. In order to realize the SII, we must achieve the following three conditions: (1) extraction of single ions by chopping; (2) detection of the SEs ejected from the target upon single ion incidence to count the implanted ions; (3) high precision aiming of a single ion incident site.

Among these the single ion extraction has almost been achieved. We have defined the singularity S as the probability that the number of extracted ion by chopping is one. The number of ions extracted by chopping obeys a Poisson distribution; for example, S is estimated to be only 60% at the average number of extracted ions by a single chop N_{ext} (1 ion/chop), which means that the probability that the plural

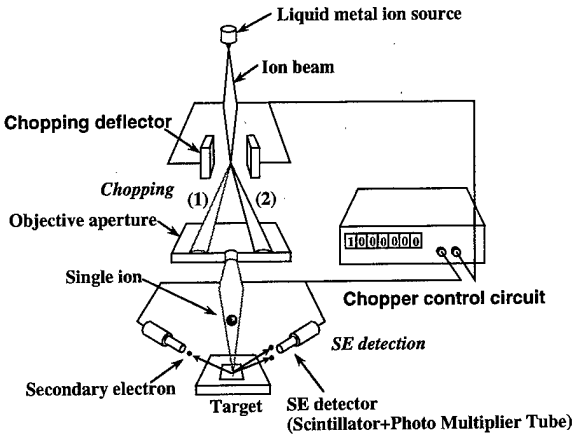


FIG. 1. Principle of single ion implantation.

[a]Electronic mail: shina@ohdomari.comm.waseda.ac.jp

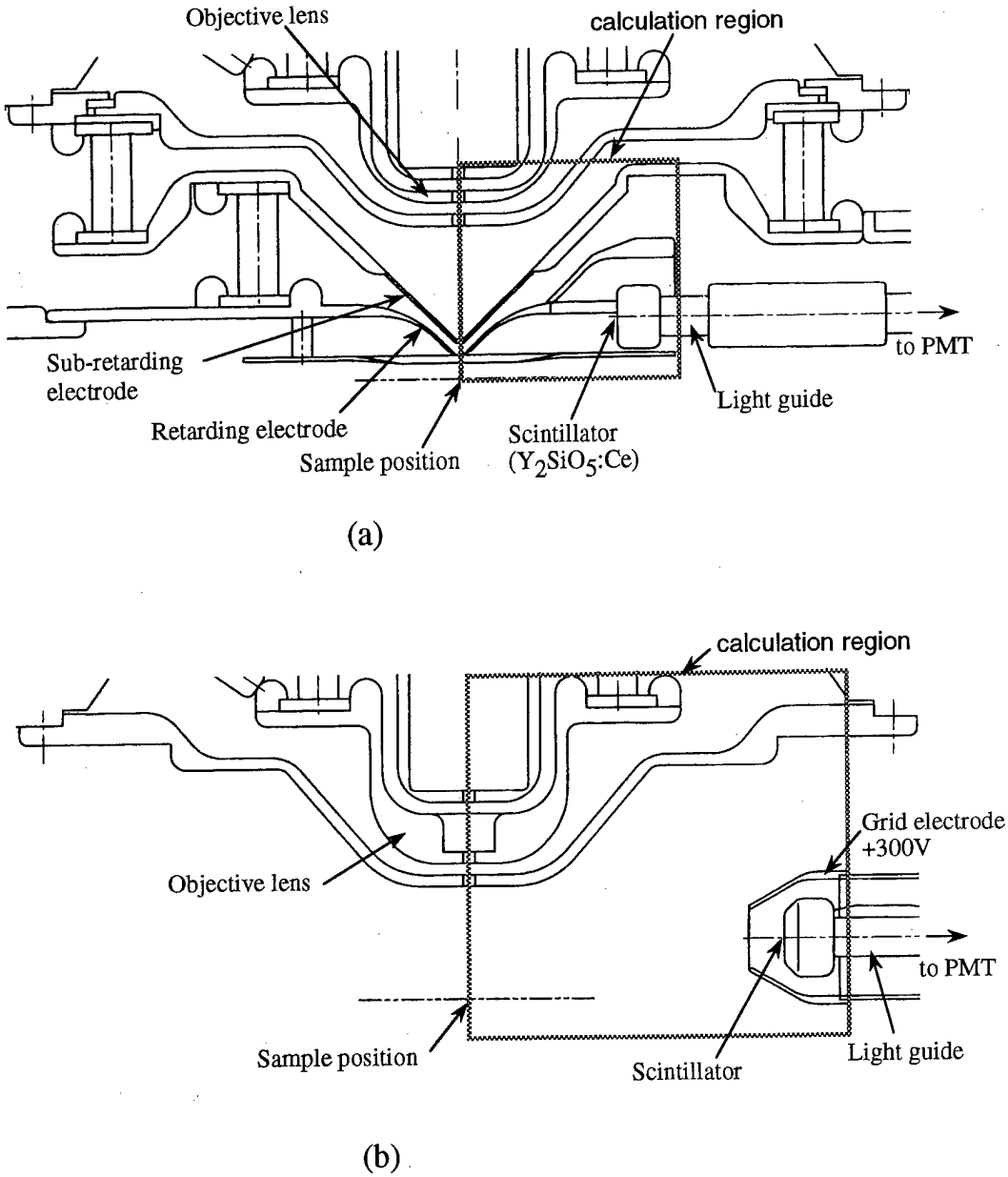


FIG. 2. Ion optics in the vicinity of sample stage of our FIB system. (a) Previous system; (b) new system.

number of ions extracted is 40%. In order to satisfy the first condition, it is necessary that the number of extracted ions by a single chop is sufficiently smaller than one.[4] For example, we have evaluated the number of extracted ions as 0.017 ions/chop under the condition that the ion beam current of 4.0 pA, the transient time of beam chopping is 25 ns, and the opening of the objective aperture is 20 μm. This means that 58.8 chops ensures the extraction of a single ion; namely the number of extracted ions per a single chop (0.017 ions/chop) is sufficiently small for the singularity to be one.

III. IMPROVEMENT OF SECONDARY ELECTRON DETECTION EFFICIENCY

In this section, we discuss the SE detection efficiency which determines the number of ions actually implanted into a target. SE detection efficiency P_d is defined as follows:

$$P_d = \frac{N_{SE}}{N_{ext}}.$$

Here N_{SE} is the average number of the detected SEs by a single chop, and N_{ext} the average number of extracted ions by a single chop. N_{ext} is proportional to the ion beam current and the transient time of beam chopping, and N_{SE} is proportional to N_{ext}. In order to evaluate N_{SE}, we have chosen SiO_2 as a target material which is commonly used as a passivation film of Si surface for ultralarge scale integrated (ULSI) fabrication. A 60 keV Si^{2+} ion beam was chopped with a frequency of 100 kHz and N_{SE} was estimated by dividing the number of SE counts per second by 10^5. To evaluate N_{ext}, we used CR-39 which is well known as a fission track detector. Energetic ions penetrate into the CR-39 film and generate potential damage along the ion

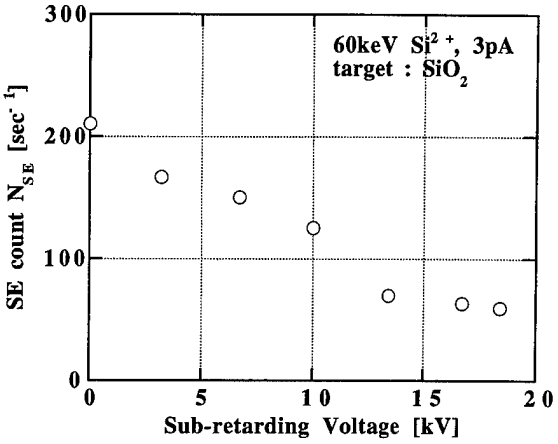


FIG. 3. Subretarding voltage dependence of SE count.

track which is revealed as an etch pit by etching in a heated NaOH solution. Finally N_{ext} was estimated by the relation between the number of repeated chopping and that of created etch pits. Further detail of the method of evaluating N_{SE} and N_{ext} is reported elsewhere.[4]

So far, we have had a rather low SE detection efficiency. Figure 2(a) shows the previous ion optics in the vicinity of the sample stage of our FIB system. A retarding electrode had been set in the vicinity of the sample in order to reduce the effect of chromatic aberration of the condenser lens (CL) and objective lens (OL) at a low acceleration energy. The SEs ejected from the target upon single ion incidence are detected by a photomultiplier tube (PMT) with a scintillator and a light guide. The detected PMT signals are counted by our original chopper control circuit (not shown in the figure). However, SE detection efficiency was very low for the case where a positive bias was applied to the subretarding electrode. The subretarding voltage dependence of SE count per second for SiO_2 is shown in Fig. 3. SE counts were examined under the condition that the ion beam current was 3 pA and the transient time of chopping was 25 ns. The number of the SE count remarkably decreases with increasing the subretarding voltage.

In order to understand the effect of the subretarding electrode on the SE detection efficiency, we have calculated the trajectory of SEs emitted from the sample by solving the two-dimensional Laplace equation. The calculation has been done on the square shown in Fig. 2(a). In this simulation, the subretarding electrode, the retarding electrode, and the scintillator were set to 200 V, 0 V, and 10 kV, respectively. The calculated SE trajectory is shown by the solid lines in Fig. 4(a). The hatched regions indicate the position of electrodes and the sample stage. The initial energy of SEs was set to be 3 eV[6] and the emission angle of SEs was varied between 10° and 90°. It was found that nearly half of the SEs emitted from the sample did not reach the sensitive part of the scintillator because the positive subretarding potential guided the emitted SEs upwards. The reason why four electrons hit the lower part of the scintillator is that the subretarding field strongly expands to the sample stage and the SE detector. It

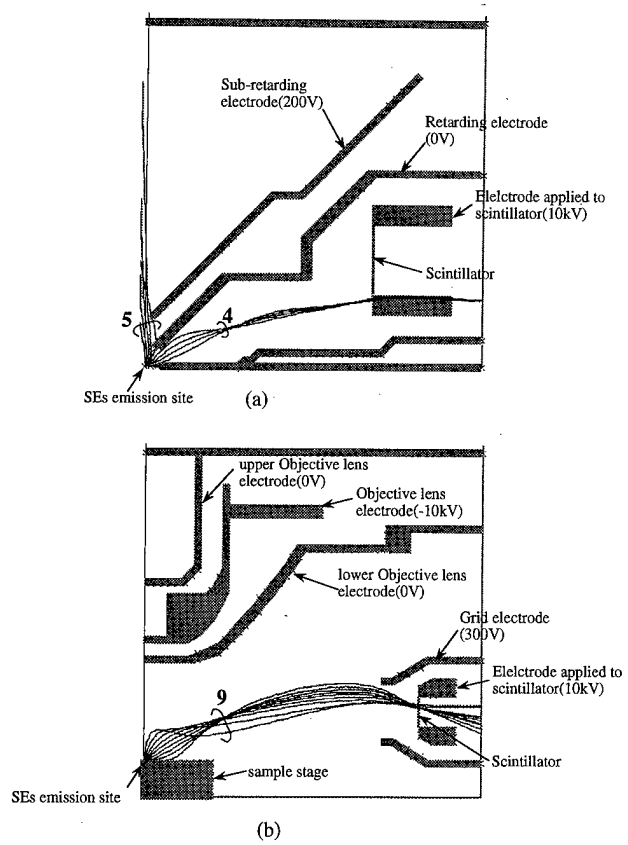


FIG. 4. SE trajectory calculated. The initial energy of SEs was set to be 3 eV and the emission angle of SEs was varied between 10° and 90° at an interval of 10°. (a) Old system; (b) new system.

is clear that the low SE detection efficiency is due to the absorption of the SEs to the subretarding electrode. However, even without using the subretarding, the SE detection efficiency remained at most 80% for SiO_2. This means that there is another reason for the loss in the SE detection. It is likely that the space around the sample stage was too small for SEs to arrive at the scintillator.

To solve these issues, we decided to remove the subretarding electrode. In addition we newly added a grid electrode to the SE detector in order to guide SEs to the sensitive part of the scintillator. The configuration of the objective lens electrode has also been changed to maintain the necessary function of focusing without the subretarding electrode. Figure 2(b) shows the new ion optics for improving the SE detection. The SE trajectory calculated for the new SE detection system on the square shown in the Fig. 2(b) is shown in Fig. 4(b). The scintillator and the grid electrode was set to be 10 kV and 300 V, respectively. The initial condition of SE emission was the same as the previous one. It can be seen that all SEs reached the scintillator.

The 60 keV Si^{2+} ion was implanted into SiO_2 to evaluate the SE detection efficiency in the new system. The results are summarized in Table I. N_{SE}, N_{ext}, and P_d in the previous system are also shown in Table I. The dark count of the PMT was estimated under the condition that ion beam is blocked by the Faraday cup. Since the ion beam current and the tran-

TABLE I. SE detection efficiency.

		New system	Old system
Average number of SE detection N_{SE}(count/chop)	SiO_2	0.01538	0.069
	dark noise	0.000193	0.0008
SN		80	86
Average number of extracted ions N_{ext} [ion/chop]		0.017	0.083
SE detection efficiency N_{SE}/N_{ext}(count/ion)		90	83

sient time of beam chopping were 7.5 pA and 40 ns for the old system, and 4.0 pA and 25 ns for the new system, respectively, N_{ext} evaluated in the new system was smaller than that in the old system and accordingly N_{SE} in the new system was smaller than that in the old system. The ratio of N_{SE}/N_{ext} in the new system was compared with that in the old system under the same condition of S/N ratio. The SE detection efficiency P_d of the single ions has been improved to 90% for SiO_2. There are two reasons for P_d smaller than 100%. One is that the number of emitted SEs upon each ion incidence obeys a Poisson distribution, which means that SEs are not always emitted from the sample upon each ion incidence. Another is that the SE trajectory shown in Fig. 4(b) has been obtained by two-dimensional calculation in which the solid angle spanned by the detector is much higher than the actual three-dimensional configuration. Consequently, the 10% loss in SE detection is due to the nature of SE emission and the inadequate SE detection capability of the new system.

We previously assessed the relation between the detection efficiency and the standard deviation of implanted ion number by SII,[4] and found that the fluctuation in the implanted ion number by SII degraded to the same level as that by the conventional ion implantation at $P_d = 50\%$ but was suppressed completely at $P_d = 100\%$. For the 90% detection, the fluctuation in the dopant number by the SII can be suppressed to only 30% of that by the conventional ion implantation. Thus, the SII is promising for eventually suppressing the fluctuation in dopant number.

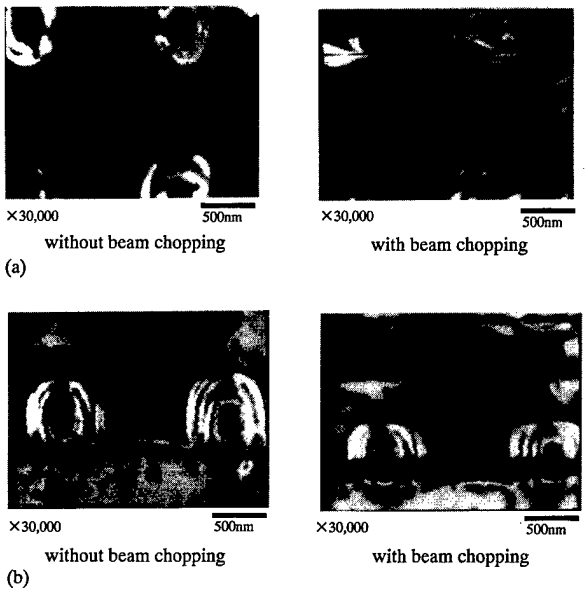


FIG. 5. SE images of a target for beam alignment. (a) Old system; (b) new system.

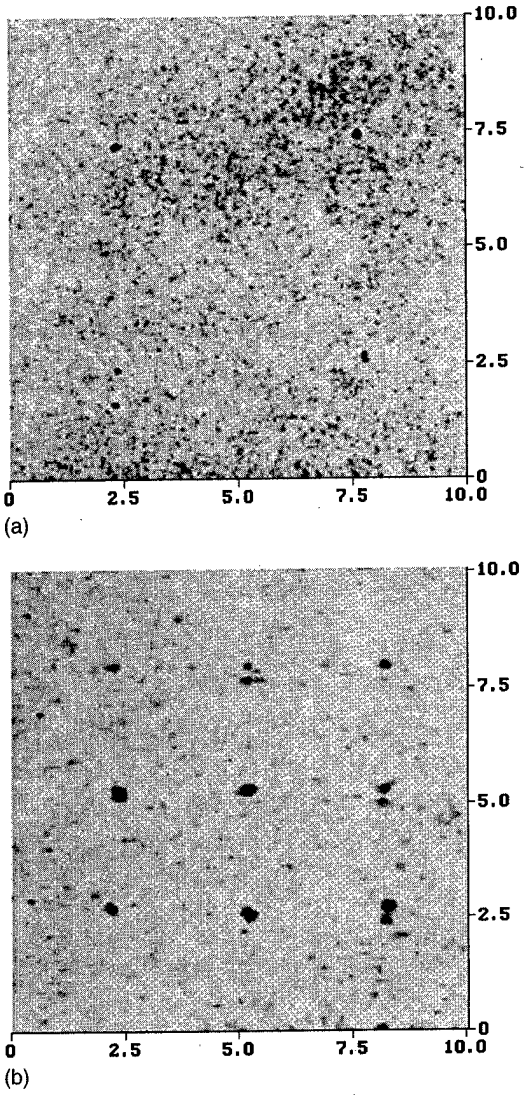


FIG. 6. AFM images of the etch pits in CR-39. (a) Old system; 60 keV Si^{2+} ion with beam current of 3 pA was implanted at an interval of 5 μm and the repetition of chopping was 300 chop/point. (b) New system; 60 keV Si^{2+} ion with beam current of 4 pA was implanted at an interval of 3 μm and the repetition of chopping was 100 chop/point.

IV. IMPROVEMENT OF AIMING PRECISION

The effect of the modifying SE detection system on the aiming precision has been tested. The left-hand side image of Fig. 5(a) shows a SE image of a target which is used for beam alignment. The dot shown in the center of Fig. 5(a) is a hole made by FIB sputtering to evaluate the beam diameter. The SE image during chopping in the old system is severely distorted as shown in the right-hand side image of

Fig. 5(a). Since the distortion synchronized with the beam chopping, the deformation is not due to the external fluctuation such as the mechanical vibration. The SE image is distorted severely because the cross-over point was not precisely aligned at the center of the chopping electrode. Figure 5(b) shows the SE images taken in the new system. The hole for beam diameter evaluation is made elsewhere. Irrespective of with or without beam chopping, the SE images are not distorted. This means that the cross-over point was precisely set at the center of the chopping electrode. This is because the mechanical misalignment originally existed in FIB optics has been moved away by removing the subretarding electrode.

Figure 6 shows AFM images of the etch pits in CR-39 obtained to estimate N_{ext} for the old and the new systems. For the old system the 60 keV Si^{2+} ion was implanted at an interval of 5 μm and the repetition of chopping was 300 chop/point. One or two etch pits at each ion incident site observed in Fig. 6(a) mean that one or two ions were extracted by repeating chopping 300 times. It can be seen that some of the etch pits deviated from the aimed position considerably. The accuracy of aiming was nearly 1 μm because of the fluctuation of ion beam orbit due to chopping, and is not adequate for aiming semiconductor fine regions. Figure 6(b) shows an AFM image of the etch pits obtained in the new system. The Si^{2+} ion beam was implanted at an interval of 3 μm and the repetition of chopping was 100 chop/point. It is clear that the deviation of the ion incident site is greatly suppressed. The single ion incident position can now be controlled with an error of less than 0.3 μm.

V. SUMMARY

The current status of the SII has been reported to control the number of dopant atoms in semiconductor fine structures. The SE detection efficiency has increased to 90% by improving the SE detection system. It can be expected to suppress fluctuation in the dopant number to only 30% by the SII compared with the conventional ion implantation, under the condition of 90% SE detection efficiency. The modification of FIB optics for the better SE detection efficiency has turned out to also be effective for the precise beam alignment. The single ion incident position can be successfully controlled with an error of less than 0.3 μm at present.

ACKNOWLEDGMENTS

This work has been supported by a Grant-in-Aid for Scientific Research (B) from the Ministry of Education, Sports, Science and Culture, Japan and by a Research for the Future from the Japan Society for the Promotion of Science (JSPS).

[1]K. Nishinohara, N. Shigyo, and T. Wada, IEEE Trans. Electron Devices **39**, 634 (1992).
[2]T. Mizuno, J. Okamura, and A. Toriumi, IEEE Trans. Electron Devices **41**, 2216 (1994).
[3]I. Ohdomari, in *Proceedings of the 1st International Symposium on Control of Semiconductor Interfaces, Karuizawa, Japan* (North-Holland, Amsterdam, 1994), p. 223.
[4]T. Matsukawa, T. Fukai, S. Suzuki, K. Hara, M. Koh, and I. Ohdomari, Appl. Surf. Sci. **117/118**, 677 (1997).
[5]T. Shinada, H. Kimura, Y. Kumura, and I. Ohdomari, Appl. Surf. Sci. **117/118**, 684 (1997).
[6]M. Rosler, W. Brauer, J. Devooght, J.-C. Dehase, A. Dubus, M. Cailler, and J.-P. Ganachaud, *Particle Induced Electron Emission I*, Springer Tracts in Modern Physics 122 (Springer, Berlin, 1991).

Chemically and geometrically enhanced focused ion beam micromachining

P. E. Russell,[a] T. J. Stark,[b] D. P. Griffis, J. R. Phillips, and K. F. Jarausch
Materials Science and Engineering Department, Analytical Instrumentation Facility, North Carolina State University, Raleigh, North Carolina 27695

(Received 17 April 1998; accepted 1 June 1998)

Improvements in focused ion beam (FIB) material removal rates utilizing geometric and chemical enhancement were investigated. Geometrical optimization of FIB micromachining of Permalloy and diamond was investigated to determine the magnitude of material removal rate gains that could be attained by increasing the angle of the ion beam with respect to the sample surface normal. The combination of geometrical optimization with chemical enhancement (C_2Cl_4 for Permalloy and H_2O and XeF_2 for diamond) was then investigated to determine whether additional gains in material removal rate could be attained. FIB sharpening of a diamond nanoindenter tip is also presented as a practical example of diamond micromachining with H_2O as the removal rate enhancing species.
© *1998 American Vacuum Society.* [S0734-211X(98)13704-6]

I. INTRODUCTION

The past decade has seen many advances in Ga^+ focused ion beam (FIB) tools including improvements in both ion optics and tool automation. Ion beam optics have improved rapidly to the point where the beam size and profile is no longer the limiting factor in many applications of FIB ranging from failure analysis to manufacturing.[1–4] Developments in FIB can therefore focus on development of techniques which increase removal rates, enhance selectivity, and improve surface finish. This work presents various efforts to improve FIB micromachining (FIBM) through the use of optimized beam, sample geometry,[5–9] and/or the use of chemically enhanced FIBM (CE-FIBM)[10–20] with the goal of maximizing the material removal rate, minimizing material redeposition, and optimizing surface finish.

Chemical enhancements of material removal rates result from chemical reactions ideally initiated by impinging Ga^+ ions. These reactions involve a chemical injected into the vacuum which then adsorbs to the sample surface and the surface constituents of the material being micromachined. Ideally, these reactions both increase the material removal rate and result in a volatile reaction product, thus reducing redeposition of micromachined material. Previous work has shown that for Permalloy[17,18] and carbon based materials,[9,20] C_2Cl_4 and H_2O provide useful improvements in material removal rates. It is also well known that increasing the angle of a sputtering beam with respect to the sputtered sample surface normal results in an increase of sputter (material removal) rate with the rate going through a maximum between 60° and 90°.

In this work, geometrical optimization of FIB micromachining of Permalloy (for example, high aspect ratio and other microelectromechanical systems, magnetorestrictive read/write heads, etc.) and diamond (shaping and sharpening of nanoindenters) is investigated to determine the magnitude of material removal rate gains that can be attained by increasing the angle of the ion beam with respect to the sample surface normal. The combination of geometrical optimization with chemical enhancement (C_2Cl_4 for Permalloy and H_2O and XeF_2 for diamond) is then investigated to determine whether additional gains in material removal rate can be attained. Finally, FIB sharpening of a diamond nanoindenter tip is presented as a practical example of diamond micromachining with H_2O as the chemical material removal rate enhancing species.

II. EXPERIMENT

All samples were micromachined with a 25 keV Ga^+ ion beam using a turbopumped FEI FIB 610 focused ion beam workstation. The various Ga^+ beam diameters and beam currents utilized are described below. A JEOL 6400F cold field emission scanning electron microscope (SEM) was used to obtain all micrographs in this study.

A. Permalloy micromachining

For the Permalloy (80% Ni/20% Fe), three micromachining geometries were used as illustrated in Fig. 1. For all Permalloy micromachining geometries, the Ga^+ beam was scanned over a 10×5 μm area. An approximately 0.13 μm diam, 500 pA Ga^+ beam with a pixel dwell time of 0.8 μs and 50% beam overlap was scanned along the 10 μm axis. For box raster micromachining [(BRM), see Fig. 1(a)], the Ga^+ beam is rastered over the area to be micromachined in a serpentine pattern for 300 s. For box polish micromachining[9] [(BPM), see Fig. 1(b)] and free edge polish micromachining [(FEM), see Fig. 1(c)], the beam is scanned back and forth along the 10 μm axis for 5 s after which the line is stepped 0.13 μm or approximately 1 beam diam along the 5 μm axis. The BPM and FEM patterns required 200 s to complete. Micromachining was performed using the above patterns with and without C_2Cl_4 chemical enhancement.

For C_2Cl_4 chemically enhanced focused ion beam micromachining (CE-FIBM) of Permalloy, a 0.25 mm inner diameter gas jet[20] was positioned to within approximately 100 μm of the sample. After insuring that the sample chamber base

[a]Electronic mail: prussell@ncsu.edu
[b]Also with: Materials Analytical Services.

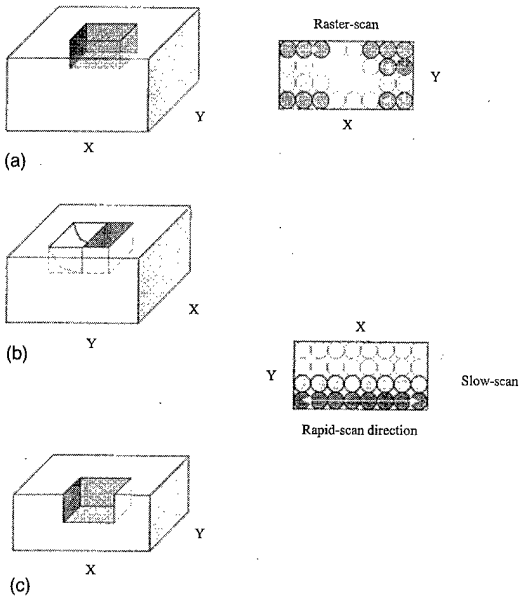


FIG. 1. Illustration of FIBM scan patterns for (a) raster box micromachining (RBM), (b) box polish micromachining (BPM), and (c) free edge micromachining (FEM).

pressure was $<1\times10^{-5}$ Torr, the C_2Cl_4 was directed toward the sample surface through the gas jet. The C_2Cl_4 throughput from the gas jet was regulated using a needle valve such that the total pressure in the sample chamber was 3×10^{-5} Torr.

B. Diamond micromachining

For XeF_2 and H_2O CE-FIBM, the 0.25 mm inner-diameter gas jet was positioned to within approximately 100 μm of the sample. After insuring that the sample chamber base pressure was $<1\times10^{-5}$ Torr, one of the chemical enhancement gasses was directed toward the sample surface through the gas jet. Upon introduction of XeF_2 or H_2O into the sample chamber, the chamber pressure increased to approximately 4×10^{-4} and 8×10^{-4} Torr, respectively. Both final chamber pressures were limited by the geometry of the gas introduction system and the vapor pressures of the respective chemicals.

Diamond BRM [Fig. 1(a)] and FEM [Fig. 1(b)] were performed using an approximately 0.16 μm diam, 1 nA Ga^+ beam with a pixel dwell time of 0.5 μs and 0% beam overlap. All BRM and FEM experiments were performed using no chemical enhancement and both XeF_2, and H_2O CE-FIBM. For BRM, the Ga^+ beam was rastered in a serpentine pattern over 10×10 μm areas for 15 and 7 min. For FEM, two sets of line scan conditions were used. The Ga^+ beam was scanned along a 5 μm line for 14 and 7 s, respectively, and then stepped in 0.16 μm steps, i.e., one beam diameter, for a total of 5 μm.

For the diamond indenter tip shaping, a micromachining lathe jig designed and fabricated at NCSU was attached to the FIB sample stage [see Fig. 2(a)] and a diamond indenter tip was then mounted on the jig. With the lathe jig, the angle

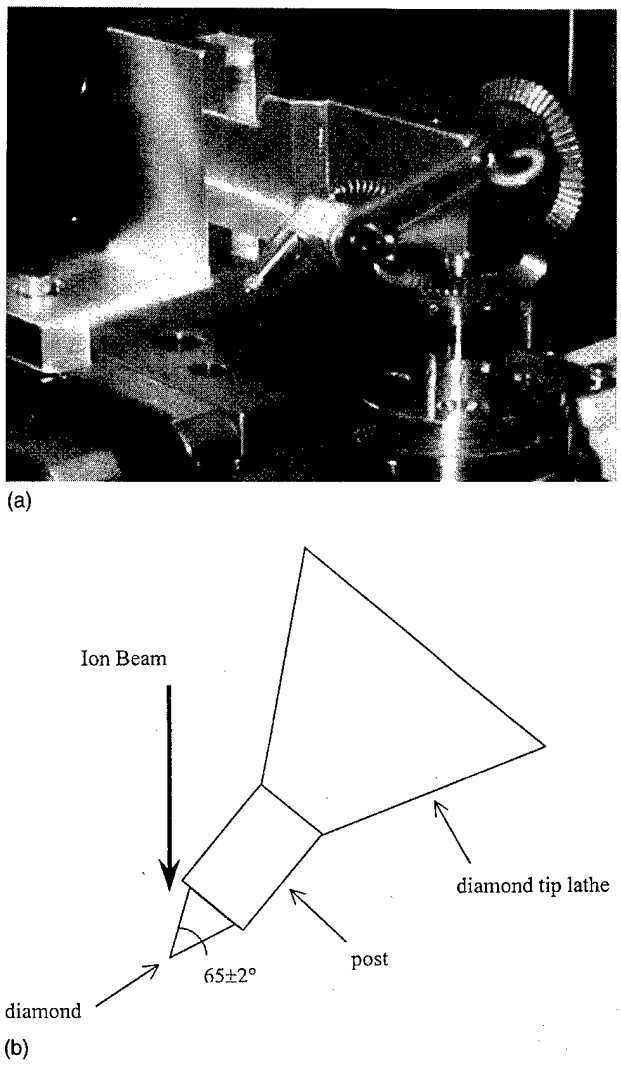
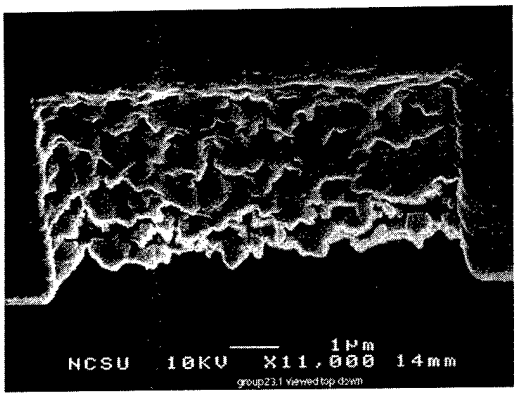


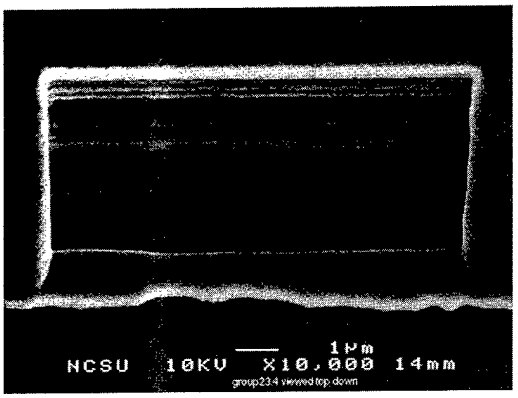
FIG. 2. (a) FIBM lathe jig mounted on the FIB stage and (b) schematic representation of the diamond indenter mounted on the lathe jig indicating angle and direction of the Ga^+ beam.

of the diamond indenter can be set with respect to the Ga^+ beam [see Fig. 2(b)] and the diamond indenter tip can then be rotated as desired (using the FIB stage rotation motor) while maintaining the desired angle. For this experiment, the lathe jig was adjusted to provide a Ga^+ beam angle of 65° (±2°) with respect to the long axis of the diamond indenter tip [see Fig. 2(b)]. The indenter tips were micromachined at 62.5° as described below at 120° rotation intervals to provide the three sided Berkovitch geometry.[21]

The diamond indenter tip was then micromachined both with and without H_2O CE-FIBM (H_2O injection conditions as described above). A rough cut was performed by scanning an approximately 0.13 μm diam, 500 pA Ga^+ beam with a pixel dwell time of 0.5 μs and 0% beam overlap along a 10 μm line. The Ga^+ beam was scanned along this line for 50 s and then stepped 0.13 μm in the direction of the indenter tip. Ten steps were performed covering a distance of 1.3 μm and requiring approximately 8 min. After the rough cuts had



(a)



(b)

FIG. 3. Free edge micromachining (FEM) of Permalloy (a) without chemical enhancement and (b) using C_2Cl_4 CE-FIBM.

been completed, a final polish was performed by scanning an approximately 0.09 μm diam, 100 pA Ga^+ beam with a pixel dwell time of 0.5 μs and 0% beam overlap along a 10 μm line. The final polish line was scanned for 30 s and then stepped 0.09 μm in the direction of the indenter tip. Six steps were performed covering a distance of 0.36 μm and requiring approximately 8 min.

III. RESULTS AND DISCUSSION

A. Permalloy micromachining

Previous studies[17,18] on the micromachining of Permalloy have shown that the material removal rate using 25 keV Ga^+ without chemical enhancement and geometrical optimization is limited by a low erosion rate and topography generated by differential sputtering of Permalloy grains of differing crystallographic orientation [see Fig. 3(a)]. Use of I_2 CE-FIBM dramatically improved the material removal rate, but surface contamination by I_2 and Permalloy containing reaction products limited the utility of this approach to instances where the surface could be cleaned by rinsing with H_2O. Use of I_2 along with geometrical enhancement further improved the material removal rate to the highest yet reported.[18] In an effort to avoid this contamination problem while still maintaining a useful material removal rate, the use of C_2Cl_4, chosen as a source of Cl based on its high Cl/C ratio and on safety issues, was investigated[17–19] for CE-FIBM. While not providing the degree of material removal rate enhancement of I_2, use of C_2Cl_4 with BRM provided a useful increase in material removal rate over FIBM, but most importantly, there was minimal production of micromachining induced topography and no discernable contamination problem. It was also shown[17–19] that optimization of beam/sample geometry can meaningfully increase the FIBM material removal rate, especially when combined with I_2 CE-FIBM. In the present study, micromachining of Permalloy using C_2Cl_4 CE-FIBM combined with geometrical optimization was studied. Micromachining rates are summarized in Table I.

TABLE I. Permalloy micromachining material removal rate summary.

Micromachining technique	Yield (μm^3/nC)	Enhancement factor
BRM	0.10	1.0
BPM	0.25	2.5
FEM	...	...
BRM with C_2Cl_4 CE-FIBM	0.19	1.9
BPM with C_2Cl_4 CE-FIBM	0.72	7.2
FEM with C_2Cl_4 CE-FIBM	0.78	7.8
BRM with I_2 CE-FIBM	0.83	8.3
FEM with I_2 CE-FIBM	3.01	30

While the minimal topography and lack of contamination produced by C_2Cl_4 CE-FIBM are ideal, the practicality of C_2Cl_4 CE-FIBM would be greatly enhanced if the material removal rate can be further improved. In this study, Permalloy BRM with and without C_2Cl_4 CE-FIBM was compared with results obtained by BPM and FEM, with and without C_2Cl_4 CE-FIBM (see Table I). The BPM material removal rate obtained was similar to the previous study.[19] Geometrical optimization using BPM without chemical enhancement provided a factor of 2.5 improvement in material removal rate. It was not possible to accurately measure the material removal rate for FEM due to the extreme topography generated [see Fig. 2(a)]. Combination of BPM or FEM with C_2Cl_4 CE-FIBM resulted in over a factor of 7 improvement in material removal rate. In keeping with previous results obtained using BRM with C_2Cl_4 CE-FIBM in this and in the previous study, the BPM and FEM [see Fig. 2(b)] with C_2Cl_4 CE-FIBM produced cuts having very low topography and no contamination was evident.

B. Diamond micromachining

In a previous study,[20] it was demonstrated that H_2O CE-FIBM dramatically increased the material removal rate for carbon based materials while retarding the removal rate for materials that formed nonvolatile oxides. In this work, the combination of H_2O CE-FIBM and geometrical optimization is investigated.

For RBM, 10×10 μm areas were micromachined with no chemical enhancement, with XeF_2 CE-FIBM, and with H_2O CE-FIBM, respectively. Material removal rate results are summarized in Table II and micrographs of the results of

TABLE II. Diamond micromachining material removal rate summary.

Micromachining technique	Yield (μm^3/nC)	Enhancement factor
BRM	0.09	1.0
FEM	0.16	1.6
BRM with XeF_2 CE-FIBM	0.32	3.5
FEM with XeF_2 CE-FIBM	0.36	4.0
BRM with H_2O CE-FIBM	0.51	5.6
FEM with H_2O CE-FIBM	1.1	12.2

FEM are shown in Fig. 4. From Table II and Fig. 4, it is clear that both XeF_2 and H_2O CE-FIBM provide significant gains in material removal rate versus no chemical enhancement with H_2O CE-FIBM providing the best material removal rate enhancement with the least sputter induced topography. The data in Table II also clearly show that geometrical optimization provides a smaller but meaningful improvement without chemical enhancement.

In the case of FEM with XeF_2 CE-FIBM, the results appear to be consistent with chemical material removal rate enhancement limited by the availability of the enhancing species. For the case of BRM with XeF_2 CE-FIBM, the material removal rate increases by a factor of 3.5 with the use of XeF_2. Although with no CE-FIBM the removal rate increases by a factor of 1.6 when comparing BRM with FEM (i.e., geometrical optimization only), the removal rate when CE-FIBM and FEM are combined is only slightly greater than BRM with CE-FIBM. As previously stated, a possible cause of this lack of a greater increase in the material removal rate may be due to a lack of XeF_2 to react with the diamond surface, i.e., the arrival and sticking to the diamond surface of the XeF_2 does not provide sufficient XeF_2 to react with all available diamond surface atoms. In this experiment, it was not possible to determine if this was indeed the case since the flow rate and thus the availability of the XeF_2 was limited by its vapor pressure. In planned future experiments, the XeF_2 container will be heated in order to provide more XeF_2 to the diamond surface in an effort to test the above hypothesis.

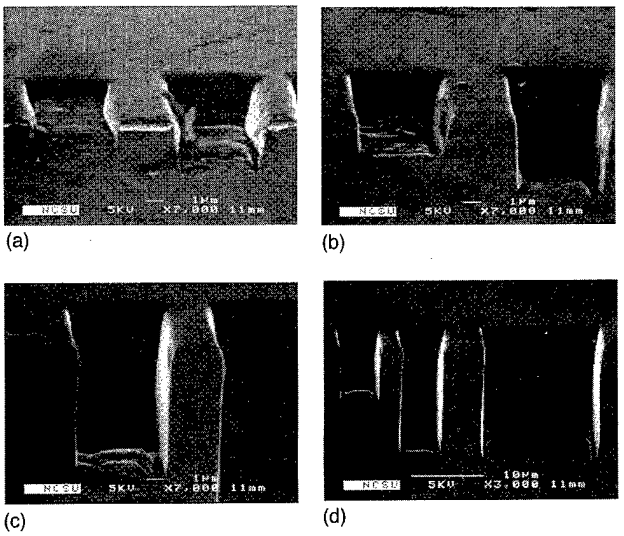


FIG. 4. Free edge micromachining (FEM) of diamond (a) with no chemical enhancement, (b) using XeF_2 CE-FIBM, (c) using H_2O CE-FIBM, and (d) lower magnification view of (c).

H_2O CE-FIBM not only provided the largest increase in material removal rate for BRM H_2O CE-FIBM, but also more that doubled the BRM H_2O CE-FIBM rate when FEM (geometrically optimized) H_2O CE-FIBM was used. In the case of H_2O CE-FIBM, it appears that for the BRM H_2O CE-FIBM case, the available H_2O is not fully utilized and excess water is sputtered away with the diamond. The inherently higher sputter rate due to the optimized sputtering geometry of FEM appears to present a larger number of unreacted diamond surface atoms for reaction with water. This increased number of unreacted surface atoms is then available to react with available H_2O resulting in a material removal rate that is higher than the combined increase obtained with FEM or H_2O CE-FIBM alone. This hypothesis can be tested by limiting the availability of H_2O until the BRM H_2O CE-FIBM rate is slightly reduced. The material removal rate obtained using FEM with H_2O CE-FIBM will then be compared with those of FEM without H_2O CE-FIBM and with BRM with H_2O CE-FIBM in this water limited regime. According to the above hypothesis, the expectation would be that the total enhancement of diamond FEM with H_2O CE-FIBM will be similar to the product of the individual enhancements obtained with FEM without H_2O CE-FIBM and with BRM with H_2O CE-FIBM.

Nanoindentation is becoming an increasingly important tool for the characterization of the mechanical properties of materials on the nanometer scale. Diamond, the hardest naturally occurring material, is an obvious choice for the tips of these indenters. Conventional lapping techniques do not provide sufficient control to produce indenter geometries of the shapes and to the precision required for optimal nanoindentation measurements. A FIB based technique utilizing H_2O CE-FIBM that can be used to modify diamond indenter tips to precisely controlled geometries with improved surface finish is described. As illustrated in Fig. 2, the diamond indenter tip was positioned in the micromachining lathe jig such that the Ga^+ ion beam trajectory was at 65° with respect to the long axis of the indenter post and directed along the post axis past the diamond tip. This orientation of the tip with respect to the Ga^+ ion beam trajectory insures that the tip is not dulled due to sputter damage from the Ga^+ ion beam tails. The diamond indenter tip was them micromachined on three sides as described above and the results are shown in Fig. 5. Figures 5(a) and 5(c) are SEM micrographs of the indenter prior to micromachining. Although the diamond tip has a very small end radius, the geometry is not of the desired Berkovitch[21] geometry and the surface of the tip is very rough. Figures 5(b) and 5(d) show the diamond tip after H_2O CE-FIBM. The desired three sided, 65° Berkovitch geometry has been achieved (angle can be controlled to within 2°) while maintaining an end radius of approximately

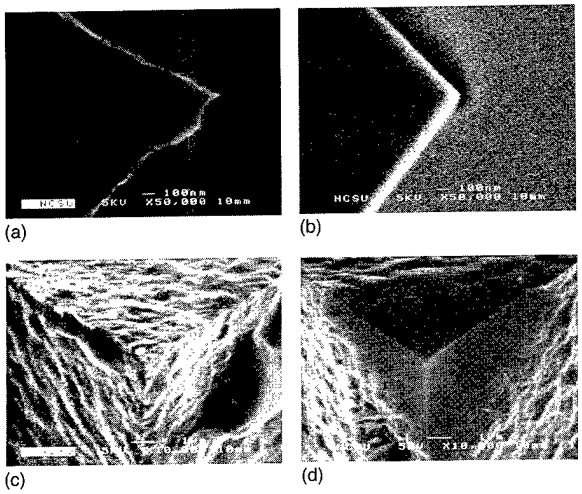


FIG. 5. Micrographs of diamond nanoindenter tip: (a) view from side before H_2O CE-FIBM shaping, (b) view from side after H_2O CE-FIBM shaping, (c) top down view before H_2O CE-FIBM shaping, and (d) top down view after H_2O CE-FIBM shaping.

100 nm. It is also clear from Figs. 5(b) and 5(d) that the surface finish of the diamond has also been greatly improved with surface roughness reduced to less than 10 nm. Comparison of diamond indenter micromachining with and without H_2O CE-FIBM was as expected. Use of H_2O CE-FIBM reduced machining times by a factor of 5 while also reducing surface roughness by a factor of about 10.

IV. CONCLUSION

For Permalloy and for diamond, it has been shown that material removal rates can be increased by employing both CE-FIBM and optimized sample/beam geometry. In the case of Permalloy, use of C_2Cl_4 CE-FIBM provides a useful increase in the material removal rate. While the Permalloy material removal rate is not as high as that provided by I_2 CE-FIBM, C_2Cl_4 CE-FIBM produces a useful material removal rate, generates negligible FIB induced surface topography, and avoids the contamination problems inherent in the I_2 CE-FIBM process. For Permalloy, future investigations will focus on identifying a Br source similar to C_2Cl_4, i.e., a source of bromine that will remain safe and essentially inert until bombarded by the Ga beam on the Permalloy surface. For diamond, it is clear that the use of H_2O CE-FIBM provides clear advantages over both nonchemical enhanced FIBM and over XeF_2 CE-FIBM. It was clearly demonstrated that by employing H_2O CE-FIBM and optimized beam sample geometry, an enhancement of the material removal rate greater than the product of the individual enhancements is obtained. Future diamond work will include continuing investigation of XeF_2 and H_2O CE-FIBM and the respective mechanisms of their interaction with surfaces undergoing micromachining.

[1]T. D. Cambria and N. P. Economou, Solid State Technol. **Sept.**, 133 (1987).
[2]L. R. Harriott, A. Wagner, and F. J. Fritz, J. Vac. Sci. Technol. B **4**, 181 (1986).
[3]J. Melngailis, C. R. Musil, E. H. Stevens, M. Utlaut, E. M. Kellogg, R. T. Post, M. W. Geis, and R. W. Mountain, J. Vac. Sci. Technol. B **4**, 176 (1986).
[4]D. C. Shaver and B. W. Ward, Solid State Technol. **Dec.**, 73 (1985).
[5]K. Gamo, Y. Ochiai, K. Shioyama, and S. Namba, Proc. SPIE **632**, 52 (1986).
[6]Y. Ochiai, K. Gamo, and S. Namba, J. Vac. Sci. Technol. B **3**, 67 (1985).
[7]J. G. Pellerin, D. P. Griffis, and P. E. Russell, J. Vac. Sci. Technol. B **8**, 1945 (1990).
[8]X. Xu, A. D. Della Ratta, J. Sosonkina, and J. Melngailis, J. Vac. Sci. Technol. B **10**, 2675 (1992).
[9]T. J. Stark, D. P. Griffis, and P. E. Russell, J. Vac. Sci. Technol. B **14**, 3990 (1996).
[10]H. Ximen and C. G. Talbet, Proceedings of the 20th International Symposium for Testing and Failure Analysis (Nov. 1994), Vol. 141, pp. 13–18.
[11]R. J. Young, J. R. Cleaver, and H. Ahmed, J. Vac. Sci. Technol. B **11**, 234 (1993).
[12]R. J. Young, Vacuum **44**, 353 (1993).
[13]M. Komuro and H. Hiroshima, J. Vac. Sci. Technol. B **9**, 2656 (1991).
[14]K. Gamo, Y. Ochiai, and S. Namba, Jpn. J. Appl. Phys., Part 2 **21**, L792 (1982).
[15]J. D. Casey, Jr., A. F. Doyle, R. G. Lee, D. K. Stewart, and H. Zimmermann, Microelectron. Eng. **24**, 43 (1994).
[16]Z. Xu, K. Gamo, and S. Namba, J. Vac. Sci. Technol. B **6**, 1039 (1988).
[17]D. M. Thaus, Masters thesis, North Carolina State University, Raleigh, NC, 1996.
[18]D. M. Thaus, T. J. Stark, D. P. Griffis, and P. E. Russell, J. Vac. Sci. Technol. B **14**, 3928 (1996).
[19]D. M. Thaus, T. J. Stark, D. P. Griffis, and P. E. Russell, Appl. Phys. Lett. **68**, 26 (1996).
[20]T. J. Stark, G. M. Shedd, J. Vitarelli, D. P. Griffis, and P. E. Russell, J. Vac. Sci. Technol. B **13**, 2585 (1995).
[21]E. S. Berkovitch, Indust. Diamond Rev. **11**, 129 (1951).

Focused ion beam technology applied to microstructure fabrication

Michael J. Vasile,[a] Raja Nassar, and Jushan Xie
Institute for Micromanufacturing, Louisiana Tech University, P.O. Box 10137, Ruston, Louisiana 71272

(Received 19 November 1997; accepted 1 May 1998)

Focused ion beams (FIBs) have found a place in several research thrusts for the manufacture of mini or micro mechanical objects. This article reports the use of FIB in three distinct applications in microfabrication: prototype structures, micron-sized machine tools and microsurgical manipulators, and ion milling of three dimensional features. Examples of each of these applications are given with the FIB component identified as the enabling or critical component in the technology. The possibility of using FIB milling as part of a production method for micron-sized machine tools is discussed, and the mass production consequences of molds fabricated by three dimensional ion beam milling is also considered. The mathematical procedure and programming steps needed to accurately control FIB three dimensional milling are outlined. © *1998 American Vacuum Society.* [S0734-211X(98)06404-X]

I. INTRODUCTION

Focused ion beam (FIB) technology has potential applications in high volume microstructure fabrication processes as well as in more limited applications such as prototype construction and low-volume, high-value specialty devices. The Institute for Micromanufacturing (IfM) has a focused ion beam program which is devoted entirely to exploring these possibilities, with an emphasis on developing efficient processes which will allow the FIB technique to be established as a useful method of microstructure fabrication. This direction represents a significant departure from the usual applications of FIB technology in the electronics industry, viz, the sectioning of integrated circuits and the preparation of TEM sections.

There are three approaches to micromanufacture fabrication that are being pursued at present: The first is the use of FIB techniques to produce small quantities of prototypes or specialty structures which are key elements in high risk projects, and would require substantial investments to produce by other methods such as microelectromechanical (MEMS) lithographic/etch technologies. Such prototypes would be one or two of a kind, made under circumstances which tolerate long, specialized FIB machining procedures. The object here is to produce or contribute to the production of a prototype for testing and proof of concept: once the viability of the prototype is established, alternative methods of production which were previously too expensive for the risk factor may now be considered. Prototype construction is the most specialized and lowest volume application in the IfM ion beam program.

The second application is the use of FIB methods to produce microscaled tools for ultraprecision, numerically controlled machine tools and micron-scale manipulators for microsurgical research applications. The limits of mechanical machining processes have been dictated by the resolution of the machine tool in terms of spindle run out, and x, y, and z motion control in drilling and milling applications. The ultraprecision machine center at the IfM has broken through the classical limits by employing interferometrically controlled x and y axes, and the utilization of air bearings for support of all three axes. The resulting resolution on the x and y axes is 20 nm. The smallest features that can possibly be machined with such a device then become dependent on the diameter of the tool bit. FIB techniques have been developed to make specialized tool shapes at dimensions that are considerably smaller than commercially available tools. The fundamentals of this technique have been established within the past year. The methods for rapid tool production needed to supply an ambitious ultraprecision machining program are under development. The FIB techniques are an enabling technology in ultraprecision or micromilling application, and the manufacture of tools can be a relatively high cost, low volume process without serious consequences. The production of microsurgical manipulators may also tolerate relatively high cost considering the costs of other factors in this application.

The third application is the use of focused ion beam milling to produce micromolds, which can serve as positive or negative tone masters in a hot-embossing mass production scheme. Micromolds may also be used as masters in molding applications for producing objects such as microlenses or microlens arrays. These applications imply a high degree of dimensional control in the mold fabrication and surface finish smooth enough to satisfy the optical requirements. Three dimensional variation of the depth of ion milling is obviously required to produce a mold for a microlens. The thrust of this effort is to control the FIB milling process so that a three dimensional cavity can be accurately milled to a prespecified geometric shape and dimensions. This part of the FIB program fits into the more general IfM objective of developing microfabrication processes which have control over the depth or Z dimension. The approach taken for FIB milling with depth variation is to write ion beam deflection software that varies the pixel dwell time in a manner that sputters the desired geometric shape.

[a]Electronic mail: mjv@engr.LaTech.edu

This article will give an illustration of each of these techniques: prototype fabrication, microtool fabrication as an enabling technology, and three dimensional control of ion milling. The relationship of the FIB process to the remainder of the processes necessary to obtain the final object will be elucidated. None of these techniques have reached a point where they can be considered a mature technology, and the path for remaining development work will be indicated.

II. EXPERIMENTAL APPARATUS

The FIB instrument used in this work was originally developed by Harriott and co-workers[1] at Bell Labs. It consists of a Ga liquid metal ion source (30 keV maximum energy) employing a single focusing lens and octapole deflection. The extraction aperture used for this work is 150 μm diam, which yields a 3 nA ion beam (Faraday Cup measurement) of 0.4 μm diam full width at half-maximum (FWHM) at 20 keV beam energy. The stage is commercially supplied with 35 and 25 cm motion in the X and Y axes with 1 μm positioning accuracy, 0.13 μm resolution and submicron vibration stability. Material to be ion milled is introduced into the work chamber on cassettes which lock into position on the stage. Loading is done through a vacuum load lock, and the load chamber, work chamber, and ion gun column are all isolated by valves. One of the cassettes has been equipped with an in-vacuum stepper motor acting through a 40:1 reduction gearbox, yielding a full 360° sample rotation apparatus with 0.37° pulse rotational resolution.[2]

Vector scan ion beam deflection is employed and ion milling areas are defined from the superposition of an outlined pixel field over the 512×512 pixel digitized image of the scan. Scan fields typically are 100, 50, or 25 μm squares for ion milling, yielding corresponding pixel sizes of 0.2, 0.1, and 0.05 μm; all of which are smaller than the ion beam FWHM. The ion beam address at any given pixel therefore has a dose component at adjacent pixels, which has significant consequences when ion milling three dimensional features.

III. MICROFABRICATION APPLICATIONS

A. Prototypes

A good example of prototype fabrication is the split microtube shown in Fig. 1. This tube (32 μm o.d.) is made of fused silica and comprises the critical test component in a proprietary application. The inner diameter (i.d.) of the tube is 5 μm (nominal) and the dimensions of the rectangular slit are 6 μm×50 μm. The microtube shown in Fig. 1 was achieved by etching commercially available microtubes to an outer diameter (o.d.) of ~30 μm and then incorporating the microtube into a relatively simple macroscopic test structure. The tube then had to be cut at a distance of 300 μm from the end, to ensure that the new terminal section would include an inside diameter that was not expanded by the etch process. The channel or rectangular slit was then cut (without sample rotation). The entire FIB processing time needed to produce one of these microtube prototypes was 3 h. Four such microtubes were processed in a single loading, with variations in the channel dimensions. Successful prototype operation was established after the ion beam machining was completed, and a total of twelve attempts were required to produce the functioning elements.

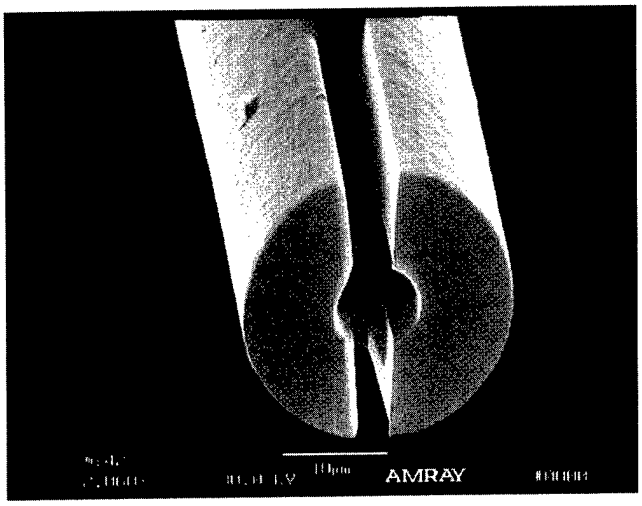


FIG. 1. Nominal 30 μm diam fused silica microtube that has been cut and machined lengthwise by FIB milling. The nominal channel dimensions are 3 μm×40 μm.

The final product will bear no resemblance to the object shown in Fig. 1, nor will it be made by FIB techniques. The preferred method to mass fabricate these devices is deep x-ray lithography, which requires an expensive mask set and synchrotron exposures. The important contribution that FIB prototyping has made here is that the parameters needed to verify a theory and also produce a workable device were identified before any investment of capitol or time was made in the x-ray process. Further refinements of the geometry are being explored using FIB machining in order to minimize any development efforts that involve the x-ray process.

B. Micron-size machine tools and microsurgical devices

The production of micron-size machine tools and manipulator devices for microsurgical applications centers around the ability to rotate the work piece 360° around its principal axis, which is orthogonal to the focused ion beam. This idea is an outgrowth of the scanning probe microscopy (SPM) probe tip fabrication efforts initiated at Bell Labs[2–5] and rotational capabilities were developed essentially simultaneously with similar efforts at Hitachi.[6,7]

1. Machine tools

The key element in making milling tools that can be used to mechanically machine with micron or even submicron resolution is the sharpness of the cutting edge. Tools with submicron radii of curvature cutting edges can be produced by FIB milling in the "lathe" apparatus as a consequence of the smoothness of the surface finish of the starting material, and the sputtering of the final facet surface by grazing incidence ions.

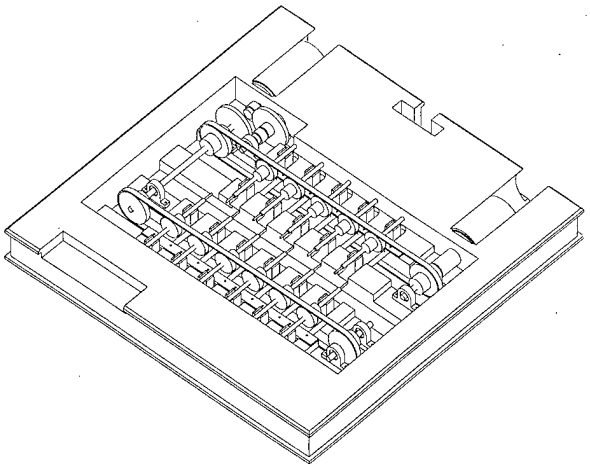
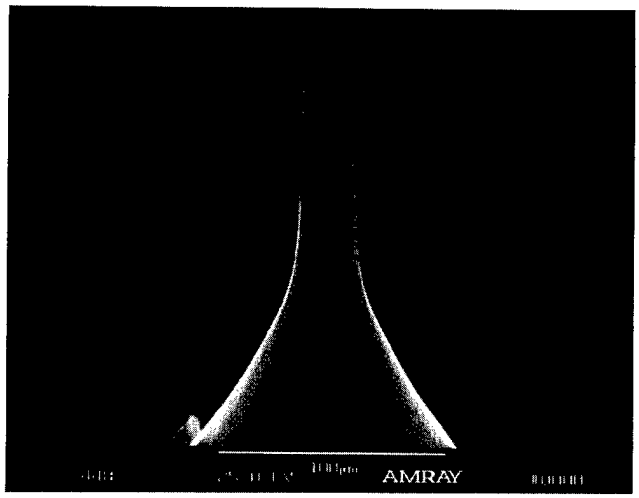


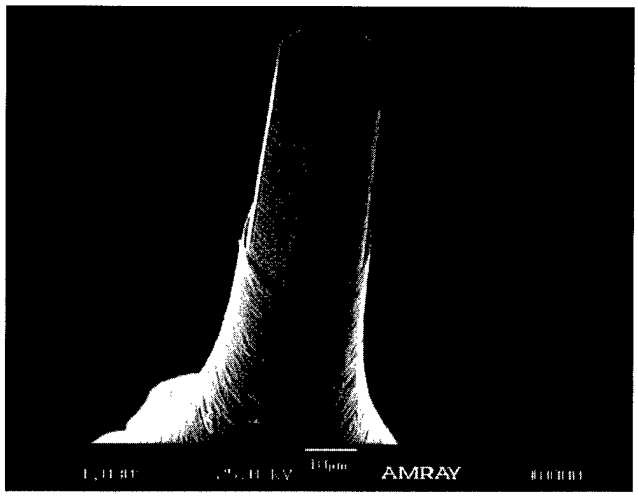
FIG. 2. One of two completed designs of a multiple load cassette for machining microtools. Eleven tool blanks can be mounted on diamond bearing vee-block assemblies, with simultaneous rotation by an in-vacuum stepper motor. The alternative design uses a single vee block with automated feed from a bank of twenty tool blades.

The process for making microtools begins with the tool blank which is a 2.3 mm o.d. centerless ground mandrel containing a 1 mm o.d. cylinder of M42 cobalt hardened high speed tool steel. The end of the 1 mm o.d. segment is reduced to an o.d. of 25 μm by diamond turning, which results in the smooth surface finish. This tool blank is mounted in the rotation device without regard to runout: The only requirement is that the axis of the tool be perpendicular to the ion beam. A facet of desired dimensions is milled on the tool blank according to procedures given in the literature,[8] while the tool is stationary. The tool is then rotated through an angle appropriate for the number of cutting edges or facets desired and realigned with the reference axes, if necessary (since the tool mounting process usually results in "run out" or eccentricity, there will be motion of the tool in the X direction resulting from rotation). The number of cutting edges and the machining sequence are important and the edge effects from the Gaussian intensity of the ion distribution beam will produce at least one rounded edge which will not function as a cutting edge. Tools with one, two, four, and six facets have been produced by this technique, and these tools have been used to machine poly(methylmethacrylate) (PMMA) with a high degree of dimensional tolerance, as shown in Refs. 8–10.

Work in progress is aimed at making tools more rapidly, by automated or semiautomated procedures. The IfM has begun a collaborative research with Sandia National Laboratories to expand the micromachining effort and the demand for tools will be larger than the current method of one-at-a-time can provide. Designs for a multiple-load cassette have been completed, which will allow the semiautomatic loading of tools, thus eliminating a great deal of pump-down time and manual manipulation. One of these designs is shown in Fig. 2, and a combination of this cassette with a pattern recognition program for centering the work piece will significantly reduce the time needed to produce a batch of tools. Approximately 1 h is required to machine the six-facet, five



(a)



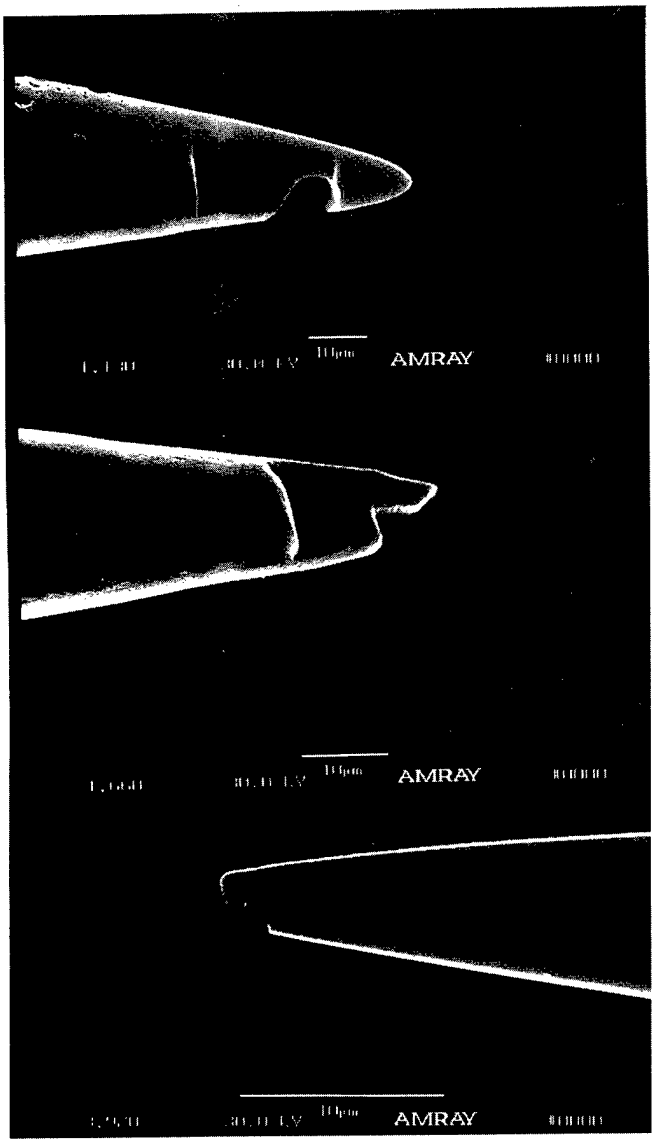
(b)

FIG. 3. SEM perspective images of a six facet cutting tool for ultraprecision machining. The tool measures 23 μm across the points, 20 μm across the facets with five cutting edges, each 60 μm long. (a) 45° view; (b) 30° view.

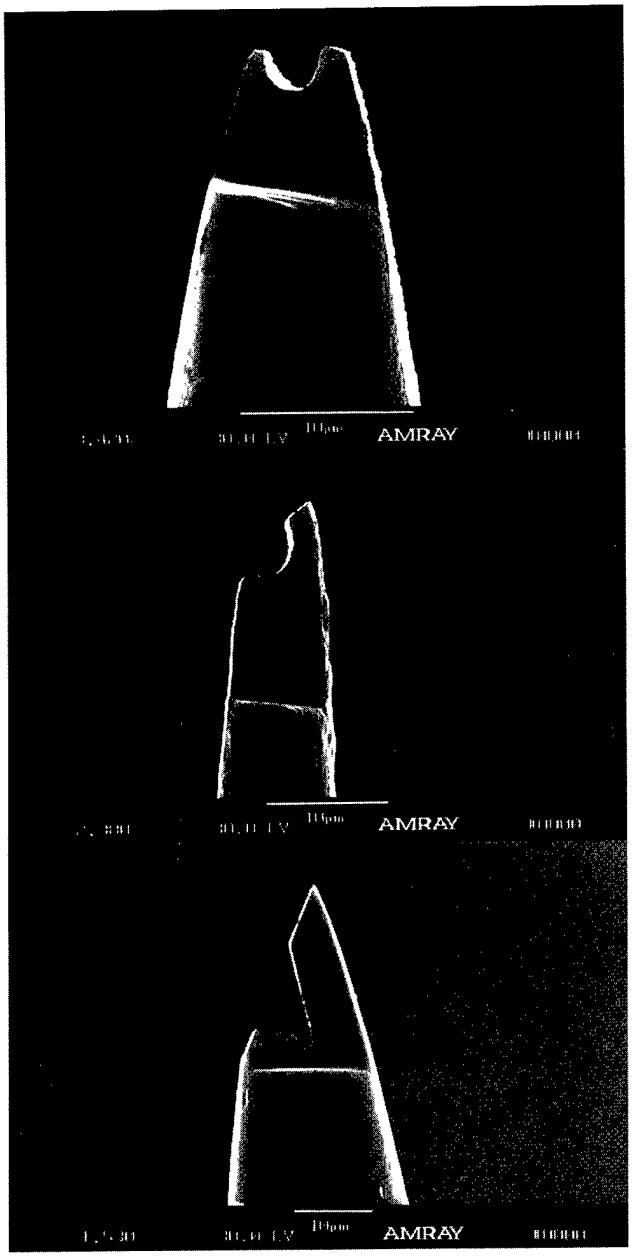
cutting edge tool shown in Fig. 3. The cost of the starting blank is $40, thus minimized additional machining time is necessary to keep added costs from FIB machining as low as possible.

2. Microsurgical tools

Specially shaped microsurgical tools have been produced by FIB methods in response to a need generated by circulatory research performed on mice. Genetically engineered mice are the preferred test animal for many reasons, but the consequence of using such a small mammal is the small size of organs and especially blood vessels. The usual techniques for manipulating tissue, arterioles, and venules under microscopic observation involve the use of pulled glass capillaries. The ends of these capillaries have uniform taper and submi-



(a)



(b)

FIG. 4. Tissue stabilizers and microartery (venule) manipulators for research on the circulatory system. These features are ion milled in the ends of anodically tapered stainless steel needles in a matter of minutes. (a) Shows two structures designed to capture blood vessels and one designed to stabilize motion. (b) Shows three structures designed to stabilize motion of blood vessels in the 5 to 10 μm diam range.

cron radii of curvature but no variation in geometric shape for gripping or other manipulations is possible. The capillary tips are also extremely fragile.

A method to make similar geometric shapes in stainless steel from an inexpensive starting material has been established. Commercially available stainless steel needles costing less than \$0.25 each, with ground tips have been reduced to the desired geometry by an anodic dissolution process. The tips of these anodically tapered needles have radii of curvature ranging between 2 and 4 μm and an extremely smooth surface finish which means a very quick ion milling procedure to yield a specialized device such as a tissue stabilizer, or an artery occluder or other manipulative devices as shown in Fig. 4. These tools are to be used on blood vessels in the 5–15 μm diam range: Fig. 4(a) shows two devices intended to hook or move small blood vessels and one device intended to stabilize the position of a blood vessel. Mobility of these small blood vessels presents significant problems for the surgical procedures, and Fig. 4(b) shows three additional devices intended to localize blood vessels.

The process for making microsurgical tools is more operator intensive than the process for making milling tools,

however, some advantages in cost reduction can be expected by adapting the multiple load cassette to work with 0.25-mm-diam needles. Additional cost control factors are the low cost of the starting material and a batch process for the anodic electroforming.

C. Three dimensional ion milling

The possibility of three dimensional control over ion-milled cavities makes the FIB technique a potential source for mass production of microdevices, or microcomponents. Ion milling time is not an issue here, since molds or embossing masters on the microscale are costly to produce by any technique, and true three dimensional control has not yet been achieved by any technique, either. Two FIB methods have been taken in this work: the direct programming approach and the "exact" solution approach.

1. Direct programming approach

The initial attempts at three dimensional ion milling used a sputter time (pixel dwell time) control method, called the direct programming approach. The ion beam deflection control program was rewritten to allow for variable dwell times as a function of a length or a characteristic dimension within geometric deflection patterns such as rectangles, squares, circles, or annuli. The ion beam dwell time was incremented or decremented on each pixel according to a mathematical function resulting in a regular, systematic increase, or decrease in the dose along the characteristic length in the deflection pattern, thereby milling to greater or lesser depths as a function of position in deflection pattern. The actual depths of ion milled features are determined *a priori* by a maximum dwell time returned from the program and an empirically determined depth calibration versus dose for the material being sputtered. This programming approach assumes that the dose given to any pixel contains no contribution from adjacent pixels, and that the yield is invariant with angle of incidence. Deflection routines were written to produce a variety of geometric cross sections (26 total, with symmetry inversion) such as sinusoids, parabolas, and hemispheres. Ion-milled sinusoidal features based on rotational symmetry around the Z axis and on planes of symmetry (which produce sinusoidal troughs), and other features such as hemispherical domes and parabolic troughs have been ion milled and reported in Refs. 11 and 12.

The direct programming method is successful at producing a first approximation to the shapes required, but the absolute depth and the exact shape are difficult to obtain. The reasons are clear when the relationship between the pixel size and the ion beam FWHM is considered, as well as the changing yield with angle of incidence, as the three dimensional feature develops in time. The changing sputter yield situation is illustrated in Figs. 5(a) and 5(b); Fig. 5(a) shows the additional dose received by pixels that correspond to the apex (i.e., the deepest part) of a parabolic cross section due to the overlap effect from the ion beam intensity distribution. The depth of the parabola will be greater than expected, even when using carefully calibrated empirical sputter rates, resulting in an elongation of the parabola with respect to the originally programmed geometry. Figure 5(b) illustrates the second main source of deviation from ideality when using the direct programming approach to produce three dimensional features. The profiles through the feature result in variable angles of incidence between the ion beam and the feature surface. The variation occurs as the feature develops in time, starting from a planar surface. Figure 5(b) is illustrated for a parabolic trough and the change in the angle of incidence increases the sputter yield disproportionately at the outer edges relative to the focus. The net result is excessive milling in this region, and a concomitant departure from the desired dimensions. There are additional effects that cause deviations from the programmed, expected geometry and these have been discussed in a recent publication.[11]

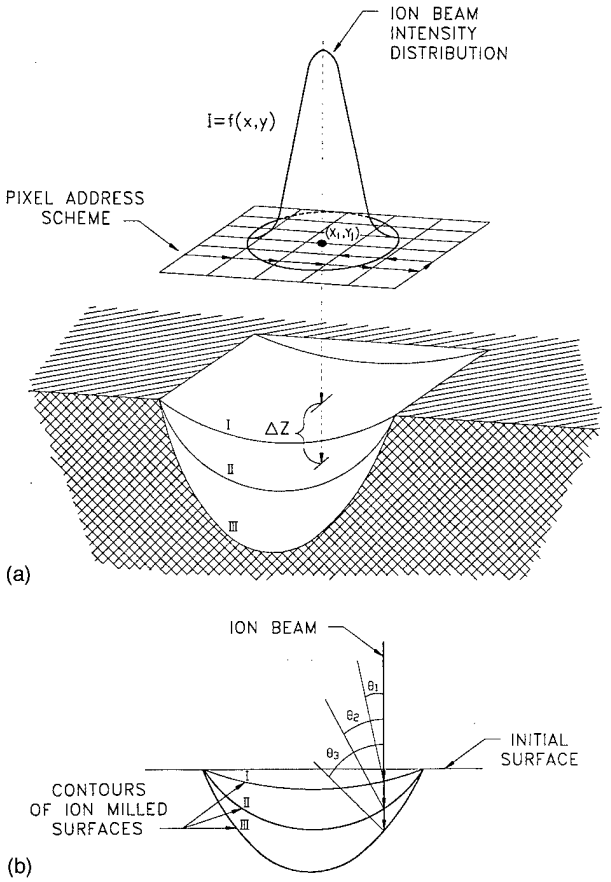


FIG. 5. (a) Schematic representation of the three dimensional ion-milling program. The direct programming approach varies the pixel dwell time in the address pattern with a 1:1 correspondence between the dwell time needed to obtain increments ΔZ. (b) Schematic representation of the change in angle of incidence as a parabolic feature develops. The exact solution corrects for the changing sputter yield as a function of the angle of incidence, and for the additional dose at pixel (i,j) from all other pixels in the address scheme.

2. Exact solution approach

The control over absolute dimensions afforded by the direct programming method is too unpredictable for micromanufacturing applications, although the surface finish and

dimensional resolution (in single crystal silicon) are excellent. The exact solution approach incorporates corrections for the overlap effect and the variable yield with changing angle of incidence illustrated in Fig. 5. These are not incremental corrections on the direct programming approach, but rather a completely different solution to the problem.

The process begins by specifying the final geometric shape of the cavity, e.g., a parabolic trough on a 10×10 μm square with a 12 μm depth. This shape can be represented by curve III in Fig. 5. The equation of the profile of the desired cavity is then obtained, and the milling process is divided into N sequences, or slices, two of which can be represented by curves I and II in Fig. 5. Given the material parameters and the ion beam parameters, the dwell time at each pixel is solved to produce ΔZ, the depth increment needed to transform curve I into curve II. The depth increment at the (i,j) pixel can be expressed as

$$\Delta Z_{ij}=\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\frac{\Phi(x_k,y_l)}{\eta}f_{x_k y_l}(x_i y_j)S(\theta_{x_i y_j})t_{kl}\Delta x_k \Delta y_l, \tag{1}$$

where $\Phi(x,y)$ is the ion flux at (x,y), η is the atom density of the solid, $f(x,y)$ is the ion beam intensity distribution function, $S(\theta)$ is the angle-dependent sputter yield, and $t_{k,l}$ are the dwell times. Δx and Δy are unit increments along the pixel address scheme and the summation over the indices k,l accounts for dose received at (x_i, y_j) from all pixels in the address scheme.

The dwell time needed for the depth increment thus takes into account the contribution of ion flux from all pixels on the deflection plane, through the intensity distribution. The sputter yield is also adjusted on a point-to-point basis for the change in angle of incidence. The solution for each depth increment can be obtained from a matrix relationship[12]

$$\{C_{k,l}(i,j)\}\{t_{k,l}\}=\{\Delta Z_{i,j}\}, \tag{2}$$

where $C_{k,l}(i,j)$ is the local sputter relationship, $t_{k,l}$ is the dwell time, and $\Delta Z_{i,j}$ is the depth increment.

The sputtering process was simulated and solved by Eq. (2) off-line for a number of geometric shapes.[12] Absolute yields were computed using the semiempirical relationships of Matsunami *et al.* and Yamamura *et al.*[13,14] The model solutions provided by the matrix inversion routine required to solve Eq. (2) were verified as unique solutions.

The present status of the exact solution approach is embodied in an ion beam control program called MILL, resident on the FIB control computer, which has:

(1) Yamamura and Matsunami's semiempirical yield functions programmed for a choice of Si, Fe, Cu, Ag, and Au targets, with Ga+ as the projectile, with variable beam energy.
(2) The matrix inversion routine which can be solved for a 30×30 pixel field, for hemispherical, parabolic or sinusoidal cross sections; the final geometry is divisible into N slices or sputter intervals.

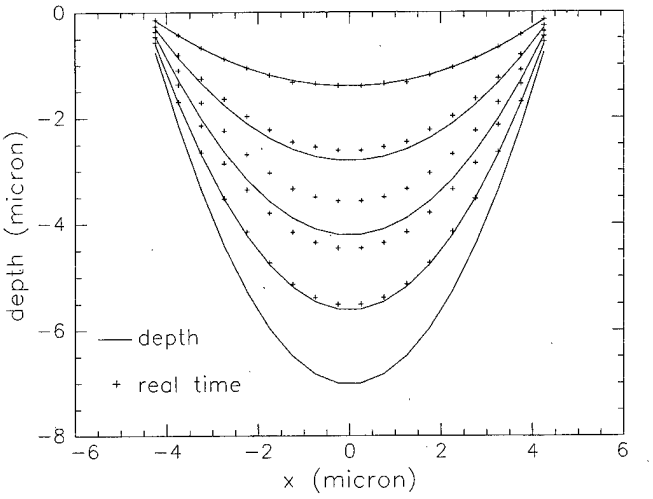


FIG. 6. Plots of the dwell time solutions (+) for each of the surface contours (solid lines) solved by Eq. (1) for ion milling a parabolic trench in silicon with 20 keV Ga^+.

(3) Transfer of each of the N sputter intervals as one scan to the ion beam deflection control.

Solutions obtained on the FIB computer match those obtained off-line, exactly. Figure 6 shows a plot of surface contours obtained for a parabolic trench to be ion milled in silicon. The solid curves are the surface contours and the discrete points are the dwell time solutions obtained from Eq. (1), normalized to the maximum depth. Note that on the first contour, there is a 1:1 correspondence between the individual dwell times and the surface contour. The correspondence breaks down for the deeper contours, as a result of the additional dose received by the pixels due to the beam intensity distribution. (The dwell times that appear to match the fourth contour are actually the dwell times required to produce the fifth contour.) The intensity overlap contribution dominate the yield variation with angle of incidence for the solutions obtained for Fig. 6, hence the dwell times adhere to a parabolic relationship at any point in the surface development.

IV. CONCLUSION/FURTHER WORK

The most immediate problem is to pass the dwell time solutions to the deflection control hardware and test the accuracy of the three dimensional ion milling routine. The next priority in microstructure fabrication is to construct a multiple loading cassette and write pattern recognition software so that the tool making process can be semiautomatic.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of their colleagues at Bell Labs, Lucent Technologies: Lloyd Harriott, Chris Biddick, and Leslie Hopkins. This work was supported in part by the National Science Foundation under Grant No. 9503204 and by the Institute for Micromanufacturing.

[1]L. R. Harriott, Proc. SPIE **773**, 190 (1987).
[2]M. J. Vasile, C. J. Biddick, and S. A. Schwalm, J. Vac. Sci. Technol. B **12**, 2388 (1994).
[3]M. J. Vasile, D. A. Grigg, J. E. Griffith, E. A. Fitzgerald, and P. E. Russell, Rev. Sci. Instrum. **62**, 2167 (1991).
[4]M. J. Vasile, C. J. Biddick, and H. Huggins, Appl. Phys. Lett. **64**, 575 (1994).
[5]M. J. Vasile, C. J. Biddick, and S. A. Schwalm, Proc. ASME DSC-Vol. 46, Micromechanical Systems, edited by A. P. Pisano and J. Jara-Almonte, Book No. H00832 (1993) (unpublished), p. 81.
[6]T. Ishitani, T. Ohnishi, and Y. Kawanami, Jpn. J. Appl. Phys., Part 1 **29**, 2283 (1990).
[7]T. Ishitani, J. Vac. Sci. Technol. B **9**, 2633 (1991).
[8]C. R. Friedrich and M. J. Vasile, J. MEMS **5**, 33 (1996).
[9]M. J. Vasile, C. R. Friedrich, B. Kikkeri, and R. McElhannon, Precis. Eng. **19**, 180 (1996).
[10]C. R. Friedrich, P. J. Coane, and M. J. Vasile, Microelectron. Eng. **35**, 367 (1997).
[11]M. J. Vasile, Z. Niu, R. Nassar, and W. Zhang, J. Vac. Sci. Technol. B **15**, 2350 (1997).
[12]R. Nassar, M. Vasile, and W. Zhang, J. Vac. Sci. Technol. B **16**, 109 (1998).
[13]N. Matsunami, Y. Yamamura, Y. Itakawa, N. Itoh, Y. Kazumata, S. Miyagawa, K. Morita, and R. Shimizu, *Energy Dependence of Sputtering Yields of Monatomic Solids*, IPPJ-AM-14 (Institute of Plasma Physics, Nagoya University, Nagoya, Japan, 1983).
[14]Y. Yamamura, Y. Itakawa, and N. Itoh, *Angular Dependence of Sputtering Yields of Monatomic Solids*, IPPJ-AM-26 (Institute of Plasma Physics, Nagoya University, Nagoya, Japan, 1983).

Focused-ion-beam-assisted etching of diamond in XeF_2

Jun Taniguchi and Naoto Ohno
Department of Applied Electronics, Science University of Tokyo, Noda, China 278, Japan

Shuuichi Takeda
Komatsu Limited, Hiratsuka, Kanagawa 254, Japan

Iwao Miyamoto
Department of Applied Electronics, Science University of Tokyo, Noda, China 278, Japan

Masanori Komuro[a)]
Electrotechnical Laboratory, Tsukuba, Ibaraki 305, Japan

(Received 25 November 1997; accepted 1 May 1998)

Ga focused-ion-beam (FIB)-assisted etching of single-crystal diamond and thin film diamond in XeF_2 was studied. The etch yield in FIB-assisted etching of diamond in XeF_2 is enhanced some six times over the physical sputtering yield. In the crystal orientation dependence of the etch yield in FIB-assisted etching, the (100) face produced the highest etch yield of the three faces—(100), (110), and (111). Thin film diamond produces the lowest etch yield. A diamond field emitter with a tip radius of less than 100 nm was obtained using Ga FIB spot exposure. © *1998 American Vacuum Society.* [S0734-211X(98)06204-0]

I. INTRODUCTION

Diamond is a promising material for application such as wide-gap semiconductor devices, micromachining tools, negative-electron-affinity field emitters, and thin x-ray mask membranes.[1,2] Because of diamond's hardness, however, it is very difficult to fabricate bulk diamond and/or thin films with miniaturized dimensions. In previous work, we studied electron-beam (EB)-assisted chemical etching of diamond with oxygen or hydrogen gas.[3,4] This has the disadvantage of a low etching rate, however, focused-ion-beam (FIB) etching has been used to fabricate miniaturized electronics devices, to repair photomasks, and to modify integrated circuits. FIB physical sputtering has been conventionally applied to the local removal of target materials, but has disadvantages such as redeposition and radiation damage. These problems can be solved by ion-beam-assisted etching where reactive gas molecules are introduced simultaneously with ion bombardment, which produces an ion-enhanced surface reaction.[5]

Fluorine is known to react with graphite to produce carbon fluoride. To fabricate miniaturized dimensions and increase the etching rate, we conducted FIB-assisted etching of diamond in XeF_2, and studied the morphology of the etched surface. In this work, we report beam-assisted etching characteristics of diamond using Ga FIB in XeF_2.

II. EXPERIMENTS

Experiments were conducted using a FIB column with a differential pumping system between a Ga ion source chamber and a reaction chamber at a vacuum pressure of 10^{-7} Torr. XeF_2 gas was introduced from a 0.12 mm inner diameter nozzle about 5 cm long whose exit was about 500 μm from the substrate surface. Synthetic single-crystal diamond substrates with (110)-, (100)-, and (111)-oriented faces and thin film diamond were used as samples. The thin film diamond was produced by chemical vapor deposition with mixtures of H_2 and CH_4 on Si substrates. After deposition, the thin film diamond was polished by mechanical lapping to a 26 μm thickness and its surface roughness was 66 nm R_{max} (Fig. 1). Thin film diamond is polycrystalline, with a grain size distributed from 2 to 10 μm.

III. RESULTS AND DISCUSSION

A. Etched groove morphology

Grooves were etched into the thin film diamond samples using a focused Ga ion beam at a 30 kV acceleration voltage, 475 pA beam current, 10 min irradiation time, 10 μm scan length, and 0.7 μm beam diameter (Fig. 2). The line dose was 285 μC/cm. Despite the polycrystalline thin film diamond structure (Fig. 1), etched grooves in thin film diamond show no edge roughness due to the grain structure, and these shapes are similar to etched grooves in single-crystal diamond (Figs. 3 and 4). Ga drops were observed on both samples at 0 Torr nozzle pressure, but at a 4 Torr nozzle pressure, Ga drops were not observed and the groove bottoms were very smooth. XeF_2 gases react with Ga and causes Ga fluoride vapor, thereby removing the Ga drops.

Rectangular patterns were etched into thin film diamond in FIB-assisted etching with XeF_2 at a 30 kV acceleration voltage, 475 pA ion beam current, 0.7 μm beam diameter, 10 min exposure time in a 6×6 μm^2 scan area (Fig. 5). The dose was 0.79 C/cm^2. The square on the right is a 0 Torr nozzle pressure etched groove. The square on the left is a 4 Torr nozzle pressure etched groove. The bottoms of the etched areas were smooth for both types of etching groove.

In FIB-assisted etching, the etched area is conductive because ion beam bombardment causes an amorphous conductive layer and Ga ion implantation. The chargeup was very weak and the positioning accuracy high, enabling sharp pat-

[a)]Electronic mail: Komuro@etl.go.jp

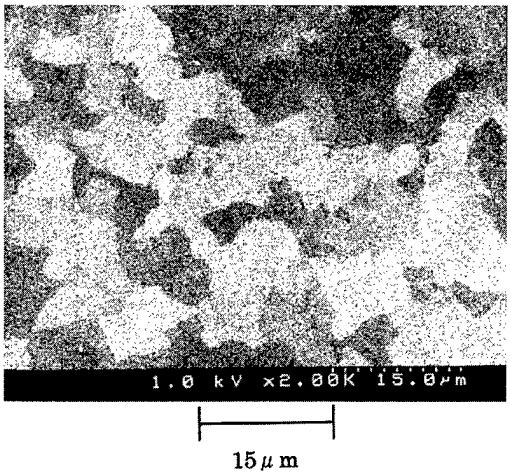
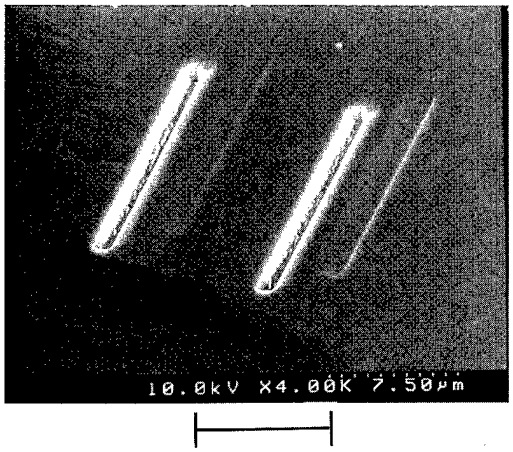


FIG. 1. Thin film diamond sample.



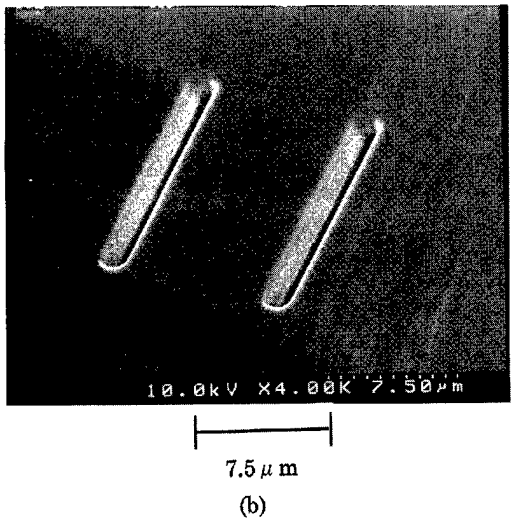
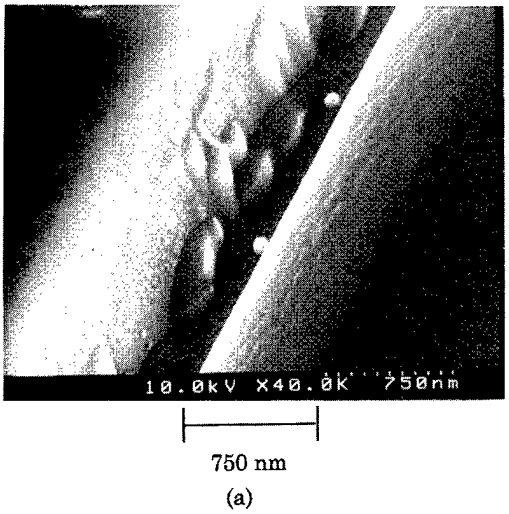


FIG. 2. Micrographs of Ga FIB-assisted etching into thin film diamond samples with XeF_2 at (a) 0 and (b) 4 Torr nozzle pressure.



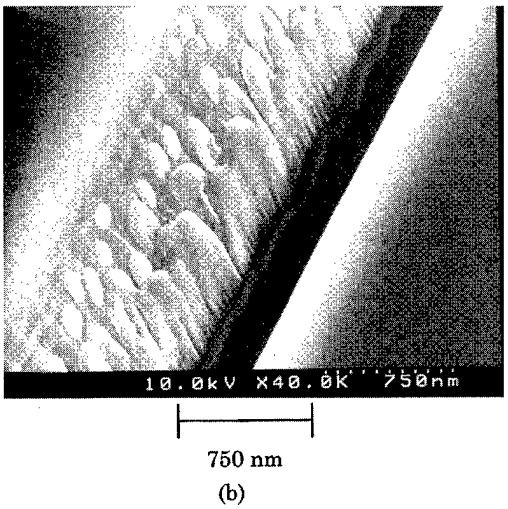


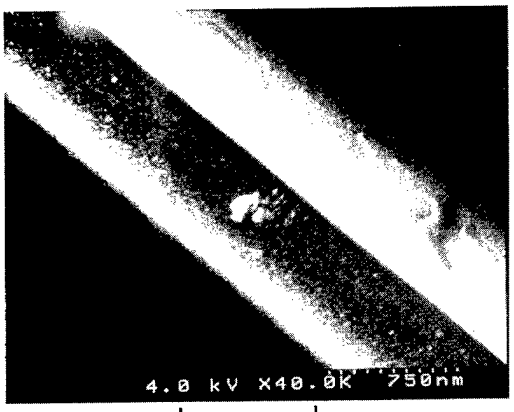
FIG. 3. Magnified micrographs (a) and (b) respectively, of the micrographs in Fig. 2 at 0 and 4 Torr nozzle pressure.

terns using FIB-assisted etching. In EB-assisted chemical etching, chargeup is strong and EB positioning accuracy poor.[3,4] The bottom and sidewalls were very rough.

An amorphous layer was produced with ion bombardment in the thin film diamond (Fig. 6). Ion-beam bombardment caused the amorphous surface layer.

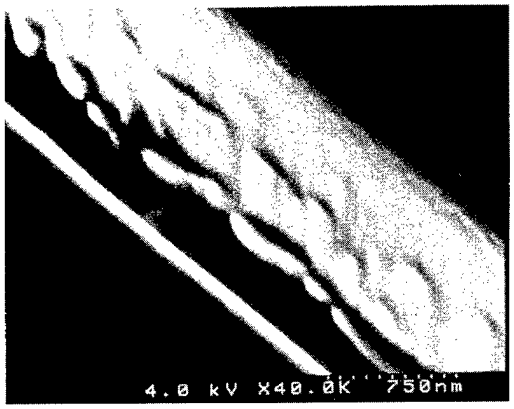
B. Nozzle pressure dependence of the etch yield

The etch yield for crystal and thin film diamond samples depends on nozzle pressure (Fig. 7). The etching conditions were 30 kV acceleration voltage, 500 pA beam current, and 0.7 μm beam diameter. The areas scanned were $10 \times 10\ \mu m^2$ for 10 min and $20 \times 20\ \mu m^2$ for 20 min. The 4 Torr nozzle pressure equaled nearly 3.2×10^{16} molecules/scm^2. The etched depths were measured by atomic force microscopy (AFM). The etch yield (number of C atoms removed per ion) was enhanced about six times over the physical sputtering yield. With increasing nozzle pressure, the etch yield became saturated. The larger scan area exhibited a higher etch yield than the smaller scan area.



750 nm

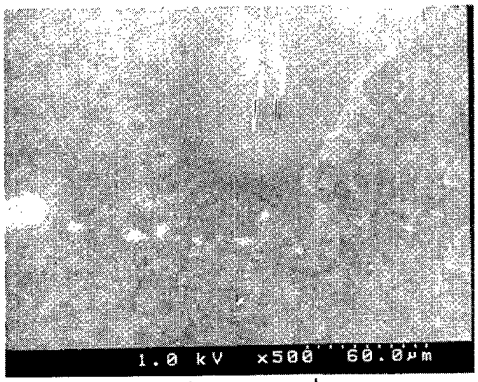
(a)



750 nm

(b)

FIG. 4. Micrographs of Ga FIB-assisted etching into crystal diamond samples with XeF_2 at (a) 0 and (b) 4 Torr nozzle pressure.



60 μ m

(a)

12 μ m

(b)

FIG. 6. (a) Scanning electron microscopy (SEM) micrographs of the amorphous layer produced with ion bombardment into thin film diamond; (b) an enlarged view at the etched grooves.

This is because the gas flux exceeds the ion-beam flux in the $20\times20\ \mu m^2$ etched area. The (100) crystal surface exhibits a slightly higher etch yield than that for the thin film diamond.

The etch yield depends on the crystal-oriented faces of the thin film diamond. The etch yield is small, so the thin film

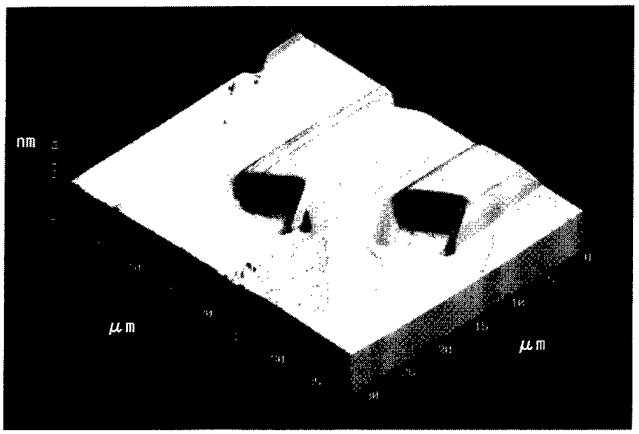


FIG. 5. AFM image of rectangular patterns etched into thin film diamond in FIB-assisted etching.

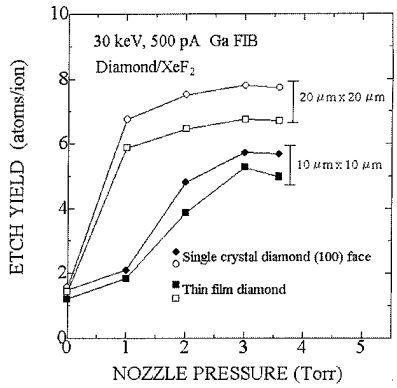


FIG. 7. Nozzle pressure dependence of the etch yield for the crystal and thin film diamond samples.

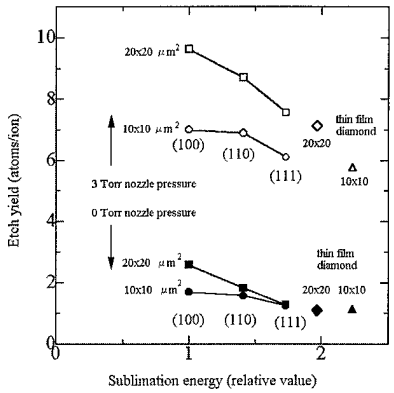


FIG. 8. Crystal orientation dependence of the etch yield for the crystal and thin film diamond samples.

diamond etch yield is less than the sample with one crystal-oriented face. We therefore examined the etch yield crystal orientation dependence.

C. Etch yield crystal orientation dependence

We studied the crystal orientation dependence of the etch yield for crystal and thin film diamond samples (Fig. 8) at etching conditions of 10 kV acceleration voltage, 115 pA beam current, and 2 μm beam diameter. The areas scanned were $10\times10\ \mu m^2$ for 20 min and $20\times20\ \mu m^2$ for 40 min. The etched depths were measured by AFM. The larger scan area exhibited a higher etch yield than the smaller scan area for the same reason as that in the case of etch yield nozzle pressure dependence. In single-crystal diamond, the (100) face had the highest etch yield of the three faces—(100), (110), and (111). The etch yield was inversely proportional to the sublimation energy. For sublimation energy, the (111) face had the highest of the three faces—(100), (110), and (111).[6] The etch yield decreased with increasing sublimation energy. This result is reasonable, considering the sublimation energy. The thin film diamond exhibited the lowest etch yield under the same etching conditions. If the thin film diamond were mixed [(110), (111), and (100)], the thin diamond film would take the medium etch yield. This reason is not clear at present; however we guess that impurity atoms at the diamond grain boundaries might be effective in lowering the etch yield.

D. Diamond field emitter tip fabrication

Diamond is useful as a field emitter tip because of its negative-electron affinity. We therefore fabricated a field emitter tip using a single-crystal diamond edge. Two spots of a Ga FIB were exposed into XeF_2 at the diamond edge. The etching conditions were 10 kV acceleration voltage, 100 pA ion-beam current, 2 μm beam diameter, 30 min exposure time for one-spot Ga FIB exposure, single-crystal diamond with a (111)-oriented face, and 4 Torr nozzle pressure (Fig. 9). Despite the large beam diameter (2 μm), a very small radius field emitter tip was obtained. This emitter tip which was ion-beam bombarded has high conductance, but the other area remains an insulator. When we investigate field emission from a diamond emitter tip, the whole diamond surface needs high conductance. Hydrogen-terminated diamond is known to be conductive,[7] so we plan to fabricate

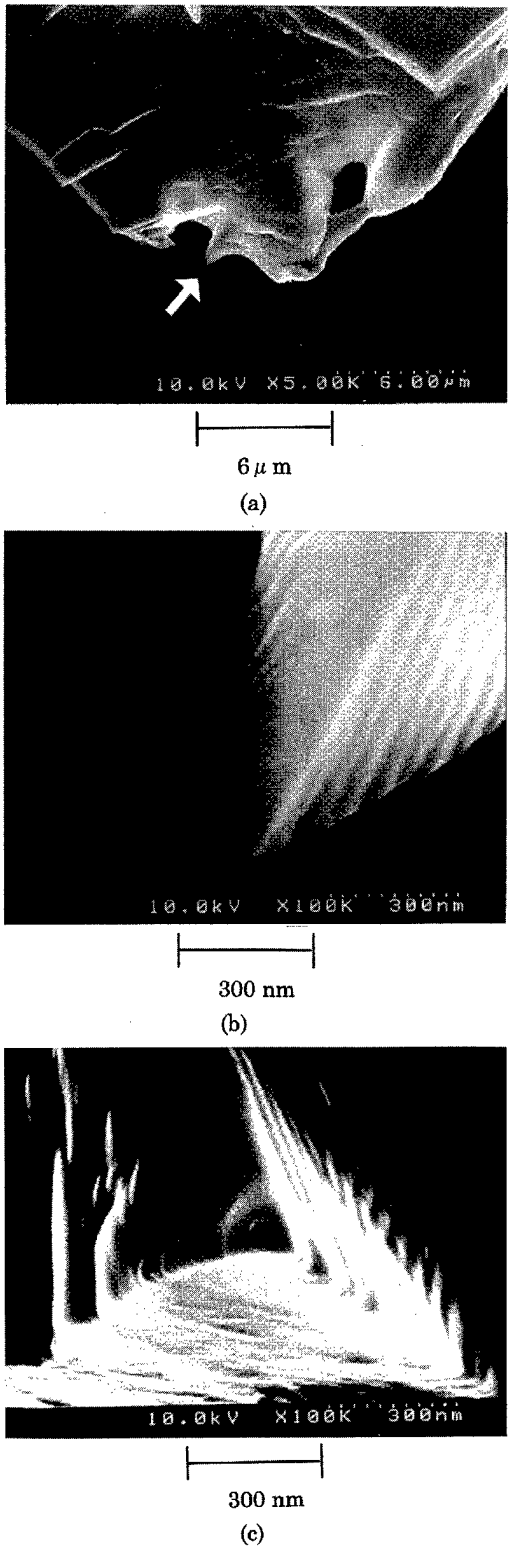


FIG. 9. SEM micrographs of the field emitter tip fabricated; (b) an enlarged view at the emitter in (a); (c) the plan view from the apex in (b).

hydrogen-terminated diamond using hydrogen annealing. We will then measure the emission properties of this emitter tip.

IV. CONCLUSION

We studied beam-assisted etching characteristics of diamond using a Ga FIB in XeF_2. FIB-assisted etching produced smooth grooves. The etch yield was enhanced about six times over the physical sputtering yield. The etch yield became saturated with increasing nozzle pressure. In the crystal orientation dependence on etch yield in single-crystal diamond, the (100) face had the highest etch yield of the three faces—(100), (110), and (111). The thin film diamond has the lowest etch yield. A diamond field emitter with a less than 100 nm tip radius was obtained using Ga FIB spot exposure.

ACKNOWLEDGMENT

The authors thank Dr. S. Matsui of NEC Corporation for providing the samples of thin film diamond.

[1]M. W. Geis, N. N. Efremow, J. D. Woodhouse, M. D. McAleese, M. Marchywka, D. G. Socker, and J. F. Hochedez, IEEE Electron Device Lett. **12**, 456 (1991).
[2]K. Okano, K. Hoshima, M. Iida, S. Koizumi, and T. Inuzuka, Appl. Phys. Lett. **64**, 2742 (1994).
[3]J. Taniguchi, I. Miyamoto, N. Ohno, and S. Honda, Jpn. J. Appl. Phys., Part 1 **35**, 6574 (1996).
[4]J. Taniguchi, I. Miyamoto, N. Ohno, K. Kantani, M. Komuro, and H. Hiroshima, Jpn. J. Appl. Phys. (in press).
[5]M. Komuro, N. Watanabe, and H. Hiroshima, Jpn. J. Appl. Phys., Part 1 **29**, 2288 (1990).
[6]T. Kuroda, *Kesshyo hyomen no kisobussei* (*Surface and Crystal Characteristics*) (Nikkan Kogyo Shinbunsha, Tokyo, 1993), pp. 43–55 (in Japanese).
[7]M. Itoh and H. Kawarada, Jpn. J. Appl. Phys., Part 1 **34**, 4677 (1995).

Focused ion beam etching of resist/Ni multilayer films and applications to metal island structure formation

Masayoshi Nakayama,[a] Fujio Wakaya, Junichi Yanagisawa,[b] and Kenji Gamo[c]
Department of Physical Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

A new method to fabricate single-electron tunneling structures with magnetic materials using focused ion beam (FIB) sputtering and lithography techniques is proposed. By using this method, a small metal island with a size comparable to or smaller than the FIB diameter, connected with source and drain electrodes via tunnel junctions, can be fabricated. In the present article, some important parameters in this method, such as sputtering yields of the photoresist (AZ 1350) and nickel (Ni) layers, were measured and cross-sectional images of the grooves delineated on a AZ/Ni/SiO_2 substrate, which is the basic structure in this method, were observed. At present, a 0.4-μm-wide groove is obtained. However, this width is limited not by a fundamental process but by noise in the scan signal and stage vibration. © *1998 American Vacuum Society.* [S0734-211X(98)06304-5]

I. INTRODUCTION

Single-electron tunneling (SET) effect originates in a Coulomb blockade of electron tunneling[1] and has been mainly observed in the double-junction system which has a small island connected to the source and drain via tunnel junctions. Such structures have been fabricated by a multiple-angle deposition method,[2,3] local oxidation of materials using a scanning tunneling microscope (STM)[4] or an atomic force microscope (AFM).[5] For previous investigation, SETs were fabricated using semiconductors or nonmagnetic metal films and less attention has been paid to the effect of electron spin.

Recently, the control of SET by the spin of the tunneling electron or spin blockade has been proposed[6] and some attempts have been made.[3,7] One possible structure to observe this spin blockade is to fabricate the SET structure using magnetic materials.

Nickel (Ni) is one of the most familiar magnetic metals and the electrical property of its oxide is semiconductor-like (i.e., nonconductive at low temperature). However, microfabrication of such materials for an SET structure has not been established up to now.

A new method to fabricate the SET structure using Ni is proposed and described in detail in the following section. The method consists of focused ion beam (FIB) sputtering and lithography techniques. In our previous work,[8] a lift-off process for a very narrow resist pattern fabricated by FIB lithography was investigated using the self-developed grooves on a nitrocellulose resist, from which 0.16-μm-wide aluminum lines were successfully fabricated. This result indicated that the very narrow pattern with the same size as the FIB spot size can be fabricated by the lift-off process.

In the present article, some important parameters of this method, such as sputtering yields of the photoresist (AZ 1350) and Ni layers, were investigated. The key process in this method is to fabricate a narrow groove by FIB sputtering on a substrate. Therefore, cross-sectional images of the grooves delineated on a AZ/Ni/SiO_2 substrate by 100 keV Ga^+ FIB were observed using a scanning electron microscope (SEM).

II. FABRICATION FLOW AND EXPERIMENTAL PROCEDURES

A. Fabrication flow

The proposed new method to fabricate a small metal island structure is as follows. First, a patterned narrow metal thin layer is formed by metal evaporation and lithography steps on an insulator and embedded in a spin-coated resist layer. (See Fig. 1.) The metal strip, as well as resist layer, is then sputtered by FIB scanning in a line across the strip to form a narrow groove with the size near the beam diameter or smaller. After the sputtering, the sample is exposed to oxidation atmosphere to form a thin oxide. Only the exposed cross-sectional area of the film at the groove may be oxidized and oxidation at other surfaces is minimized due to the coverage of the resist layer.

After the second metal layer is evaporated over the surface, the resist layer is removed for a lift-off step leaving the groove filled with the second metal. By this lift-off step, an island structure connecting with the separated metal strips (acting as source and drain of the SET devices) via its oxide is fabricated.

If the width of the island structure (i.e., the length of the groove and/or the width of the initial strip) is not narrow (short and/or narrow) enough, FIB sputtering may be further used again to narrow the channel whose width should be limited by the beam resolution. Using a fine focused (<10 nm) beam[9] an approximate 10 nm sized SET can be expected, which is desirable for high temperature operation of SET devices. The process described above is schematically shown in Fig. 1.

[a]Electronic mail: masayoshi@nano.ee.es.osaka-u.ac.jp
[b]Corresponding author; electronic mail: yanagisawa@ee.es.osaka-u.ac.jp
[c]Also at Research Center for Materials Science at Extreme Conditions, Osaka University, Toyonaka, Osaka 560-8531, Japan.

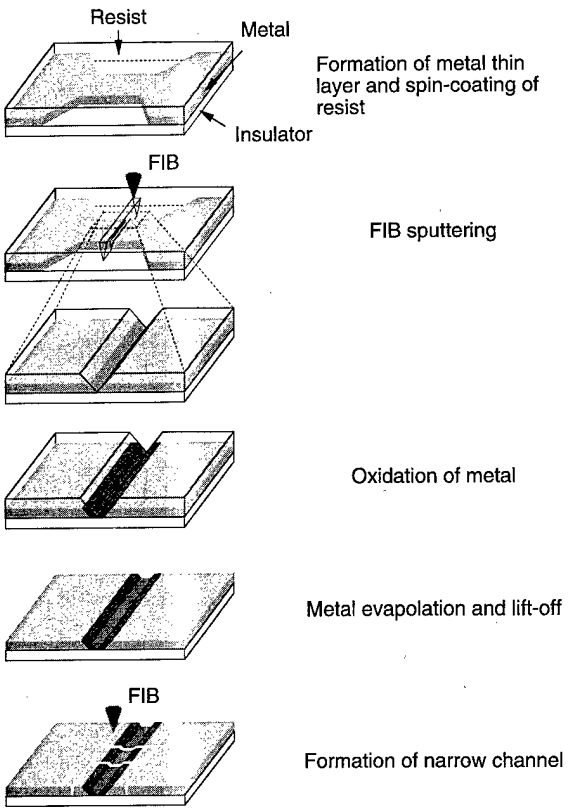


FIG. 1. Schematic of the fabrication process of the small island structure using FIB etching and liftoff. The final step reduces the size of island further.

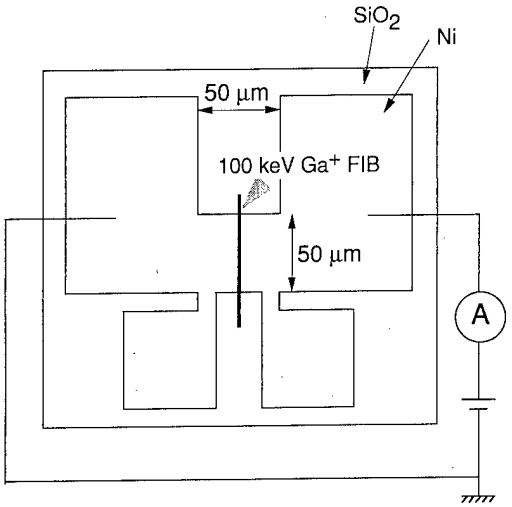


FIG. 2. Resistance measurement of Ni strip during sputter etching by FIB.

B. Experimental procedures

In the present study, as described in Sec. I, Ni and AZ 1350 photoresist were used for the metal and the resist, respectively.

The most important process in this method is the fabrication of the narrow groove across the initial Ni strip. To decrease the groove width, the relation between the sputtered depth and the FIB dose should be well known. Therefore, the sputtering yields of the AZ resist and Ni layers were first measured separately.

AZ resist was spin coated on a Si wafer with a thickness of 140 nm. A postbake was performed at 70 °C for 2 min. Ni was evaporated on SiO_2 substrate with a thickness of 230 nm. 100 keV Ga^+ FIB was used to irradiate over 50 $\times 10\ \mu m^2$ areas, changing the dose in both samples. The FIB spot size was about 0.1 μm in diameter with the ion beam current of about 40 pA. Depth of the sputtered region was measured by a stylus (DEKTAK).

Narrow grooves were formed as follows. First, a patterned Ni layer with a thickness of about 30 nm was evaporated on a silicon-oxide layer using a lithographic technique. It has a small strip (50 μm wide and 50 μm long) connected with large contact pads to measure the resistance of the strip, as shown in Fig. 2. AZ resist was spin coated over the surface with a thickness of 140 nm and postbaked under the same condition described above. A small area of AZ resist on the contact pads was removed using a conventional photolithography to make a contact for measuring the resistance of the Ni strip.

Then, a 100 keV Ga^+ FIB was repeatedly scanned in a line across the Ni strip. FIB conditions were the same as described above. The ion irradiated region was sputtered physically to cut the Ni strip, as well as the resist layer, into two regions. If the cutting process is controlled precisely, it should be possible to form a groove with a width less than a beam diameter.[8] To control the cutting process precisely, the resistance of the strip was monitored during the FIB sputtering, as shown in Fig. 2. The cross section of the grooves was observed using a SEM.

III. RESULTS AND DISCUSSION

The sputtering yields of the AZ resist and Ni layers for 100 keV Ga^+ FIB were shown in Figs. 3(a) and 3(b), respectively, as a function of the ion dose. From the results, it was found that the resist sensitivity of AZ resist for 100 keV Ga FIB was less than that of PMMA and nitrocellulose resists[10] and the doses of about 5×10^{17} cm^{-2} were needed to sputter the 140 nm AZ resist layer. The sputtering rate of AZ resist showed a nonlinear dependence with dose. In contrast, a linear relation between the sputtered depth and the ion dose was observed for Ni and the sputter yield was estimated to be 2.5 nm/10^{16} ions cm^2. A dose of about 1×10^{17} cm^{-2} should be needed to sputter the whole 30-nm-thick Ni layer. This corresponds to a sputtering yield of 2.5 atoms/ion, a value of which is in reasonable agreement with reported value.

AZ film shows a rapid sputtering rate at low dose and a low etching rate at doses higher than 1×10^{16} cm^{-2}. A similar behavior was also observed for sputtering PMMA resist. During the initial stage of sputtering, the organic resist films decompose by ion irradiation and volatile fragment molecules are removed by evaporation. At high doses, the resi-

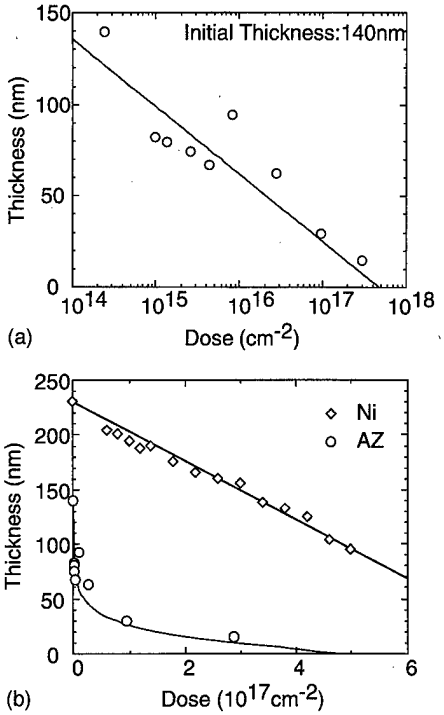


FIG. 3. Sputtering yields of (a) AZ resist and (b) Ni film for 100 keV Ga^+ FIB.

due is carbonized and show low sputtering yield of ~4 atoms/ion, a value of which is roughly similar to the etching rate of C.

Dose dependence of the linewidth of grooves delineated on the AZ/Ni/SiO_2 were shown in Fig. 4. By assuming a Gaussian distribution of the beam and neglecting the angular dependence of the sputtering yield, the linewidth $2w$ should be given as

$$w = \sqrt{2\sigma^2 \ln\left(\frac{N_D}{\sqrt{2\pi}\sigma N_c}\right)},$$

where N_D is the dose, σ is the standard deviation, and N_C is the critical dose to sputter the total thickness. The result is shown as a solid curve in Fig. 4.

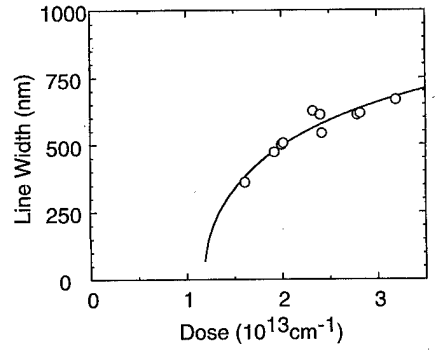


FIG. 4. Dose dependences of the linewidth of grooves delineated on the AZ/Ni/SiO_2 by 100 keV Ga^+ FIB.

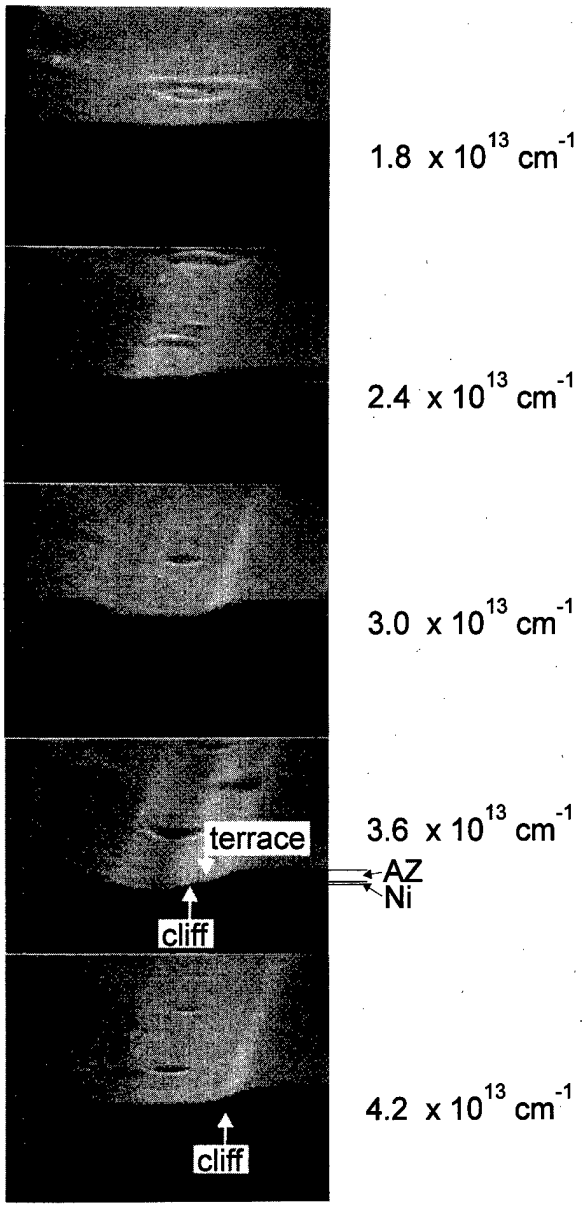


FIG. 5. Cross-sectional SEM images of the lines delineated on the AZ/Ni/SiO_2 by 100 keV Ga^+ FIB for several ion doses.

Using the obtained σ, the effective spot size (FWHM) was calculated as about 0.47 μm. However, the observed linewidth was about four times larger than the beam diameter. This was caused by the beam fluctuations due to noise in scan signal. In fact, we observed a periodic waving of line patterns delineated in PMMA with a peak-to-peak height of about 0.4 μm and a frequency corresponding to 60 Hz. It can be expected that we can obtain narrower grooves if the noise is eliminated. Recently, we have obtained 0.1-μm-wide grooves, which are comparable to the FIB spot diameter, by eliminating the scan noise successfully.

The cross-sectional SEM images of grooves formed at several ion doses are shown in Fig. 5. The groove was grown by increasing the ion dose. At a dose of 3.6×10^{13} cm^{-1}, a

terrace of the Ni layer was observed at both sides of the groove. At higher dose, the terrace was sputtered and the cliff was formed at both sides of the groove. The width of the groove which determines the size of the Ni island in the successive process was rather wide in the present result, but it was mainly because of the scanning noise (and can partly be the vibration of the sample stage), as described above, and these problems are not fundamental but technical ones.

IV. CONCLUSION

A new fabrication process for SET structures using magnetic materials by FIB sputtering and lithography techniques was proposed and some important parameters of this process were measured.

A 0.4-μm-wide groove was formed which is required for metal island formation by lift-off process. However, this width is limited not by a fundamental process but by a technical problem.

ACKNOWLEDGMENT

This work was supported, in part, by a Grant-in-Aid for Scientific Research in Priority Area ''Nanoscale Magnetism and Transport'' from the Ministry of Education, Science, Sports, and Culture, Japan.

[1]K. K. Likharev, IBM J. Res. Dev. **32**, 144 (1988).
[2]Y. Nakamura, C. D. Chen, and J.-S. Tsai, Jpn. J. Appl. Phys., Part 2 **35**, L1465 (1996).
[3]K. Ono, H. Shimada, S. Kobayashi, and Y. Ootuka, J. Phys. Soc. Jpn. **65**, 3449 (1996).
[4]K. Matsumoto, S. Takahashi, M. Ishii, M. Hoshi, A. Kurokawa, S. Ichimura, and A. Ando, Jpn. J. Appl. Phys., Part 1 **34**, 1387 (1995).
[5]E. S. Snow, D. Park, and P. M. Campbell, Appl. Phys. Lett. **69**, 269 (1996).
[6]S. Maekawa and J. Inoue, J. Magn. Magn. Mater. **156**, 315 (1996).
[7]K. Ono, H. Shimada, and Y. Ootuka, J. Phys. Soc. Jpn. **66**, 1261 (1997).
[8]H. Kaneko, Y. Yasuoka, K. Gamo, and S. Namba, J. Vac. Sci. Technol. B **6**, 982 (1988).
[9]R. L. Kubena, J. W. Ward, F. P. Stratton, R. J. Joyce, and G. M. Atkinson, J. Vac. Sci. Technol. B **9**, 3079 (1991).
[10]Y. Yasuoka, K. Harakawa, K. Gamo, and S. Namba, Jpn. J. Appl. Phys., Part 2 **26**, L1453 (1987).

Focused ion beam direct deposition and its applications

Shinji Nagamachi[a] and Masahiro Ueda
Technology Research Laboratory, Shimadzu Corporation, Hikaridai 3-9, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Junzo Ishikawa
Department of Electronic Science and Engineering, Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto 606-01, Japan

(Received 10 December 1997; accepted 9 March 1998)

We developed focused ion beam direct deposition as a new method for fabricating patterned metal films directly on substrates. We designed and constructed a focused ion beam apparatus which satisfied demanded capabilities for direct deposition such as low energy and fine focused beam, high beam current density, high vacuum condition, changeability of ion species, precise and wide range patterning, sample observation by an optical microscope, and quick sample exchange. We also developed liquid alloy–metal ion sources for conductive materials, superconductive material and magnetic material. We tried to apply the focused ion beam direct deposition method to IC modification, surface acoustic wave (SAW) devices, SQUIDs, multilayers, and probing on small crystals. In SAW devices, SQUIDs, and multilayers, fabricated devices had comparable performance to devices fabricated by ordinary photolithographic processes. In IC modification and probing on small crystals, a low resistant and flexible connection was confirmed. We proved that focused ion beam direct deposition method is a useful tool for research and development such as prototyping. © *1998 American Vacuum Society.* [S0734-211X(98)04604-6]

I. INTRODUCTION

We developed focused ion beam direct deposition as a new method for fabricating patterned metal thin film directly on substrates.[1–4] The principle of this method is to perform ion beam deposition by using a very low energy (30–200 eV) focused ion beam. If we compare the direct deposition method to the focused ion beam induced deposition method,[5] some merits can be expected. Both methods are maskless and resistless processes. In focused ion beam induced deposition, a metalorganic gas is supplied to the surface of a substrate and absorbed gas molecules are dissolved into metal atoms on a substrate by irradiation of an accelerated (20–50 keV) Ga ion beam. In the direct deposition method, a high vacuum condition maintained during deposition prevents a substrate from being contaminated with absorbed gas molecules. It also prevents a thin film from being mixed with fragments of gas molecules and irradiated ions. These contaminations are predicted in the induced deposition method. To use a very low energy beam in direct deposition prevents a substrate and a thin film from inducing irradiation damage which is also predicted in induced deposition. A large current density of a focused ion beam has an influence on the purity of direct deposited thin film. We estimate the expected purity of deposited films from a current density and a residual gas pressure in the next section.

There were two major technical subjects for focused ion beam direct deposition. One was to develop a low energy focused ion beam apparatus and the other was to develop ion sources for direct deposition. We designed and constructed a low energy focused ion beam apparatus and developed liquid–alloy ion sources for conductive, superconductive, and magnetic materials. We measured some basic properties of directly deposited films and started development of applications. In this article, we mainly report the applications of direct deposition which were tried.

II. EXPERIMENTAL APPARATUS AND BASIC PROPERTIES OF DEPOSITED FILMS

A. Experimental apparatus

There are several capabilities which are demanded in a direct deposition apparatus. The most important capability is to get a fine focused ion beam in the energy range of 30–200 eV. We calculated optical properties of various lens shapes and arrangements to find an optimized lens system for a very low energy beam.[4]

The beam current density is a very important parameter for direct deposition. We can estimate the expected purity of a deposited film from the beam current density and the residual gas pressure. If we assume P as the expected impurity concentration, P can be written by the flux density of residual gas (J_{gas}), the sticking probability of residual gas (S_{gas}), the flux density of ion beam (J_{ion}), and the sticking probability of ion beam (S_{ion}) as shown in Eq. (1):

$$P=\frac{J_{gas}\times S_{gas}}{J_{gas}\times S_{gas}+J_{ion}\times S_{ion}}. \tag{1}$$

The flux density of residual gas is calculated by Eq. (2) where p is the pressure of residual gas, m is the mass of residual gas molecules, k is Boltzmann's constant, and T is the temperature:

[a]Electronic mail: nagamach@shimadzu.co.jp

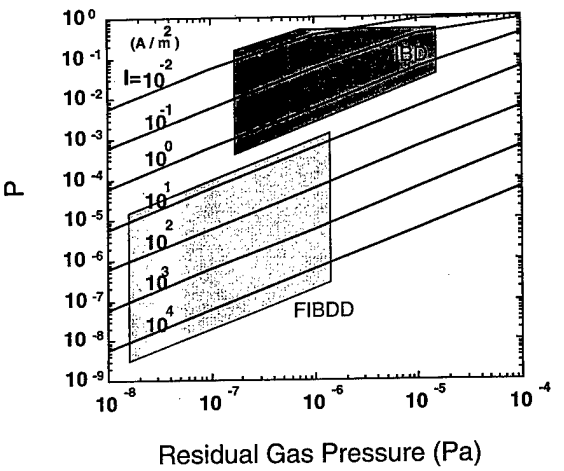


FIG. 1. Expected impurity concentration (P) as a function of residual gas pressure and beam current density on the assumption that sticking probabilities of gas molecules and ions are unity.

$$J_{\text{gas}}=\frac{p}{\sqrt{2\pi mkT}}. \tag{2}$$

If we assume the main contribution of residual gas as H_2O, which was the main contribution in our apparatus, the temperature as room temperature and express the pressure in Pa, we get

$$J_{\text{gas}}=3.59\times10^{22}\ \text{p};\quad (\text{molecules m}^{-2}\ \text{s}^{-1}). \tag{3}$$

The flux density of ion beam is calculated from a beam current density (J_I) and a charge of ion (q) as shown in Eq. (4):

$$J_{\text{ion}}=\frac{J_I}{q}, \tag{4}$$

and if we assume the ion as a single charged ion and express the current density in A/m^2, we get

$$J_{\text{ion}}=6.02\times10^{18}J_I\quad (\text{ions m}^{-2}\ \text{s}^{-1}). \tag{5}$$

If we assume S_{gas} and S_{ion} as unity, we get results shown in Fig. 1. Expected impurity concentrations of both focused ion beam direct deposition and ordinary ion beam deposition with a plasma ion source are shown in the figure. It promises that focused ion beam direct deposition is a noble method to fabricate pure metal films. Beam current density and residual gas pressure should always be considered in focused ion beam direct deposition and ordinary ion beam deposition.

The usage of various ion species is also important for direct deposition. It can be accomplished by using liquid–alloy ion sources and a mass filter. Other capabilities, such as precise and wide range patterning by deflector and a sample stage, sample observation by an optical microscope and quick sample exchange, are necessary for the practical direct deposition apparatus.

We designed and constructed a focused ion beam apparatus which satisfies the described demanded capabilities. Figure 2 shows a schematic view of our apparatus. By our apparatus, the beam diameter of Au^+ ion beam in the energy

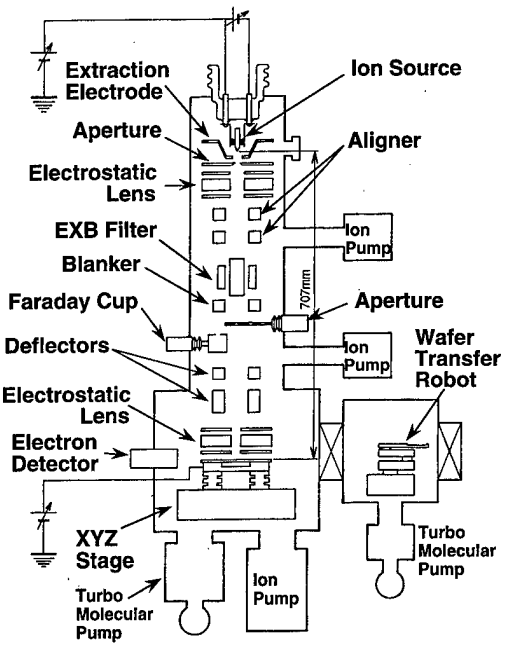


FIG. 2. Schematic view of focused ion beam direct deposition apparatus.

range of 30–200 eV could be tuned between 0.4 and 7 μm while the beam current varied from 40 pA to 10 nA. The beam current density of Au^+ beam was about constant value in this beam current region and was 300 A/m^2. The residual gas pressure was $2-7\times10^{-7}$ Pa during operation. The precise explanations about our focused ion beam apparatus and its beam characteristics were reported in Refs. 3 and 4.

We also developed liquid metal ion sources for direct deposition. We adopted an impregnated-electrode-type liquid metal ion source tip[6] made of tungsten. For conductive materials (Au, Cu, and Al) we developed a Au–Si alloy source, a Au–Cu alloy source, a Au–Cu–Si alloy source, a Au–Cu–Ge alloy source, an Al–Au–Ge alloy source, a Au–Cu–Al alloy source, and a Au–Cu–Al–Ge alloy source. For superconductive material (Nb) we developed a Nb–Au–Cu alloy source and for magnetic material (Co) we developed a Co–Cu–Nb–Au alloy source. These ion sources satisfy practical stability of extracted beam and have long lifetime (>100 h). A detailed report on the developed liquid–alloy ion sources will be published elsewhere.

B. Basic properties of directly deposited films

We measured some basic properties of directly deposited films such as the purity of Au film, resistivities of Au, Cu, and Al films and the superconductivity of Nb film.[1–3] We measured the impurity concentration in the deposited Au film on the Si substrate by Auger electron spectroscopy and secondary ion mass spectroscopy. The concentration of C and O was estimated below 100 ppm and was consistent with the expected value shown in Fig. 1. Resistivities of directly deposited Au and Cu films were 1.2–1.6 times larger than those of bulk metals. In the case of Al film, it was 2.2–2.7 times larger than that of bulk Al. The reason why the resistivity of Al film is so large is not clear. The superconductive

TABLE I. Relation between merits of direct deposition and related applications.

Merits of FIBDD \ Applications	IC modification	SAW device	SQUID	multi layer	probing on small crystal
maskless process	○	○	○	○	○
damage-free process	○	○			○
clean process	○	○			
high purity	○	○	○	○	○
3D fabrication	○	○			○
combination of ion species		○	○	○	
combination of ion energy	○	○	○		

critical temperature of deposited Nb film was 8.5–8.7 K for appropriate deposition conditions. These results for basic properties promise us wide range applicability of focused ion beam direct deposition method.

One problem of ion beam deposition is a charging-up phenomenon which increases the surface potential and disturbs deposition on insulative substrates. In the case of focused ion beam direct deposition, there is a way to perform deposition without charging-up. If we start a deposition from a conductive area to an insulative area, the deposited metal film makes electric connection between the conductive area and the depositing position. This principle can be applied to a dielectric substrate with a metal pattern which is connected to common potential through some resistance. We adopted this technique to fabricate surface acoustic wave (SAW) devices as described below.

III. APPLICATIONS OF FOCUSED ION BEAM DIRECT DEPOSITION

We are sure that unique characteristics of direct deposition enable it to apply unique applications. The relation between merits of direct deposition and related applications which we tried is displayed in Table I. In the following, each application listed in Table I is discussed separately.

A. IC modification

IC modification for failure analysis is one of the established applications of Ga focused ion beam.[7] Two independent conductive lines which are covered by passivation layer are connected to each other by the following procedure. The first step of IC modification is sample observation to find the point to be modified by the function of scanning ion microscope. The second step is milling the passivation layer to expose the conductive lines. The third step is making a connection between two lines by focused ion beam induced deposition. We apply direct deposition method to the third step. Small electric resistance of the connection can be expected. We tried to make a connection by 54 eV Au^+ beam between two Al conductive lines whose distance was 200 μm. The fabricated connection pattern was 60 μm in width and 0.15 μm in thickness. The observed resistance of the connection was below 1.0 Ω and was consistent with the calculated resistance (0.8 Ω) from the measured resistivity of directly deposited Au (3.7 $\mu\Omega$ cm).

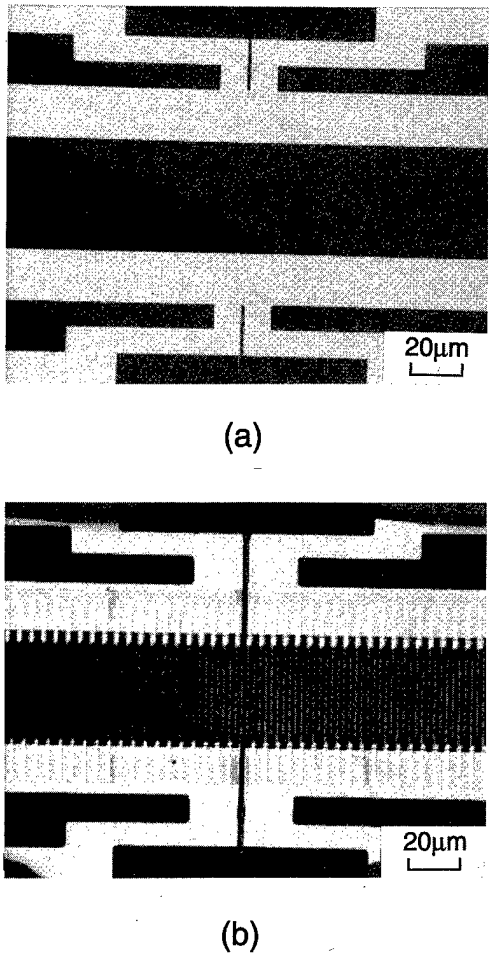


FIG. 3. Photograph of fabricated SAW filter (a) before and (b) after Al IDT direct deposition.

B. SAW device

SAW device[8] is widely used as a rf filter. Input electric power is transduced to mechanical vibration called SAW by an interdigital transducer (IDT) which looks like a pair of comb, on a piezoelectric substrate. SAW is transduced to electric power again by another IDT. A large flexibility on the design is expected for directly deposited IDTs.

We fabricated UHF band-pass filters with directly deposited Au and Al IDTs on $LiNbO_3$ ($64°YX$) substrates. Before IDT deposition, we prepared feeder electrodes and contact pads on the substrate as shown in Fig. 3(a). These electrodes and pads were fabricated by Al sputtering deposition, photolithography, and wet etching process. We started the deposition of the IDT from the feeder electrode to the piezoelectric surface. Figure 3(b) shows an example of a fabricated SAW

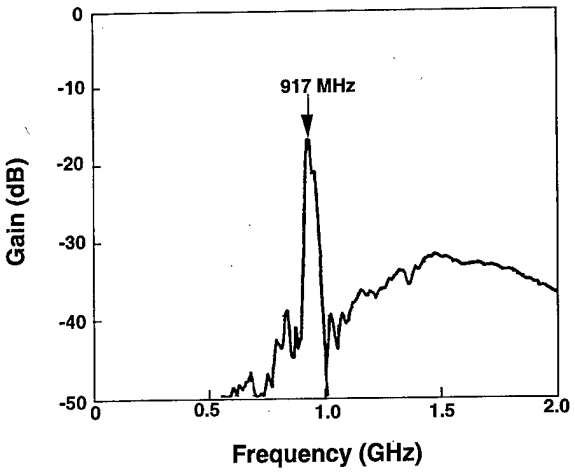


FIG. 4. Measured characteristic of the fabricated sample which has a directly deposited Al IDT. The number of the IDT is 41 pairs for input and 41 pairs for output. The periodic length is 4.8 μm.

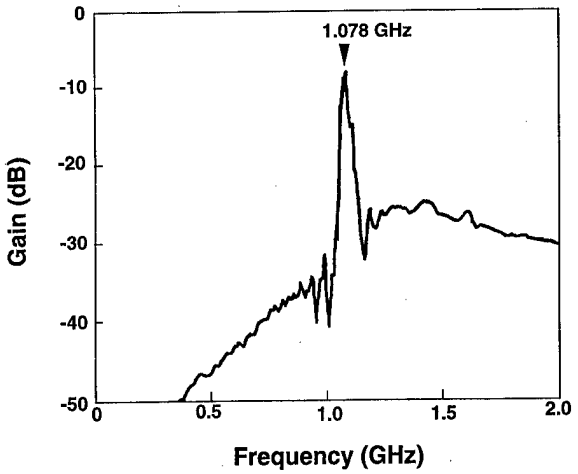


FIG. 6. Measured characteristic of the fabricated sample which has a directly deposited Au IDT. The number of IDT is 41 pairs for input and 41 pairs for output. The periodic length is 4.0 μm.

filter whose IDT was formed by a 32 eV Al^+ beam deposition. The periodic length of this IDT was 4.8 μm (so the interval to next IDT was 2.4 μm). The mean thickness of IDT was 10–20 nm. For step coverage, the thickness was modified to 60 nm around the step. We also fabricated Al IDT filter by lithographic process as a referenced sample.

We measured the characteristics of fabricated filters with a network analyzer without packaging. Figure 4 shows a measured relation between frequency and electric power gain through the filter, which has directly deposited Al IDTs (41 pairs for input/41 pairs for output, 10 nm in thickness, 4.8 μm in periodic length, and 48 μm in length). A peak is observed whose height is about 25 dB and the peak frequency is consistent with the calculated value from the periodic length. Figure 5 shows a measured characteristic of the referenced sample fabricated by lithographic process which has Al IDTs (81 pairs/81 pairs, 80 nm in thickness, 4.8 μm in period, and 48 μm in length), and it looks almost the same as the characteristic of the directly deposited sample except the gain loss at the peak frequency. The gain loss of the referenced sample is little less than that of the Al IDT sample. It may be caused by a difference of number of IDT. Figure 6 shows a result of the directly deposited Au IDT sample (41 pairs/41 pairs, 20 nm in thickness, 4.0 μm in period, and 48 μm in length). The result is also almost the same as that of Al IDT sample, but the gain loss is less than those of the Al IDT sample and the referenced sample. These evaluations indicate that electric properties of directly deposited SAW filters are almost comparable to lithographic patterned SAW filters, though the precise conditions of fabrication are slightly different from each other. The broad peak whose peak frequency was about 1.5 GHz was caused by capacitive coupling between input and output feeders, because the same peak was observed in measurement of a sample without IDT.

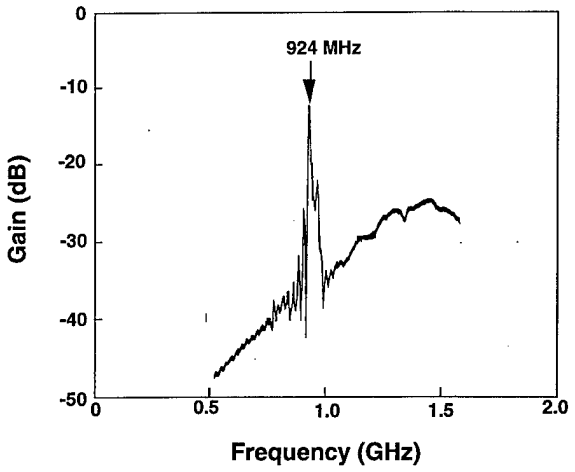


FIG. 5. Measured characteristic of the referenced sample which has an Al IDT. The number of the IDT is 81 pairs for input and 81 pairs for output. The periodic length is 4.8 μm.

C. SQUID

To apply a direct deposition method to the fabrication of the weak link of quasiplanar Josephson junction[9] was what first motivated us to start our research about focused ion beam direct deposition. We expected good controllability and reproducibility of the critical current in Josephson junction. Figure 7 shows a schematic view of a weak link fabrication and a quasiplanar Josephson junction. Before direct deposition of the weak link, all the structures except weak link were prepared by the sputtering deposition of Nb, photolithography, reactive ion etching, and anodic oxidization. The thickness of the lower Nb electrode, Nb_2O_5 insulator and upper Nb electrode was 300 nm, 30–40 nm, and 200 nm, respectively.

The area to be directly deposited was sputter etched by the irradiation of a 20 keV Au^+ beam to remove the naturally oxidized surface layer. The weak link whose size was 0.6 μm in width and 10 nm in thickness was deposited by a 108 eV Nb^{2+} beam. The weak link was covered by a Au film

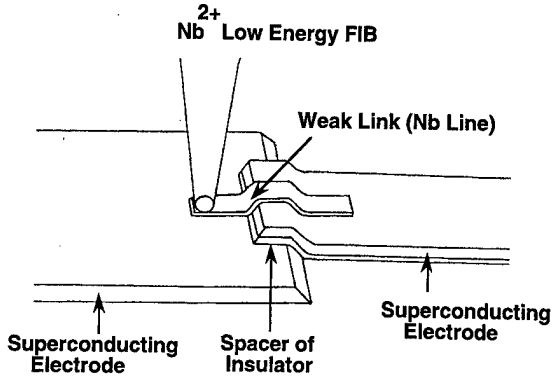


FIG. 7. Schematic view of quasiplanar Josephson junction and fabrication of a weak link by direct deposition.

($1\ \mu$m in width and 4 nm in thickness) deposited by a 54 eV Au^+ beam, to prevent the surface of Nb weak link from being naturally oxidized. Figure 8 shows an example of current–voltage characteristic of fabricated Josephson junction measured in a He cryostat. The critical current of this sample was 42 μA. We constructed a dc-SQUID flux meter by directly deposited Josephson junctions. Measured flux noise level was $7.5\times10^{-6}\phi_0$ $Hz^{-1/2}$ at 100 Hz and was sufficient for biomagnetic measurement.[10]

The expected controllability and reproducibility were not accomplished yet, because of probable poor reproducibility of prepared structure, especially the shape of step.

D. Multilayer

In the case of electrostatic lens system, every ion species extracted from the same ion source has the same optical property in principle. So, a change of ion species and getting an optimized beam can be done by tuning a mass filter without turning the lens system. This feature enables a direct deposition method with an alloy ion source to fabricate patterned multilayer directly on a substrate. We used a Co–Cu–Nb–Au alloy ion source and tried to fabricate magnetic/nonmagnetic (Co/Cu) multilayers.[11] If controllability of thickness is an order of an atomic layer, we can expect exhibition of giant magnetoresistance (GMR) effect.[12]

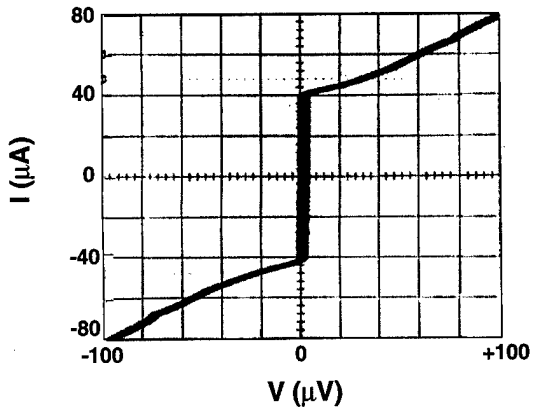


FIG. 8. Example of measured current–voltage relation of the fabricated Josephson junction.

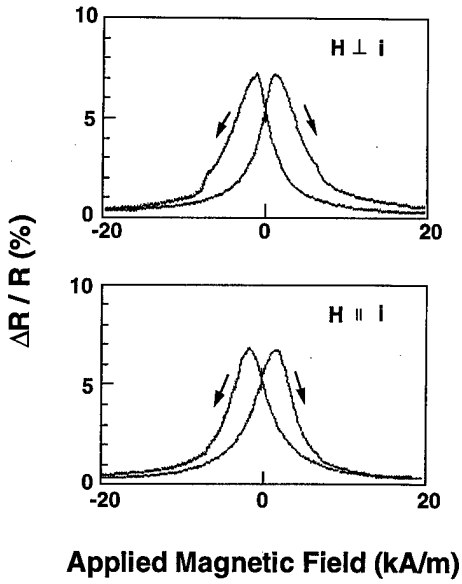


FIG. 9. Measured MR ratio ($\Delta R/R$) at room temperature as a function of applied magnetic field in the case of Co(2.0 nm)/Cu(2.1 nm) multilayer. Magnetic fields are applied in the plane of the sample perpendicular (upper) and parallel (lower) to the supplied current direction. Arrows in the figure show the direction of magnetic field scanning.

We fabricated multilayers which were $14\times76\ \mu m^2$ in size and consisted of 12 Co layers and 11 Cu layers by 108 eV Co^{2+} beam and 54 eV Cu^+ beam. The thickness of Co was 2.0 nm and the thickness of Cu was varied from 1.3 to 2.2 nm. After fabricating a multilayer, we deposited four-probe electrodes by 54 eV Au^+ beam across the multilayer.

MR ratios of the fabricated multilayers were measured by the four-point-probe method with 0–40 kA/m dc magnetic field applied in the film plane parallel and perpendicular to the supplied current direction at room temperature. The magnetization curves were measured with a vibrating sample magnetometer. Figure 9 shows an example of measured MR ratios ($\Delta R/R$) as a function of applied magnetic field. The lower side of the figure (the direction of the magnetic field was parallel to the current direction) shows GMR and the upper side (the direction was perpendicular) shows a sum of GMR and anisotropic MR, because GMR is isotropic and AMR (anisotropic magnetoresistance) is anisotropic. In this case, the MR ratio of GMR was 6.7% and that of AMR was 0.4%. Figure 10 shows measured GMR as a function of Cu layer thickness. There is a sharp peak at the thickness of 2.1 nm. It corresponds to the second maximum[12] which was observed in Co/Cu multilayers fabricated by sputtering deposition method, though the peak MR ratio was smaller than reported maximum MR ratio.[13] The exhibition of GMR shows not only controllability of direct deposition but also useful quality of magnetic film, and it has opened a new capability of direct deposition.

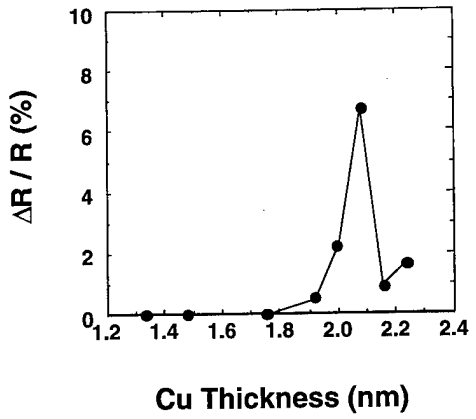


FIG. 10. Measured MR ratio ($\Delta R/R$) of GMR as a function of Cu layer thickness.

E. Probing on small crystals

In research on new material, a measurement of electric properties is often of great importance. If a sample is too small to directly attach probe electrodes and is too frail to fabricate probe electrode patterns by metal deposition, photolithography and etching process, focused ion beam direct deposition may be a unique technique to fabricate probe electrode patterns on such samples.

Figure 11 shows an example of probing on a small crystal by direct deposition of Au^+ beam. An organic crystal, which is $(DI\text{-}DCQNI)_2Li$ and is a species of a DCQNI (dicyanoquinonedi-imines) salt,[14] was 2 μm in diameter and 72 μm in length as shown in Fig. 11(a). The sample was on a semi-insulating GaAs substrate. Across the sample, four probe lines were deposited by 54 eV Au^+ beam as shown in Fig. 11(b). The thickness of probe lines was comparable to the sample diameter. These probe lines were extended and connected to pad electrodes which were also directly deposited by 54 eV Au^+ beam. Next, the substrate was fixed on an IC package and wire bonding between the pad electrodes on the substrate and electrodes on the IC package was carried out. Figure 11(c) shows the final feature of this sample. In this way, it is possible to measure electric properties of this sample from pins of the IC package.

The size and shape of sample is restricted. A large diameter or a large step influences the deceleration field and prevent a depositing line from covering the step. A diameter or a step may be limited within several microns. Too small sample size causes another restriction. We usually use the optical microscope assembled in the ion column for sample observation and beam positioning, so the sample size is limited above a few microns. It is necessary to use a scanning ion microscope to observe a smaller sized sample though it may cause irradiation damage.

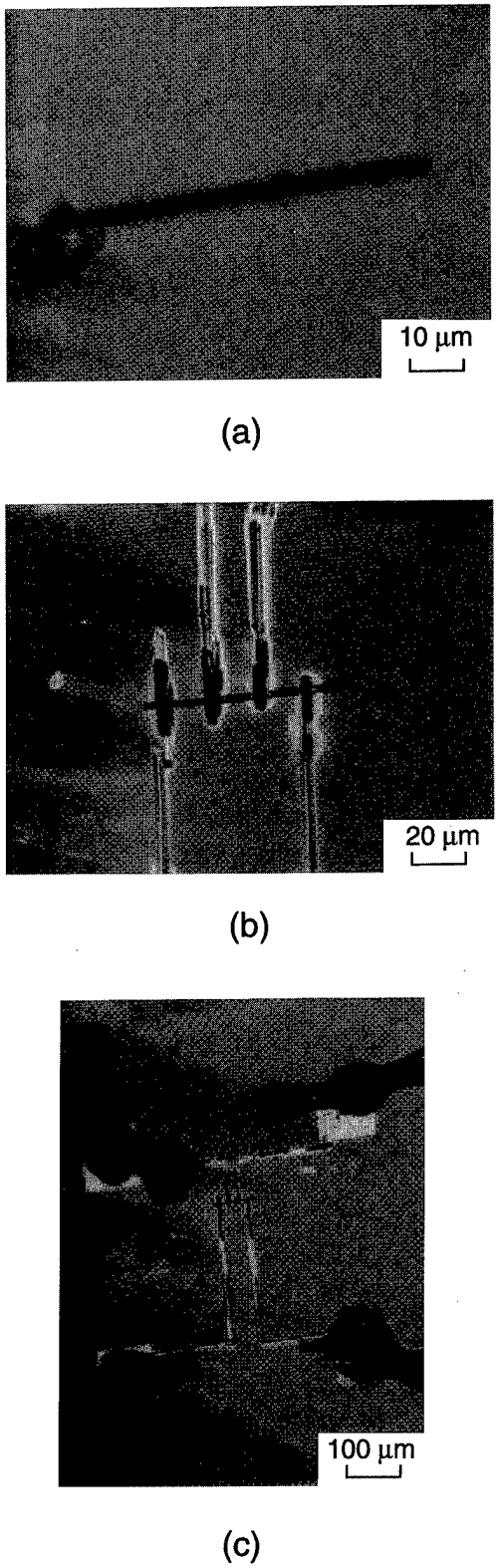


FIG. 11. Photograph of probing on a small crystal (a) before deposition, (b) after four-point probe line deposition, and (c) after mounting on the IC package and wire bonding.

F. Restrictions on applications

There are two restrictions on applications of the direct deposition method. One is sample size limitation. Even if we use a 10 nA Au^+ beam which is the maximum beam current in the current region keeping a high current density, the deposition rate is about 1 $\mu m^3/s$. Therefore, it takes 17 min to fabricate a $100\times100\times0.1$ μm^3 gold pattern by a 10 nA Au^+ beam, but it will take 116 days to fabricate a

10 mm×10 mm×0.1 μm gold pattern. The other is the inapplicability to mass production. To fabricate one sample by focused ion beam direct deposition, it usually takes within a half-day including programming time for deflector control. In the case of photolithographic process, it will take at least one week to fabricate a new sample, but it will take the same time to fabricate a million samples.

IV. CONCLUSION

We designed and constructed focused ion beam direct deposition apparatus which has sufficient capabilities for the development of practical applications. We tried to apply a focused ion beam direct deposition method to IC modification, SAW devices, SQUIDs, multilayers, and probing on small crystals. In SAW devices, SQUIDs and multilayers, fabricated devices had comparable performance to devices fabricated by ordinary photolithographic processes. In IC modification and probing on small crystals, a low resistant and flexible connection was confirmed.

Though this method is inapplicable to mass production, we proved that the focused ion beam direct deposition method is a useful tool for research and development such as prototyping.

ACKNOWLEDGMENTS

The authors are grateful to K. Shinada, Y. Fujiyama, and Y. Yamada of Shimadzu Corp. for preparing substrates and measurement of SQUID, to T. Hata, K. Totani, and H. Takasu of Rohm Co., Ltd. for preparing substrates, measurement, and discussions of SAW devices, to Dr. H. Sakakima and M. Satomi of Matsushita Electric Industrial Co., Ltd. for measurement and discussions of multilayer and to Dr. S. Asaka and K. Kato of the Institute for Molecular Science for supplying samples, measurements, and discussions of probing on a small crystal.

[1]S. Nagamachi, Y. Yamakage, H. Maruno, M. Ueda, S. Sugimoto, M. Asari, and J. Ishikawa, Appl. Phys. Lett. **62**, 2143 (1993).
[2]S. Nagamachi, Y. Yamakage, M. Ueda, H. Maruno, K. Shinada, Y. Fujiyama, M. Asari, and J. Ishikawa, Appl. Phys. Lett. **65**, 3278 (1994).
[3]S. Nagamachi, Y. Yamakage, M. Ueda, H. Maruno, and J. Ishikawa, Rev. Sci. Instrum. **67**, 2351 (1996).
[4]S. Nagamachi, M. Ueda, Y. Yamakage, H. Maruno, and J. Ishikawa, Rev. Sci. Instrum. **68**, 2331 (1997).
[5]J. Melngailis and P. G. Blauner, Mater. Res. Soc. Symp. Proc. **147**, 127 (1989).
[6]J. Ishikawa and T. Takagi, J. Appl. Phys. **56**, 3050 (1984).
[7]K. Nikawa, J. Vac. Sci. Technol. B **9**, 2566 (1991).
[8]M. G. Holland and L. T. Claiborne, Proc. IEEE **62**, 582 (1974).
[9]H. Ohta, IEEE Trans. Electron Devices **27**, 2027 (1980).
[10]D. Drung, Supercond. Sci. Technol. **4**, 377 (1991).
[11]S. Nagamachi, M. Ueda, H. Sakakima, M. Satomi, and J. Ishikawa, J. Appl. Phys. **80**, 4217 (1996).
[12]H. Sakakima, M. Satomi, Y. Irie, and Y. Kawawake, Trans. Mater. Res. Soc. Jpn. **15**, 1083 (1994).
[13]S. S. P. Parkin, Z. G. Li, and D. J. Smith, Appl. Phys. Lett. **58**, 2710 (1991).
[14]A. Aumuller, P. E. G. Klebe, S. Hunig, J. V. von Schutz, and H. P. Werner, Angew. Chem. Int. Ed. Engl. **25**, 740 (1986).

Transmission electron microscopy observation of thin foil specimens prepared by means of a focused ion beam

H. Saka[a)]
Department of Quantum Engineering, Nagoya University, Nagoya 464-8603, Japan

(Received 17 April 1998; accepted 28 May 1998)

Recent development in the application of a focused ion beam to preparation of thin foil specimens for transmission electron microscopy observation has been reviewed. It has been shown that this technique is very powerful for characterization of a variety of industrial materials. © *1998 American Vacuum Society.* [S0734-211X(98)11304-5]

I. INTRODUCTION

Transmission electron microscopy (TEM) is now widely used for characterization of the microstructure of a variety of industrially important materials. However, recent demands imposed on TEM specimens has made it increasingly difficult to rely only on the traditional methods of preparation of TEM specimens such as electropolishing, chemical polishing, and even ion milling technique. This is the reason why focused ion beam (FIB) milling is gathering much attention as a new technique to prepare TEM specimens.[1–3]

There are at least two advantages to using FIB to prepare TEM specimens. One is that it is possible to prepare a TEM specimen with pin-point accuracy from a preselected area. The other is that a TEM specimen can be easily prepared from a material such as a multilayered structure, which is composed of different substances—the physical and/or chemical properties of which are quite different from one another.

In this article recent development in the combined use of FIB and TEM(FIB-TEM) for the characterization of industrial materials will be reviewed with special reference to the aforementioned two advantages.

II. PIN-POINT SPECIMENS

A. Failure analysis of TFT-LCD

A thin-film transistor liquid crystal display (TFT-LCD) is now widely used for a personal computer display. A TFT is composed of more than ten layers (with a dimension of nanometer-scale) which are deposited and fabricated on a large LCD-grade glass substrate. FIB-TEM is now indispensable for failure analysis of TFE-LCD.

Figure 1(a) shows an optical micrograph of a faulty junction detected by an array tester. The junction has short-circuited the insulating film. In this case no gross anomaly at the junction is observed in the optical micrograph, corresponding to the position indicated by the bar in Fig. 4(a). However, cross-sectional TEM observation [Fig. 1(b)] clearly reveals that a microvoid is formed where a metal conductor passes over a step.[4] The failure analysis of LSI is described by Ishitani *et al.*[5]

B. Whisker and hillocks on Al film

In TFT-LCD, Al and Al alloys have gathered increasing attention as potential candidates for wiring materials in higher resolution TFT-LCD because of their low electric resistance. However, in the case of Al-based metals protrusions such as hillocks and whiskers are formed due to the stress evolved during processing. Formation of such protrusions gives rise to a serious problem such as short-circuit between the gate metal and the source/drain electrodes located in an upper layer.

For a better understanding of the mechanism of formation of whiskers and hillocks, it is important to have a detailed knowledge about crystallography and microstructure of whiskers and hillocks formed on Al lines. For such a purpose, FIB-TEM is very powerful. There are two methods for such an observation, i.e., cross-sectional and plan-view observations. The detailed procedures have been published elsewhere.[6–8]

Figure 2 shows a cross-sectional dark-field TEM micrograph of a single crystalline whisker. It is clear that the orientation of the whisker and that of the grain from which the whisker grew are quite different. Figure 3 shows a similar TEM micrograph but of a kinked polycrystalline whisker. Figure 3(a) was taken in the bright-field imaging mode. The whisker must have started to grow at the bottom part of the original grain inside the Al film and then suddenly have kinked at the root. Figures 3(b)–3(e) were taken in the dark-field imaging mode; the root of the whisker is composed of polycrystals. From Figs. 2 and 3 it is clear that the growth of these whiskers initiates at the interface between Al film and the glass substrate.

Figure 4 shows an example of the plan-view observation of a whisker. It is evident that the whisker grows from inside the grain *A* and not from grain boundary nor a triple point of grain boundaries. Also the size of that grain from which the whisker grew is not necessarily larger than the nearby grains.

a)Electronic mail: saka@numse.nagoya-u.ac.jp

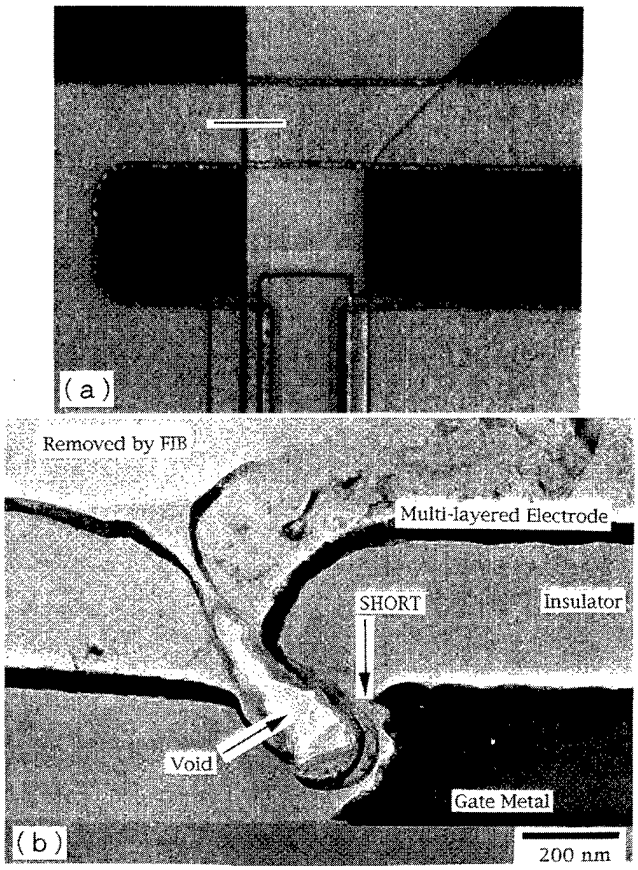


FIG. 1. (a) Optical micrograph of a faulty junction. (b) Cross-sectional TEM micrograph corresponding to the position indicated by a bar in (a).

C. Crack

The fracture of materials is of great practical as well as academic importance. Many attempts have been made to observe a crack tip using optical microscopy, x-ray topography, and TEM. Of these techniques, the resolution of optical microscopy and x-ray topography is too poor. By contrast, the resolution of TEM is high enough for detailed information to be obtained. However, in order to carry out TEM observation, a thin foil specimen must be prepared, and it is usually extremely difficult to prepare such a specimen from a preselected region containing a crack tip.

Single crystal of Si, MgO, and alumina were used. Vickers indentations were performed in air, at room temperature, using a 50 g load and a 15 s dwell time. Indentation resulted in radial cracks which are approximately perpendicular to each other. After indentation, the specimen was transferred into a FIB system and two trenches were milled in such a way that a thin wall is left behind between the two trenches, the wall being thin enough to be transparent to electrons when tilted by 90° (Fig. 5).[9–12]

Figure 6 shows an optical micrograph of cracks formed at an indent on a (001) surface of Si. The cracks were parallel to ⟨110⟩ directions in this case. Figure 7 shows an electron micrograph of the area of the indent (denoted by I). The following features are evident:

(1) Within the indent the crystal is deformed severely (P) and some dislocations (denoted by A) have glided out of the indent. They have a tendency to lie on the {111} plane.
(2) Near the indent, cracks piercing the foil are observed. There are two types of such cracks. One is almost parallel to the foil surface, i.e., (001) plane (L_1,L_2,L_3,L_4,L_5). The other type is perpendicular to the foil surface and lies on the {110} plane (H_v).
(3) On both sides of the indent and underneath the indent, fringe contrasts can be seen (denoted by H_1). It is natural to consider the areas with fringe contrast as containing cracks. The fringes can be explained by moiré fringes which result from the crystal being separated in two by the crack.

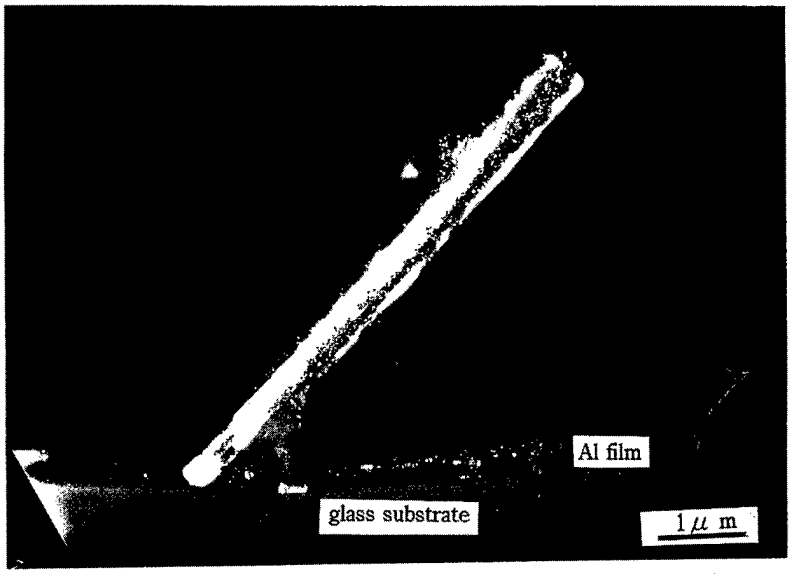


FIG. 2. Cross-sectional TEM micrograph of a single crystalline whisker imaged in the dark-field mode.

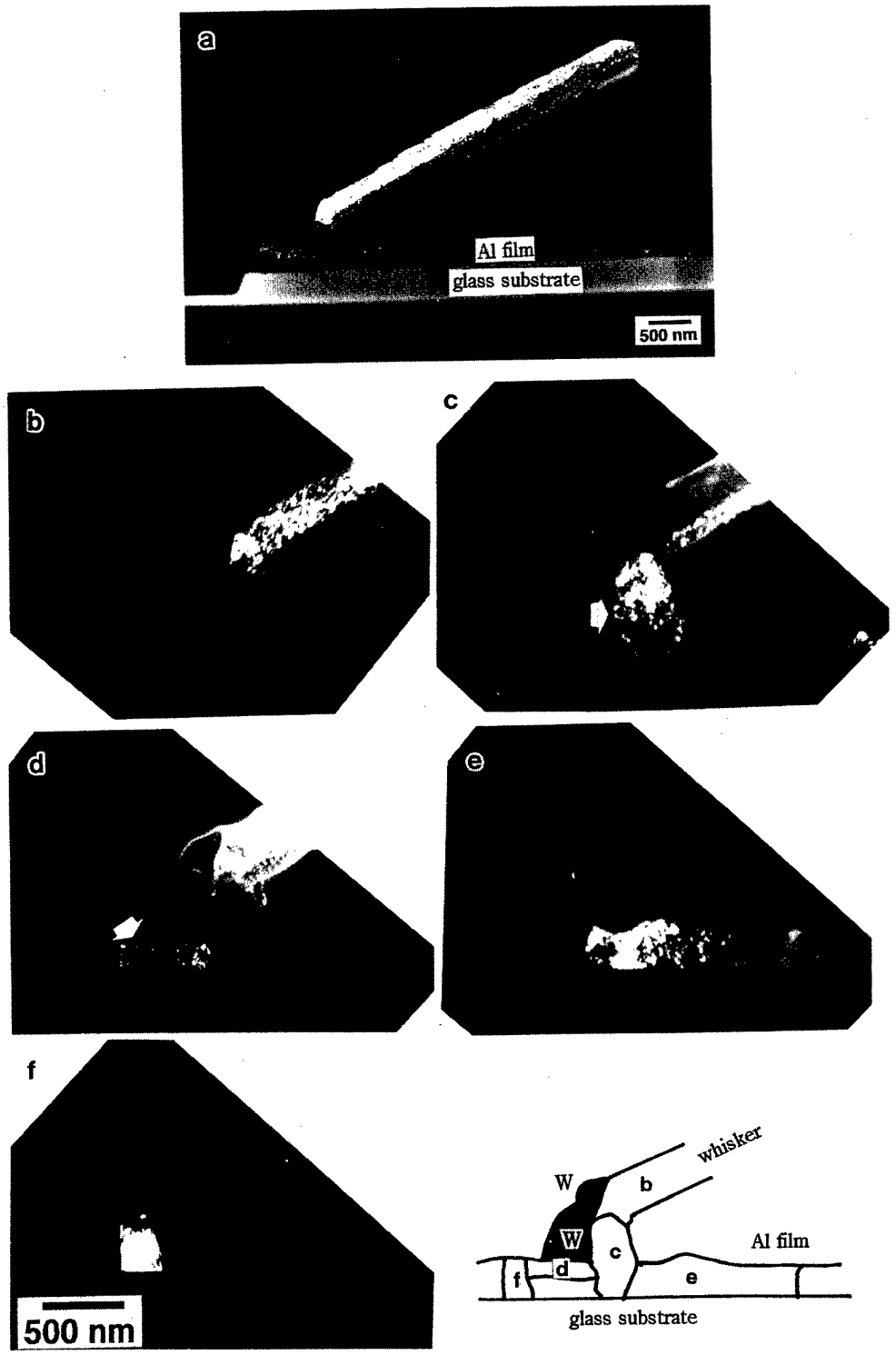


FIG. 3. Cross-sectional TEM micrographs of a kinked polycrystalline whisker. In (g) morphology of Al grains near the root of the whisker is shown. Figures in each of the grains correspond to those dark-field images (b)–(f) where the grain appears bright. W denotes a W-deposited protective layer.

III. MULTILAYERED STRUCTURE

A. GA steel

A galvanneled steel (GA steel) has been used widely for an automobile body steel sheet. A GA steel is produced by dipping a steel sheet into a molten bath of zinc, followed by a continuous annealing, where molten zinc reacts with the substrate steel to form a multilayer microstructure consisting of different kinds of intermetallic compounds.

Most of the previous studies have been carried out using optical microscopy and scanning electron microscopy (SEM) because of the difficulty with which a cross-sectional TEM

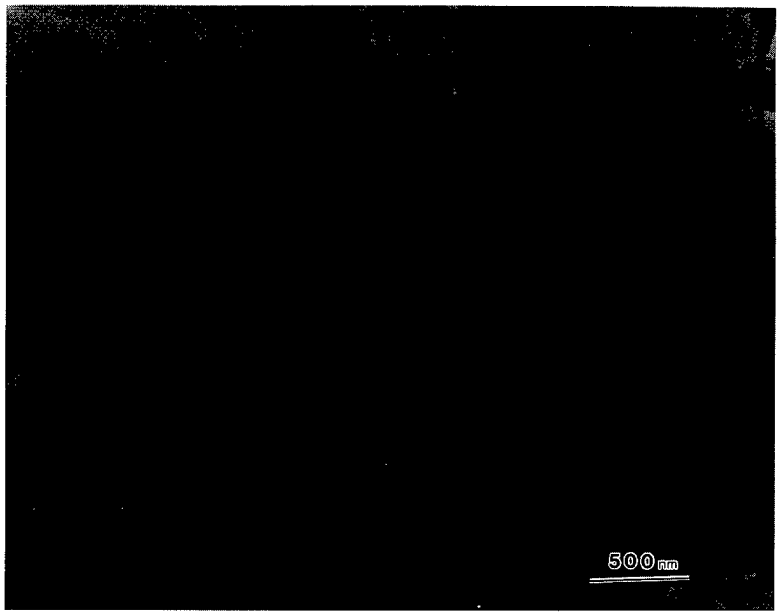


FIG. 4. Plan-view observation of a whisker.

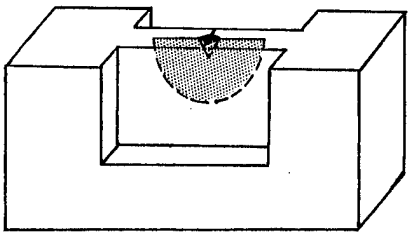


FIG. 5. Geometry of the foil specimen.

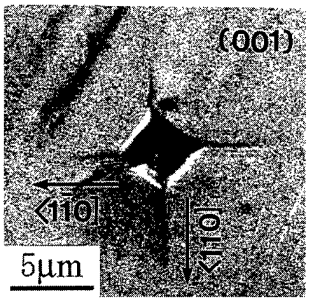


FIG. 6. Optical micrograph of an indent on a (001) surface of Si.

specimen is prepared from a GA steel. However, more recently FIB has been successfully applied to a GA steel.[13–16]

Figure 8 shows a typical TEM of a cross section of a GA steel. The whole coating from the outermost surface down to the steel substrate can be observed. Each grain was identified using electron diffractions. Figures 9(a) and 9(b) shows schematically all the layer structure of the coating of a GA steel as suggested from SEM and TEM observations, respectively. The result obtained from TEM observation is different from that suggested from SEM observation, except for the Γ layer. Especially, in SEM Γ_1 phase and δ_{1k} phases are missing.

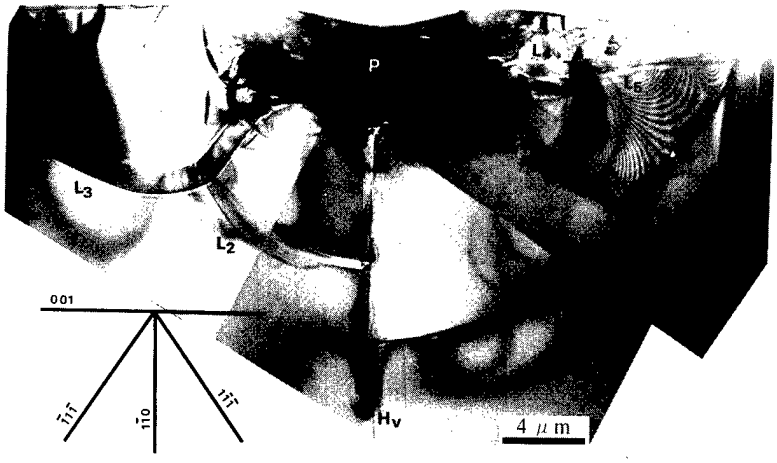


FIG. 7. Electron micrograph around an indent formed in Si at room temperature.

FIG. 8. Cross-sectional TEM micrograph of a GA steel. W is a W-deposited protective layer.

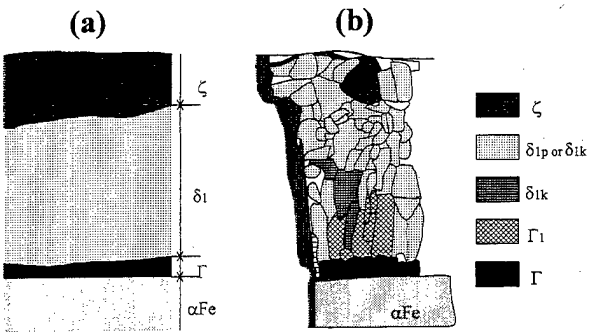


FIG. 9. Comparison of layer structure as suggested by SEM(a) and by TEM(b).

Needless to say, the result obtained from TEM observation is much more reliable.

B. Diamond coating on Si_3N_4

Diamond coating is used for a cutting tool. The performance of such a tool is governed mostly by the microstructure of an interface between diamond and the substrate. For characterization of such an interface, cross-sectional TEM observation should be very useful. However, because diamond is very hard, it is usually very difficult to prepare a cross-sectional TEM specimen from such a diamond-coated material by means of the traditional techniques. FIB is again very powerful in preparing such a specimen.[17]

Figure 10 shows an example of a cross-sectional TEM micrograph of a microstructure near an interface between diamond and Si_3N_4. The coating was made using the chemical vapor deposited (CVD) method. Microstructure of diamond, Si_3N_4 and the interface between them is clearly observed.

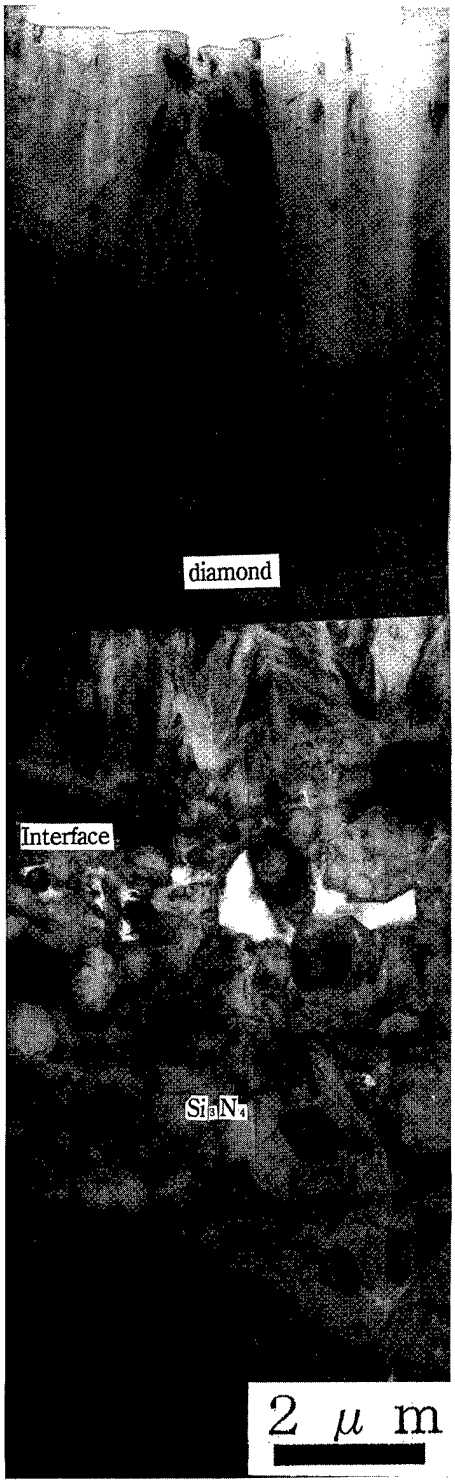


FIG. 10. Microstructure near an interface between Si_3N_4 substrate and a diamond coat.

ACKNOWLEDGMENTS

The author thanks Professor K. Kuroda, Dr. K. Sasaki, and T. Kato of Nagoya University, Dr. S. Tsuji and K. Tsujimoto and H. Takatsuji of IBM Japan for their cooperation.

[1]R. J. Young, E. C. G. Kirk, D. A. Williams, and H. Ahmed, Mater. Res. Soc. Symp. Proc. **199**, 205 (1990).
[2]K-h. Park, Mater. Res. Soc. Symp. Proc. **199**, 271 (1990).
[3]T. Ishitani and T. Yaguchi, Microsc. Res. Tech. **35**, 320 (1996).
[4]S. Tsuji, K. Tsujimoto, N. Tsutui, N. Miura, K. Kuroda, and H. Saka, Thin Solid Films **281/182**, 562 (1996).
[5]T. Ishitani *et al.*, J. Vac. Sci. Technol. B, these proceedings.
[6]K. Tsujimoto, S. Tsuji, H. Takatsuji, K. Kuroda, H. Saka, and N. Miura, Mater. Res. Soc. Symp. Proc. **480**, 207 (1997).
[7]H. Saka, K. Tsujimto, S. Fujino, K. Kuroda, H. Takatsuji, and S. Tsuji, *Proceedings of 4th International Workshop on Stress Induced Phenomena in Metallization*, AIP Conf. Proc. No. 418 (AIP, New York, 1977), p. 371.
[8]K. Tsujimoto, S. Tsuji, H. Saka, K. Kuroda, H. Takatsuji, and Y. Suzuki, Mater. Res. Soc. Symp. Proc. **441**, 421 (1997).
[9]H. Saka and G. Nagaya, Philos. Mag. Lett. **72**, 251 (1995).
[10]H. Saka, G. Nagaya, T. Sakuishi, S. Abe, and A. Muroga, Mater. Res. Soc. Symp. Proc. **409**, 45 (1996).
[11]H. Saka and S. Abe, J. Electron Microsc. **1**, 45 (1997).
[12]H. Saka and S. Abe, Mater. Sci. Eng., A **234/236**, 552 (1997).
[13]H. Saka, K. Kuroda, M. H. Hong, T. Kamino, T. Yaguchi, H. Tsuboi, T. Ishitani, H. Koike, A. Shibuya and Y. Adachi, Proceedings of the 13th International Congress on Electron Microscopy, 1994, p. 1009.
[14]H. Saka, T. Kato, M. H. Hong, K. Sasaki, K. Kuroda, and T. Kamino, *Galvatech'95* (Iron and Steel Soc., Inc., 1995), p. 809.
[15]T. Kato, M. H. Hong, K. Nunome, K. Sasaki, K. Kuroda, and H. Saka, Thin Solid Films (to be published).
[16]T. Kato, M. H. Hong, and H. Saka, J. Electron Microsc. (submitted).
[17]K. Kuroda, M. Takahashi, H. Itoh, and H. Saka, Thin Solid Films (to be published).

Dynamic Monte Carlo simulation for depth profiling by ion-sputter etching: Application to the AlAs/GaAs multilayered system

Hyung-Ik Lee[a] and Ryuichi Shimizu
Department of Applied Physics, Faculty of Engineering, Osaka University, Suita, Osaka 565, Japan

(Received 12 January 1998; accepted 4 May 1998)

We present a dynamical simulation approach by Monte Carlo calculation to describe atomic mixing phenomena in depth profiling of multilayered systems. This approach is based on the binary encounter model, taking into account (1) generation of both the interstitial atoms and vacancies and (2) annihilation of the vacancies. The results indicate that the simulation describes very well the depth profiles of AlAs/GaAs multilayered systems obtained by Auger electron spectroscopy. It predicts the existence of preferential sputtering of Al in the AlAs layer, where the Auger intensity of the As (*MVV*—32 eV) in the AlAs layer is about 1.2 times larger than that of the GaAs layer for 0.5 keV Ar^+ sputtering. © *1998 American Vacuum Society.* [S0734-211X(98)05904-6]

I. INTRODUCTION

Atomic mixing effect induced by ion bombardment has been intensively studied in experiment and theory.[1–4] In experiment, depth profiling has now been routinely used in semiconductor device technology. One of the most basic factors restricting the resolution of depth profiling is the atomic mixing effect. Theoretical approaches to describe the atomic mixing effect have also been intensively investigated so far. In particular, transport theory proposed initially by Hofer and Littmark[4] have revealed many important aspects of atomic mixing phenomena in sputter depth profiling. Although analytical approach for this theory provides a good result for simple cases, computer simulation approaches such as TRIDYN[5] or ACAT-DIFFUSE[6] codes have been applied as useful methods for samples with a complex boundary condition. We have recently developed a dynamic Monte Carlo approach for depth profiling.[7] This simulation approach is basically the same as TRIDYN in the aspect that the target sample remains of constant atomic density during sputter etching. However, in the present simulation constant atomic density is maintained by taking into account the recombination of the vacancy and interstitial atoms generated in the atomic mixing processes rather than relaxation performed in TRIDYN, enabling more realistic recombination processes based on the embedded-atom model to be incorporated.

In the present study, we have described a dynamic simulation approach in order to explain the atomic mixing effect in sputter depth profiling. The approach is based on further improvements of the dynamic Monte Carlo simulation scheme employed in the previous paper.[7] This simulation approach has been applied to describe the depth profiling of the AlAs/GaAs superlattice and verified by comparing with the experiment performed with 0.5 and 1.0 keV Ar^+ ions impinging on the specimen at angle of incidence of 55°. The results describe the experimental results very well, assuming that the surface roughness effect is negligible for the AlAs/GaAs multilayer system.[8]

[a]Electronic mail: hyungik@dragon.yonsei.ac.kr

II. SIMULATION MODEL

A. Binary encounter model

The present dynamic simulation approach is basically the same as the Monte Carlo simulation developed in our laboratory based on the binary encounter model using Lindhard's equation and the Ziegler, Biersack and Littmark (ZBL) potential to describe the electronic and nuclear stopping processes, respectively. In the practical Monte Carlo calculation procedures the functional representation of ZBL potential has been used to shorten the computing time.[9]

Application of this Monte Carlo simulation scheme to the depth profiling of AlAs/GaAs superlattice structure has revealed that the simulation still overestimates the atomic mixing effect considerably. Attempts have continuously been made to improve this situation by modifying the interatomic potential to get better agreement with the Auger depth profiles of AlAs/GaAs, leading to the consideration of the generation of vacancies and interstitial atoms, and of their annihilation during the atomic mixing calculation. Actually this dynamic Monte Carlo simulation has described atomic mixing in Cu–Pt alloys with considerable success.[10] However, the previous model leads to poorer depth resolution because atomic mixing region extended deeper. These attempts to improve the Monte Carlo simulation scheme for depth profiling have finally led to the present dynamic simulation approach based on the following improvements: (1) To describe atomic mixing in the low energy region, we estimated the maximum impact parameter P_{max} from $P_{max}=d/2$, where d is the mean interatomic distance of a specimen under ion bombardment. (2) In the previous dynamic Monte Carlo simulation approach, only those interstitial atoms that are displaced from adjacent layers are assumed to recombine with the vacancies to annihilate them.[7] In the present approach, we assume that all the atoms in the relevant layer may take part in the recombination with vacancies.

B. Atomic mixing model

The interstitials and vacancies generated by collisional cascade are distributed in pairs if the sputtered atoms are

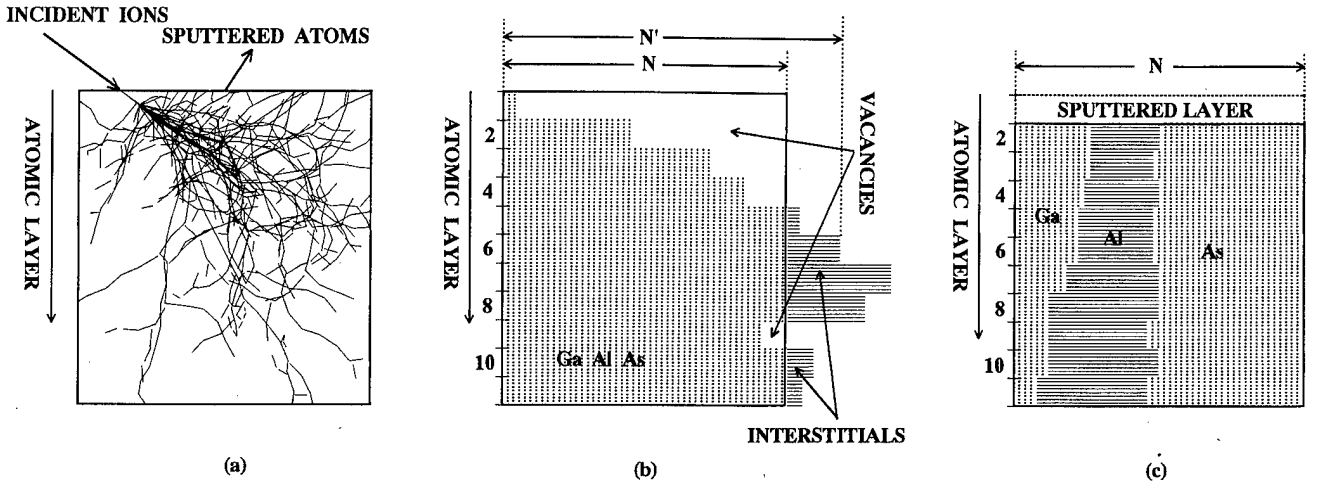


FIG. 1. Schematic diagram for redistribution of interstitials and vacancies: (a) Cascade mixing process. (b) The formation of interstitials and vacancies due to the cascade mixing process. (c) The result of redistribution in each layer. The result shows that the topmost layer is sputtered and the total number of atoms in each layer is preserved as N.

regarded as being sputtered only from the topmost atomic layer. These pairs migrate during collisional cascade and the vacancies are eventually annihilated by recombination with interstitials or absorbed into the sinks, e.g., surface, grain boundary, etc. Since the behavior of these interstitial atoms and vacancies is directly related to diffusion, sputtering and range of atomic mixing in a complicated manner, it is very difficult to describe the migration processes of these pairs accurately.[11] In the present simulation, therefore, all the interstitial-vacancy pairs are assumed to be completely recombined in the mixing zone. The total number of atoms of each layer, therefore, remains constant so that the atomic density in each atomic layer is kept invariant during sputtering. Otherwise, the atomic density tends to increase by a knock-in effect as sputtering proceeds.

The basic model used to treat the interstitial-vacancy pairs is as follows: First, it is assumed that each atomic layer is composed of N pseudoatoms, where a pseudoatoms represents $\sim 10^9-10^{10}$ atoms, depending on the number of trajectories of primary ions in the Monte Carlo simulation. The Monte Carlo calculation is performed until the total number of sputtered atoms attains the number necessary to sputter N pseudoatoms. As depicted in Fig. 1(a) atomic mixing is generated by the collision cascade of both the incident ions and recoiled atoms, and the atom displaced from original layer to another layer is counted as an interstitial atom, leaving a vacancy at its original site. The numbers of both the vacancies and interstitial atoms are easily counted for each atomic layer. When the total number of pseudoatoms in an atomic layer, N', is larger than N, $(N'>N)$, we regard the excess atoms, $(N'-N)$, as interstitials while the difference, $(N-N')$, is regarded as vacancies for $(N>N')$. This operation leads to the histogram as depicted in Fig. 1(b). This histogram shows the total number of Ga, Al and As atoms in each atomic layer. Some layers contain the vacancies and others the interstitial atoms as excess atoms. Second, redistribution of atoms in an ith atomic layer is calculated as follows. (1) Vacancies are assumed to be annihilated with interstitials in the previous $(i$-1)th layer and the composition of those atoms that fill the vacancies is proportional to atomic concentration in the relevant layer. (2) Interstitials in the ith layer are to be added to the previous $(i$-1)th layer. Note that the previous model also counted the numbers of vacancies and interstitials in each layer but it omitted annihilation of these vacancies with interstitials in the relevant layer by starting the above redistribution operation under the assumption that the vacancies are to be annihilated with only those interstitial atoms that were migrated from the other layers. In the previous work, this caused significant fluctuation of the composition in each atomic layer as sputter-etching proceeds.

The process described above is performed from the bottom layer to the top atomic layer. As a result, the total number of atoms of the top atomic layer becomes zero and the second atomic layer becomes the new top layer when N pseudoatoms are sputtered. This is illustrated in Fig. 1(c), showing concentrations of Ga, Al and As atoms in each atomic layer. This procedure is repeated for successive layers to simulate the sputter-etching process layer by layer.

Since the atomic concentration of j element, $c_j(i)$, in ith layer is easily calculated in the present approach, Auger intensities of Ga, As and Al are calculated according to the relation

$$I_j \propto \Sigma_{i=\text{surf}} c_j(i) \exp\left[-\frac{(i-0.5)d}{\lambda_j}\right], \tag{1}$$

where λ_j is the escape depth of an Auger electron emitted from the j element, d is the mean interatomic distance of a specimen. The present calculation has used the electron escape depths, 0.21 nm for Ga (MVV—55 eV), 1.06 nm for Ga (LMM—1063 eV), 0.40 nm for Al (LVV—64 eV) and 0.20 for As (MVV—32 eV) according to Ref. 12.

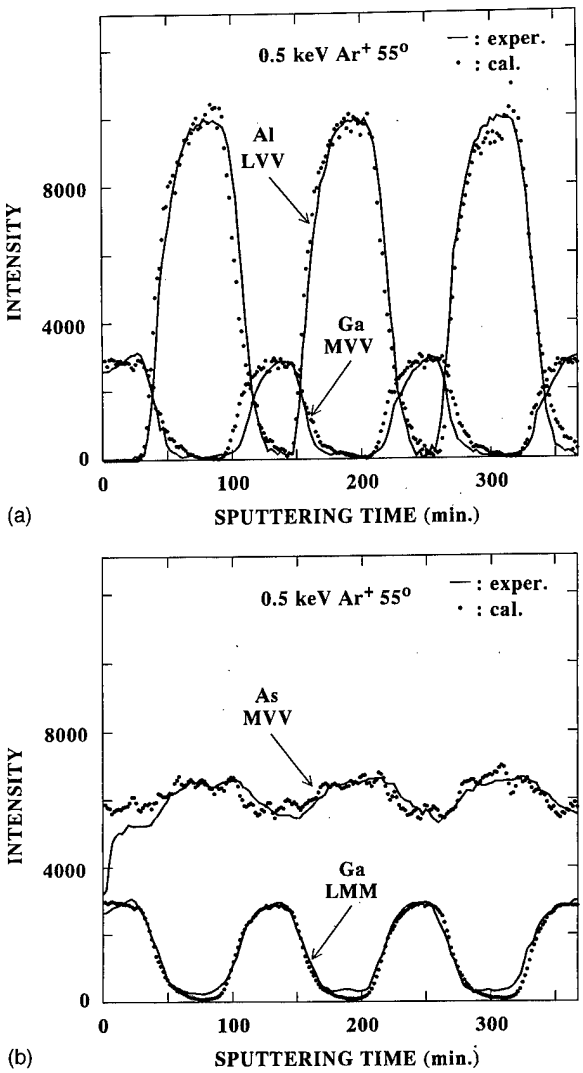


FIG. 2. Calculated (points) and experimental (lines) results of Auger depth profiles of (a) Al *LVV*, Ga *MVV* and (b) Ga *LMM*, As *MVV* for an AlAs (9.90 nm)/GaAs (8.77 nm) multilayered system, sputtered with 0.5 keV Ar^+ ions at incidence angle of 55° from the surface normal.

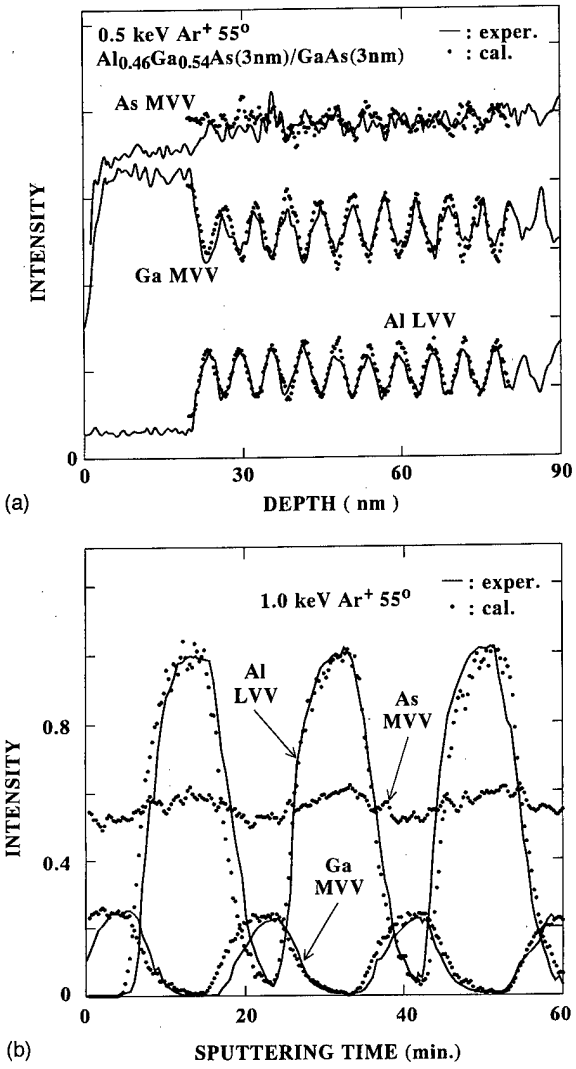


FIG. 3. Calculated (points) and experimental (lines) results (see Refs. 13 and 14) of Auger depth profiles of (a) $Al_{0.46}Ga_{0.54}As$ (3.1 nm)/GaAs (3.1 nm) multilayered system with 0.5 keV Ar^+ and (b) AlAs (9.90 nm)/GaAs (8.77 nm) with 1.0 keV Ar^+ at incidence angle, 55°.

III. RESULTS AND DISCUSSION

Figures 2(a) and 2(b) show the calculated results, compared with the experimental results[8] of the Auger depth profiles for (a) Ga-*MVV*, Al-*LVV* (b) Ga-*LMM* and As-*MVV* in the multilayered system of GaAs (8.77 nm,the topmost layer is 7.64 nm)/AlAs (9.90 nm) sputtered with 0.5 keV Ar^+ ions at an incidence angle of 55° to the surface normal. Even in the low energy sputtering (0.5 keV), each calculated profile agrees very well with the experimental result for all the elements without deterioration in describing the profiles with respect to sputtered depth. Figure 3 shows the depth profiles of thin layer superlattices, (a) $Al_{0.46}Ga_{0.54}As$ (3 nm)/GaAs (3 nm) sputtered with 0.5 keV Ar^+ and (b) AlAs (9.90 nm)/GaAs (8.77 nm) with 1.0 keV Ar^+. We find again very good agreement between the calculated and experimental[13,14] profiles.

In Fig. 2(b) the present simulation predicts that Auger intensity of As in the AlAs layer is about 1.2 times larger than that of GaAs layer, suggesting that the mixing zone becomes more enriched with As than with Al, by preferential sputtering of Al in the AlAs layer. The ratio of partial sputtering yields of Al to As for AlAs is also predicted to be 1.40 under 0.5 keV Ar^+ bombardment while that of Ga to As for GaAs is 1.08. To examine the deterioration of depth profile caused by preferential sputtering, we calculated low energy Auger depth profiling for two cases having almost the same mixing zone of ~10 monolayer for 0.5 keV Ar^+; (1) Al in the GaAs/AlAs multilayered system, (2) Al* in the GaAs/Al*As, in which pseudoatom Al* was assumed to have the same escape depth as Al atom but its mass assumed to be same as that of the Ga atom. In the latter, therefore, no effect of preferential sputtering is expected. The results, shown in Fig. 4, indicate that the Auger intensity of Al in the AlAs layer is always a little less than that of Al* in the Al*As layer. This leads to additional theoretical confirmation that the concentration of Al becomes less in the mixing zone

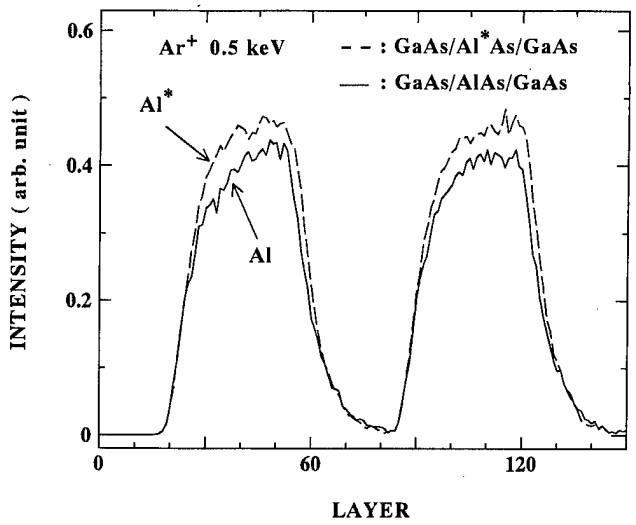


FIG. 4. Proposed simulation for GaAs/Al*As, in which pseudoatom Al* is assumed to have the same escape depth of Al atoms but its mass is assumed to be the same as that of Ga atom.

due to the preferential sputtering of Al in the AlAs layer.

The change of etching rate at the interface makes it difficult to convert the sputtering time to the depth in experimental results. However, the simulation can easily calculate the change of etching rate. Figure 5 shows variation of the etching rate in the AlAs and GaAs layer calculated for comparing the present simulation with the experimental data of Fig. 2. Al (64 eV) Auger depth profiling is depicted for reference to the relevant superlattice layers. The simulation predicts that etching rate in the AlAs layer is a little bit slower than that of the GaAs layer. This prediction is reasonable because the sputtering yield in the AlAs layer is reported to be lower than that of the GaAs layer.[15]

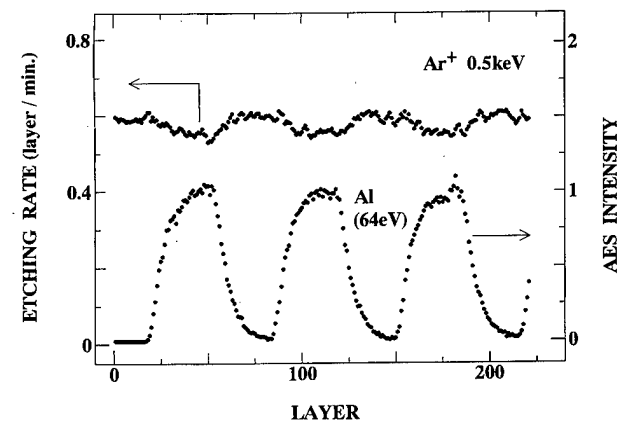


FIG. 5. Variation of etching rate in the AlAs and GaAs layer calculated in the present simulation for depth profiling of Fig. 4. Al Auger depth profiling was depicted for reference to the relevant superlattice layers.

IV. CONCLUSIONS

This article presents a dynamic simulation approach to the Monte Carlo calculation of atomic mixing caused by ion bombardment. In the present simulation we have emphasized the following features of our binary encounter model in the low energy region: (1) We have chosen the maximum impact parameter, P_{max} to be given by $P_{max}=d/2$. (2) We have paid special consideration to interstitial atoms and vacancies generated in the collisional cascade. The improvements based on the above considerations have led the present simulation to describe very well the Auger depth profiles of the AlAs/GaAs superlattice system, adding useful clarification to the basic mechanism of the atomic mixing phenomena.

We also note that the present simulation is essentially a Monte Carlo simulation based on the binary encounter model though we have incorporated some basic improvements to describe collision processes of low energy projectiles. This simulation code, therefore, can be used as easily as the conventional TRIM[16] and/or TRIMEX[17] codes, and leads to a more comprehensive understanding of depth profiling by surface analytical techniques. Another point to be noted is that the behavior of interstitials and vacancies generated in the collisional cascade is crucial in understanding atomic mixing phenomena. For this, incorporation of the recombination processes based on embedded-atom model to the present simulation code has been under examination.

ACKNOWLEDGMENTS

The authors are very grateful to Dr. K. Kajiwara of Sony Corp. Research Center for providing the recent experimental results of Auger depth profiles of AlAs/GaAs multilayered system. Thanks are also due to Dr. Wolfgang Eckstein for extending the TRYDYN calculation to the present subjects. Critical reviewing by Professor R. H. Ritchie during the stay in the laboratory as a JSPS guest professor is gratefully acknowledged.

[1]K. Wittmaack, Vacuum **34**, 119 (1984).
[2]B. V. King and I. S. R. Tsong, J. Vac. Sci. Technol. A **2**, 1443 (1984).
[3]P. Sigmund and N. Q. Lam, Mat. Fys. Medd. Dan. Vidensk. Selsk. **43**, 255 (1993).
[4]W. O. Hofer and U. Littmark, Phys. Lett. A **71**, 457 (1979).
[5]Moeller and W. Eckstein, Nucl. Instrum. Methods Phys. Res. B **2**, 814 (1984).
[6]Y. Yamamura, Nucl. Instrum. Methods Phys. Res. B **28**, 17 (1987).
[7]H. I. Lee, R. Shimizu, M. Inoue, K. Kajiwara, and S. Hofmann, Jpn. J. Appl. Phys., Part 1 **35**, 2271 (1996).
[8]K. Satori, Y. Haga, R. Minatoya, M. Aoki, and K. Kajiwara, J. Vac. Sci. Technol. A **15**, 478 (1997).
[9]H. J. Kang, E. Kawatoh, and R. Shimizu, Jpn. J. Appl. Phys., Part 1 **24**, 1409 (1985).
[10]K. Min and R. Shimizu, Surf. Sci. **341**, 241 (1995).
[11]J. Bourgoin and M. Lannoo, *Point Defects in Semiconductors* (Springer, Berlin, 1983).
[12]T. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **21**, 170 (1993).
[13]K. Kajiwara and H. Kawai, Surf. Interface Anal. **15**, 433 (1990).
[14]K. Kajiwara and R. Shimizu, J. Vac. Sci. Technol. A **13**, 1316 (1995).
[15]K. Kajiwara, Doctorate thesis, Faculty of Engineering, Osaka University, Japan, 1992.
[16]J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids* (Pergamon, New York, 1985).
[17]S. Han, Ph. D. thesis, Dept. of Nuclear Engineering and Engineering Physics, University of Wisconsin-Madison, 1984.

Proposals for exact-point transmission-electron microscopy using focused ion beam specimen-preparation technique

T. Ishitani,[a] Y. Taniguchi, S. Isakozawa, and H. Koike
Instrument Division, Hitachi Ltd., Hitachinaka 312-8504, Japan

T. Yaguchi, H. Matsumoto, and T. Kamino
Techno Research Laboratory, Hitachi Instruments Engineering Co., Ltd., Hitachinaka 312-8504, Japan

(Received 19 November 1997; accepted 23 February 1998)

A focused ion beam (FIB) has been actively applied for preparation of about 0.1-μm-thick specimens for transmission electron microscopes (TEMs). For device failure analyses, however, it is mostly difficult to prepare the exact-point TEM specimens. The reason is that the failures are mostly beneath the surface and their exact locations are unknown. Then, even step-by-step FIB cross sectioning may sputter away the failures in the TEM specimen preparation. In the present study, we review two proposals for exact-point TEM microscopy using FIB specimen-preparation technique: (1) high-voltage scanning electron microscopes (HV-SEMs) imaging in TEM and (2) energy filtering TEM (EF-TEM) imaging. The HV-SEM imaging provides information on not only the sample surface but also the inner structure up to about 1 μm deep. The EF-TEM imaging is applicable even for 0.5-μm-thick specimens at 100 kV in the accelerating voltage, in contrast with about 0.1-μm-thick specimens for conventional TEM imaging. Preliminary experiments have supported that either proposal presumably improves the yield of the exact-point TEM inspection.
© *1998 American Vacuum Society.* [S0734-211X(98)01504-2]

I. INTRODUCTION

Transmission electron microscopes (TEMs) and scanning electron microscopes (SEMs) have been used for failure analysis and process evaluation of semiconductor devices and advanced materials. A focused ion beam (FIB) is a powerful tool for cross-sectional specimen preparation of any specific area for TEM and SEM.[1–4] The TEM specimens are typically prepared as follows (see Fig. 1): First, a mesa-shaped sample of <50 μm in width is cut from a device chip or wafer using a dicing saw. Next, the sample is transferred into the FIB work chamber for milling. The milling area is edited and specified on a CRT displaying scanning ion microscope (SIM) images of the sample surface. Two trenches are milled, leaving a thin strip (approximately 0.1 μm in width), which is the section of interest for the TEM inspection.

For device failure analyses, the failures are mostly found from electrical measurements and their exact locations are unknown because they are beneath the surface. Although fine-step cross sectioning is employed to confine the failure in the final TEM specimen of about 0.1 μm in thickness, there is a possibility of overmilling the failure.

In the present study, we review two proposals for exact-point TEM observation using FIB specimen-preparation technique: (1) high-voltage SEM (HV-SEM) imaging (which is one of TEM's optional functions)[4,5] and (2) energy filtering TEM (EF-TEM) imaging (or electron spectroscopic imaging; ESI).[6,7] The HV (200 kV) SEM images provide information on not only sample surface but also inner structure up to about 1 μm deep. (Note that the information depth of SIM images is as shallow as a few nanometers.)[8] Such inner structure images are very useful not to overmill the failures during the fine-step cross sectioning. On the other hand, the EF-TEM imaging is applicable even for 0.5-μm-thick specimens at 100 kV in the accelerating voltage, in contrast with 0.1-μm-thick specimens for conventional TEM (CTEM) imaging. Thick specimens are desirable to avoid overmilling the failures. Either imaging presumably improves the yield of the exact-point TEM inspection. Here, the FIBs used in the present experiments are 30 kV gallium (Ga) ion beams.

II. HIGH-VOLTAGE SEM (HV-SEM) IMAGING IN TEM

A TEM has a HV-SEM imaging mode in addition to normal transmission imaging mode (e.g., Hitachi 200-kV TEM system: HF-2000).[4,5] When the primary electron beam strikes a sample, secondary electrons (SEs) as well as backscattered electrons are generated. These electrons travel upward through the objective lens bore by a combined force of magnetic field of the lens and electrostatic field of the SE detector (+10 kV) and are collected.

Figure 2 shows an electron-sample interaction when the sample is illuminated by a 200 kV electron beam. Devices such as 4M- and 16M-dynamic-random-access-memory (DRAM) used in the present study have transistors, and capacitors which are composed by Al and poly-Si surrounded by silicon oxide (SiO_2). When the 200 keV primary electron beam strikes Al or Si, both SEs and backscattered electrons are generated. The SEs generated deep in the sample are absorbed within the sample because their energies are so small. On the other hand, the backscattered electrons have high energies so that they can travel through silicon oxide of

[a]Electronic mail: ishitani@cm.naka.hitachi.co.jp

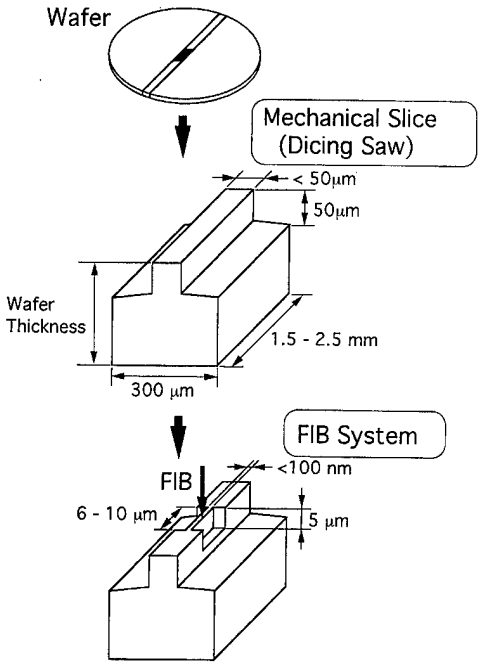


FIG. 1. Typical TEM specimen preparation using FIB cross sectioning.

1 μm thick and come out from the surface. These backscattered electrons can generate SEs as they come up to the surface. These SEs are collected to make SEM images. Thus, the HV-SEM imaging provides information on sample surfaces as well as inner structures of about 1 μm deep. (A few nm in information depth for the SIM imaging. The ion beam does not logically excite backscattered electrons. Besides, the backscattered ion yield is much smaller than the backscattered electron yield for electron bombardment.[8])

A new process flow proposed for the exact-point TEM observation on the failure analysis is shown in Fig. 3. Here, the failure is located under the device surface. First, the sample is cross sectioned so as to be a 1–2-μm-thick TEM specimen in the FIB chamber. Next, it is transferred into the TEM chamber for its HV-SEM imaging. Fine-step FIB cross sectioning is repeatedly carried out until the region of interest appears symmetrically in the HV-SEM images of the specimen's front and backsides. The TEM specimen of about 0.1 μm in thickness surely contains the region of interest. Although specimen transfer between the FIB and TEM systems is required several times, the transfer is easily done by using a FIB/TEM compatible specimen holder as shown subsequently.

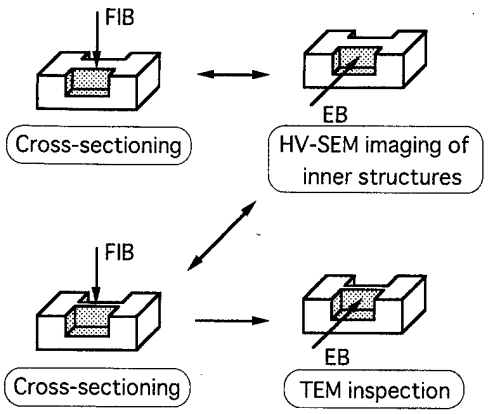


FIG. 3. Process flow of the exact-point TEM observation using HV SEM imaging and FIB cross sectioning.

Feasibility experiments[2] for the exact-point TEM specimen preparation with the help of HV-SEM imaging have been carried out using the DRAM devices. Figure 4 shows a series of HV-SEM images of the 16M-DRAM device, which have been taken at every 0.1–0.2-μm-thick FIB milling. Structures around the contact change in three dimensions (3D) as milling proceeds. Figure 5 shows a stereo pair of images of the 4M-DRAM specimen. The specimen was tilted by $\pm 8°$. The depth of field is about 1 μm. The image resolution gradually worsens with the depth and reaches about 0.1 μm at about 1 μm deep. It means that the 200 kV SEM image allows 3D microscopy with about 0.1 μm image resolution from the surface to about 1 μm deep for a Si device.

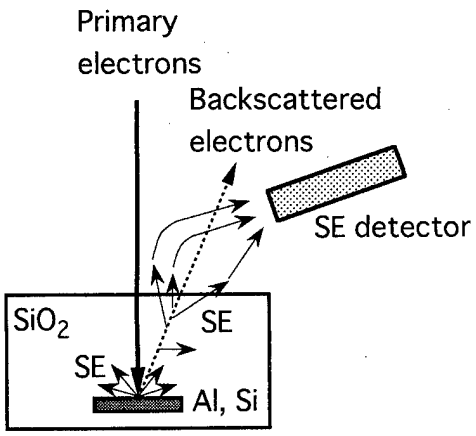


FIG. 2. Generation of secondary electrons (SE) and backscattered electrons using a high voltage (HV) electron beam.

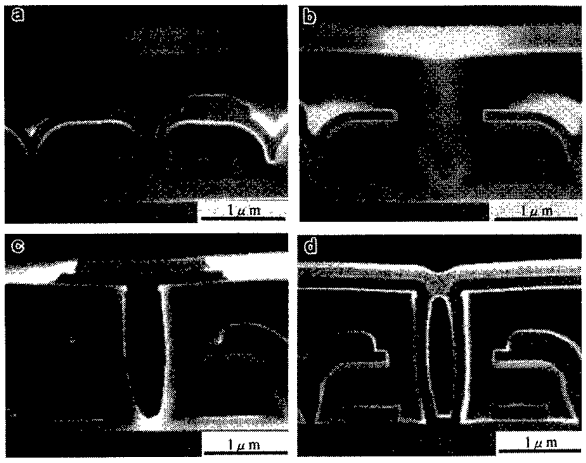


FIG. 4. Series of HV SEM images taken at every 0.1–0.2-μm-thick FIB-milling for 16M-DRAM specimen (Hitachi 200-kV TEM; HF-2000).

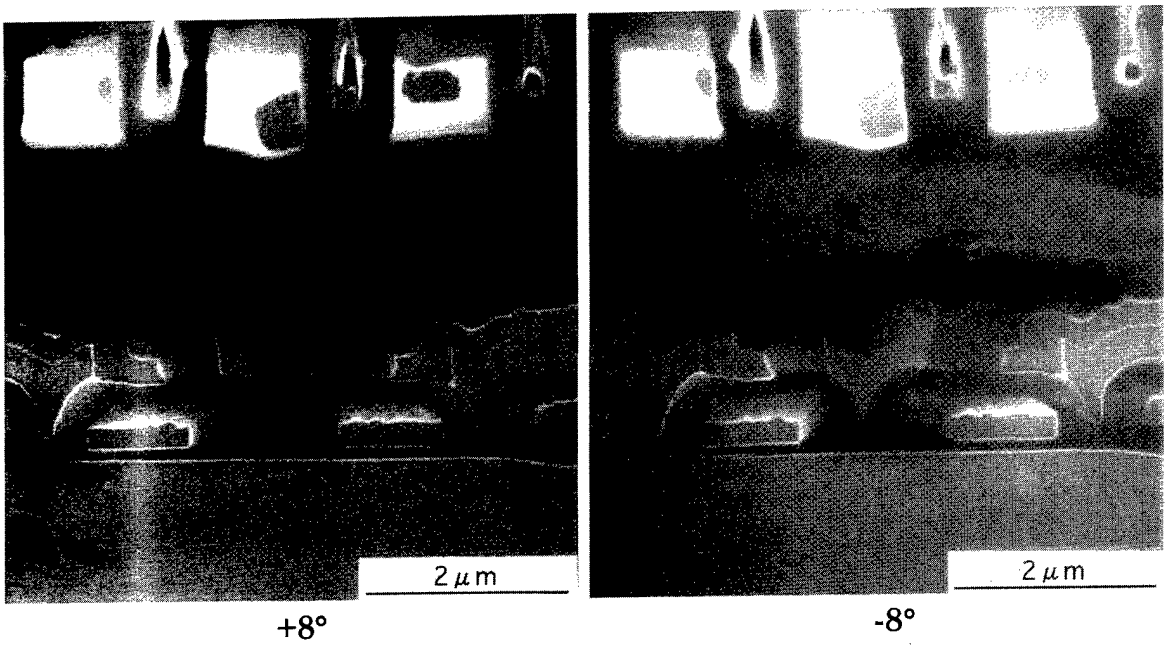


FIG. 5. Stereo pair of HV SEM images of 4M-DRAM specimen (Hitachi 200-kV TEM; HF-2000).

Figure 6 shows the Monte Carlo simulations of 30 and 200 keV electron's trajectories in a 1-μm-thick Si film. The 200 keV beam broadens with the penetrating depth and its diameter (defined as $2^{3/2}\sigma$) reaches about 0.1 μm at 1 μm deep. This beam broadening explains successfully the depth-dependent image resolution of the 200 kV SEM mentioned above. On the other hand, the 30 keV electron beam is about 0.1 μm broad even at a quarter micron deep. Besides, some electrons are backscattered and diffused. These processes also worsen the SEM image resolution. Therefore, a 30 kV SEM is insufficient to aid the submicron-step FIB cross sectioning with imaging the inner submicron-scale structures. We trust that the FIB cross sectioning aided by HV(200 kV)-SEM imaging will be a very powerful method on an exact-point TEM sample preparation for the failure analysis.

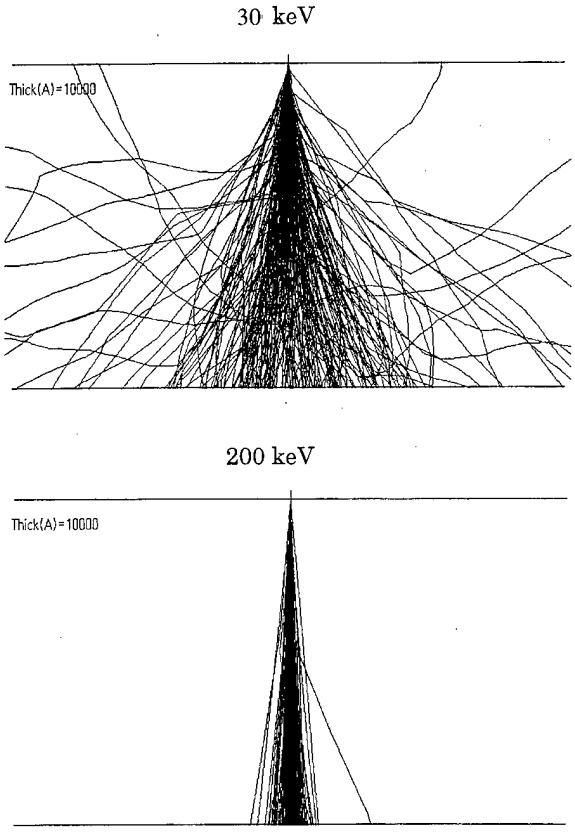


FIG. 6. Monte Carlo simulations of 30 and 200 keV electron's trajectories in a 1-μm-thick Si film. (The number of electrons is 200.)

III. ENERGY-FILTERING TEM (EF-TEM) IMAGING

Another approach to the exact-point TEM observation is the energy-filtering TEM (EF-TEM) imaging,[6,7] which is applicable for thick specimens of about 0.5 μm, in contrast to about 0.1 μm in the conventional TEM (CTEM) imaging. The electron beam transmitted through the specimen contains electrons of various energies as shown in Fig. 7. The beam includes zero-loss electrons which have no energy loss, plasmon-loss electrons which have lost some energy due to the excitation of plasma oscillations in the specimen, and core-loss electrons which have lost energy through inner shell excitation of the specimen atoms.

Since CTEM uses all these different energy-loss electrons, the images tend to be blurred due to chromatic aberrations. An energy filter, on the other hand, can be used to select electrons of specific energies. With zero-loss electrons, even a thick specimen can be observed with sharp contrast. Using specific energy-loss electrons, thin specimens can be observed with a high contrast that reflects the elemental density distribution. Furthermore, with core-loss electrons resulting from inner shell excitation of specific elements, el-

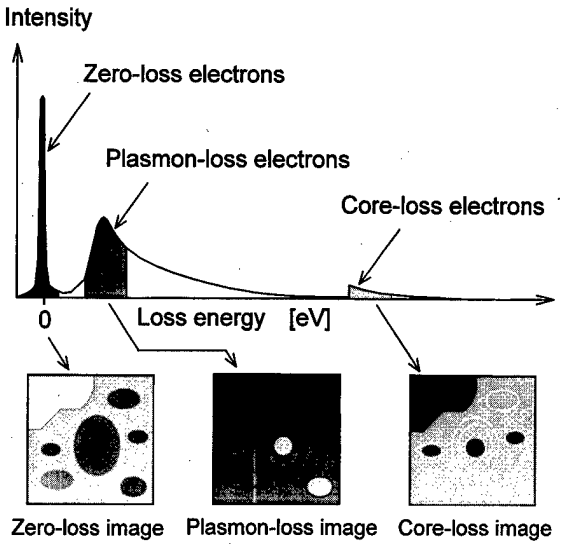


FIG. 7. Transmitted electrons containing zero-loss electrons, plasmon-loss electrons, and core-loss electrons.

emental mapping is feasible. These images reveal the elemental composition of ultramicroscopic portions of specimens, something not possible with CTEM.

An optical diagram of energy filter (Hitachi EF-1000)[4,5] used in the present study is schematically shown in Fig. 8. After being transmitted through the specimen, the electron beam passes through a magnetic objective lens and a two-stage intermediate lens, and then converges at a crossover to produce an image on the entrance image plane. The beam enters an energy filter of a sector magnetic pole, first where it is deflected 90°, then 180°, and finally 90°. The beam is

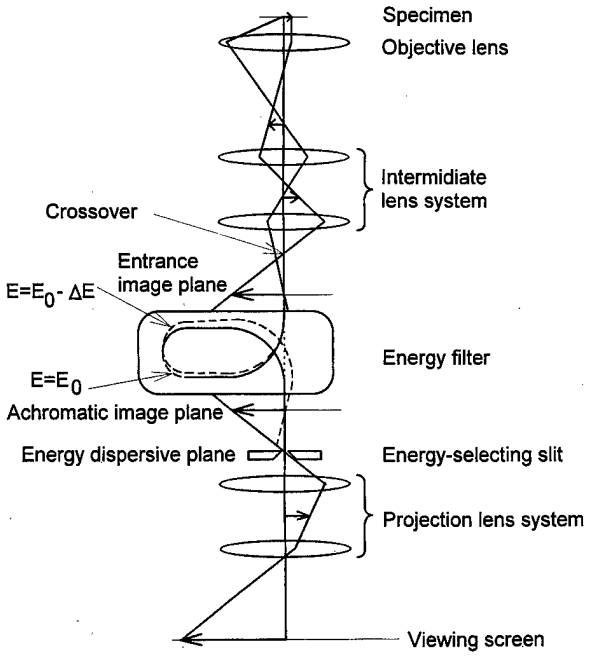


FIG. 8. Schematic diagram of energy filtering optics (Hitachi EF-1000).

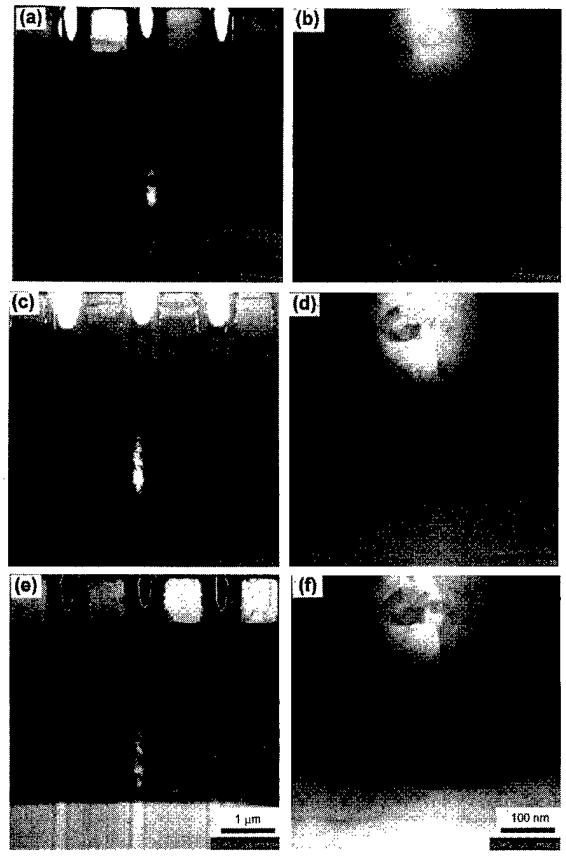


FIG. 9. 100 kV CTEM and EF-TEM images of a 16M-DRAM specimen (about 0.5 μm in thickness): (a) CTEM, (c) zero loss, and (e) plasmon loss (60 loss and 20 eV energy window); (b), (d), and (f) in the right row are high magnification images of the corresponding images in the left row.

energy dispersed and travels in γ-shaped trajectories. After leaving the energy filter, the electrons travel in different trajectories depending on the energy and converge again on the achromatic image plane. The electrons are confined in specific energies using an energy selection slit, and the resulting image shows little blurring. Finally, the achromatic image is enlarged on a viewing screen by two-stage projector lenses. Since the input beam axis for the energy filter is identical with the output, the CTEM mode is available by simply turning off the energy filter. A switchover from EF-TEM to CTEM mode and vice versa allows two-way TEM observation of the same specimen.

Typical comparisons between CTEM and EF-TEM images[4] are demonstrated using a 16M-DRAM specimen in Fig. 9. The thickness of the FIB prepared specimen is about 0.5 μm and the electron accelerating voltage of the TEM is 100 kV. Figure 9(a) shows a CTEM image of the whole area of 1-bit DRAM cell, and Figs. 9(c) and 9(e) show EF-TEM images of zero loss and plasmon loss (60 eV loss and 20 eV energy width) images, respectively. Figures 9(b), 9(d), and 9(f) are high magnification images of a contact hole in the corresponding Figs. 9(a), 9(c), and 9(e) images, respectively. The CTEM images do not provide fine structures of the specimen because the specimen is too thick for the 100 keV electrons and inelastic electrons blur the images.

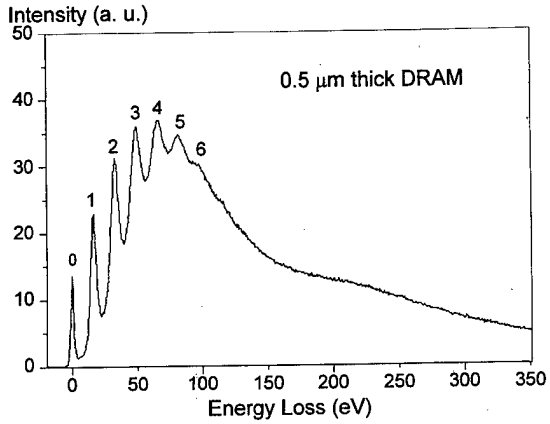


FIG. 10. Electron energy loss spectrum of the DRAM cell. The energy range of plasmon-loss images of Figs. 8(e) and 8(f) is set to the maximum of spectrum, i.e., the fourth plasmon-loss electrons.

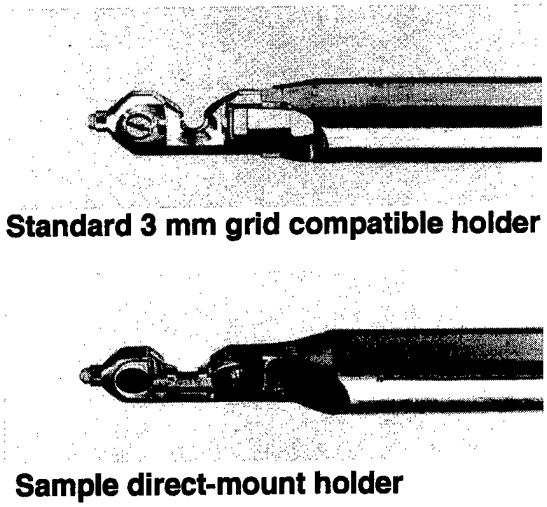


FIG. 11. FIB/TEM compatible specimen holders; (a) a standard 3 mm grid compatible holder and (b) a direct-mount specimen holder.

The zero-loss images remarkably enhance the image contrast and allow us to observe the fine structures. Here, the present specimen was still in the fabrication process so that the passivation film did not cover completely the wiring area and leaves some voids therein. Since these voids and their surrounding areas are largely different in electron densities, it is difficult to suppress a halation in the image. Last, the plasmon-loss images show removal of the halo pattern to allow us the entire area with the same contrast. The 60 eV loss images of the contact hole [Fig. 9(f)] show almost the same contrast and image quality as the zero-loss image [Fig. 9(d)]. This indicates that plasmon-loss images can reduce the specimen-thickness effect among various-thickness specimens. Bright rings at the top of Fig. 9(e) are formed by surface plasmon excitation of their corresponding openings in the specimen. Figure 10 shows an energy loss spectrum of the DRAM specimen used in Fig. 9. Comparing with a small zero-loss peak (0), plasmon-loss electrons have many peaks (1–6) counting up the sixth excitation. In the present plasmon-loss images [Figs. 9(e) and 9(f)], the energy range was set at 55–75 eV (20 eV in energy width) so as to involve the fourth plasmon-loss electrons which is the maximum of the spectrum. Since, in general, the suitable energy range for plasmon-loss image observation changes with the specimen thickness and the specimen itself, it should be defined in the operation.

Consequently, the EF-TEM provides plasmon-loss images with high contrast and high resolution even for about 0.5-μm-thick specimens. Such a thick specimen allows observation of whole structures that may be lost in CTEM specimen preparation (about 0.1 μm in thickness).

IV. LINKAGE OF FIB WITH TEM

The FIB cross sectioning is closely associated with the TEM observation as mentioned above. The specimen is transferred several times between the FIB and TEM systems, especially for a combination of the FIB milling and the HV-SEM imaging. To facilitate specimen handling between the two systems, compatible side-entry holders have been developed (e.g., Hitachi FIB system FB-2000A).[4,9] Figures 11(a) and 11(b) show two kinds of holders: a standard 3 mm grid compatible holder and a direct-mount specimen holder. The upper-side rim of the holder has been cut out for the ion beam. A double-tilt holder and an eight-specimen holder (not shown here) have also been produced. The eight-specimen holder comes with a turret on which the specimen grids are mounted. Each grid can be FIB-cross sectioned or TEM examined sequentially. The use of compatible holders provides the following three advantages:

(1) Time savings during specimen transfer between the FIB and TEM,
(2) easy to perform the specimen transfer between FIB and TEM systems,
(3) prevent specimen damage, contamination, cracking, or dropping during specimen handling.

The FIB system, mounted with a high-performance ion column designed using a new guideline,[10] allows the FIBs at high current densities in a wide range of beam diameters. Thus, the system has brought about high performance in both milling speed and positional accuracy. Consequently, the whole process from FIB milling to TEM observation is performed at a high throughput of 2–3 specimens/day, in comparison to 2–3 days/specimen for conventional specimen preparation techniques using mechanical polishing and broad-ion-beam thinning. Furthermore, the thickness allowance of about 0.5 μm for the EF-TEM imaging is rather effective to improve the yield of the exact-point TEM observation as well as shorten the FIB milling time.

V. CONCLUSIONS

Although FIB cross sectioning is very useful to prepare the TEM specimens of about 0.1 μm in thickness, there are some difficulties on exact-point TEM observations for device failure analysis. The reason is that the failures are mostly under the device surfaces and their exact locations are not determined from the SIM images providing the device sur-

face structures. Even fine-step FIB cross sectioning may overmill the failures. The following two proposals of the exact-point TEM observations combined with FIB cross sectioning are discussed: (1) HV-SEM imaging and (2) EF-TEM imaging (one of optional TEM functions). The HV-SEM imaging provides information on not only the sample surface but also inner structure up to about 1 μm deep. Such inner structure images are very useful to exactly confine the failures in about 0.1-μm-thick TEM specimens. The EF-TEM imaging is applicable even for 0.5-μm-thick specimens at 100 kV in the accelerating voltage, in contrast with 0.1-μm-thick specimens for the CTEM imaging. Thicker specimens are more desirable because they certainly retain the failures. Preliminary experiments have supported that either proposal presumably improves a yield of the exact-point TEM inspection. The specimen transfer between FIB and TEM systems is effectively carried out by employing the FIB/TEM compatible specimen holders. The whole process from FIB milling to exact-point TEM observation is able to attain high throughput of 2–3 specimens/day.

[1]K. Nikawa, J. Vac. Sci. Technol. B **9**, 2566 (1991).
[2]F. A. Stevie, T. C. Shane, P. M. Kahora, R. Hull, D. Bahnck, V. C. Kannan, and E. David, Surf. Interface Anal. **23**, 61 (1995).
[3]T. Ishitani and T. Yaguchi, Microsc. Res. Tech. **35**, 320 (1996).
[4]T. Ishitani, H. Tsuboi, T. Yaguchi, and T. Kamino, Scanning Microsc. (in press).
[5]Hitachi Scientific Instrument Technical Data, TEM No. 93 (1996).
[6]Y. Taniguchi, M. Arai, S. Taya, S. Isakozawa, and K. Asayama, Scanning Microsc. (in press).
[7]S. Taya, Y. Taniguchi, E. Nakazawa, and J. Usukura, J. Electron. Microsc. **45**, 307 (1996).
[8]T. Ishitani and H. Tsuboi, Scanning **19**, 489 (1997).
[9]T. Ishitani, T. Yaguchi, and H. Koike, Hitachi Rev. **45**, 19 (1996).
[10]T. Ishitani and Y. Kawanami, J. Vac. Sci. Technol. B **13**, 371 (1995).

GaAs microcrystal growth on semiconductor surfaces by low energy focused ion beam

Toyohiro Chikyow[a] and Nobuyuki Koguchi
National Research Institute for Metals, 1-2-1 Sengen, Tsukuba Ibaraki 305, Japan

(Received 30 November 1997; accepted 14 April 1998)

A position controlled or density controlled GaAs microcrystal growth from Ga droplets was attempted using a low energy focused ion beam system. Ga ions from a liquid Ga ion source were accelerated up to 10 kV to produce a focused ion beam. Subsequently the ions were given a positive bias to reduce their kinetic energy by a four-element retarding lens system. The Ga ions with a 30 eV kinetic energy softly landed on a S-terminated GaAs surface and formed a series of Ga droplets or Ga droplets matrix. After initiating the As molecule supply, GaAs microcrystals were found to grow from these Ga droplets. The low energy focused ion beam was also applied to create nucleation sites of Ga droplets on an As-terminated Si (001) surface at 100 eV. Ga droplets with high density were formed on a region where As atoms were removed partially on the Si surface. From these results, the low energy focused ion beam was found useful for position control or density control of GaAs microcrystals on semiconductor materials. © *1998 American Vacuum Society.* [S0734-211X(98)01404-8]

I. INTRODUCTION

Recently, an innovative method for fine structure fabrications has been introduced from the field of microelectronics.[1] An especially low dimension carrier confinement has been attempted for practical device application as well as fundamental interests.[2] For example, the quantum dot has been intensively studied because of its potential for laser application.[3] As a sophisticated fabrication method, "droplet epitaxy" was proposed, where GaAs microcrystals grow from Ga droplets on a Se- or S-terminated GaAs surface, and successful results have been reported.[4,5] However, the position of the GaAs microcrystals could not be controlled. To improve the optical properties, a position control for the GaAs microcrystals is required. For this purpose, first position control of Ga droplets must be achieved.

In the past, several types of focused ion beams with a retarding system have been proposed for etching and deposition with kinetic energy control.[6–9] However, a serious problem in the low energy focused ion beam system (LE-FIB) has been the problem in making fine structures due to its expanded beam diameter during retarding.

In this paper, a LE-FIB with a combination of the "droplet epitaxy" is proposed to fabricate position controlled fine structures and a direct GaAs microcrystal formation on the S-terminated GaAs surface is demonstrated. In addition, GaAs microcrystal growth on an As-terminated Si surface is attempted to investigate a new possibility of fine structure fabrication.

II. LOW ENERGY FOCUSED ION BEAM SYSTEM (LE-FIB)

The proposed LE-FIB system with a liquid Ga ion source has four groups of static lenses. The first one is used to extract Ga ions from the source and inject them into the accelerating column at 7.0–8.0 kV. The second lens is biased up to 10.0 kV for making a fine beam. The third lens is biased for focusing around 4.0 kV. The last one has four static electrodes for retarding using a 9.90–9.97 kV positive bias. The isotropic retarding in the column suppresses the ion beam expansion to some extent. Deflectors in the middle of the column scan the beam, and secondary electrons caused by ion bombardment are detected by a multi-channel plate attached at the bottom of the ion gun column, to obtain images as in scanning electron microscopy (SEM). This LE-FIB system is schematically illustrated in Fig. 1. The retarding part is set in the ultrahigh vacuum (UHV) chamber as shown in Fig. 1. The sample holder is shared with a scanning tunneling microscope (STM), which is situated in another UHV chamber, and with a conventional molecular beam epitaxy (MBE) system to supply Ga and As molecules. The three systems (LE-FIB, STM, and MBE) are connected in UHV. The sample holder is transferred between the three systems without breaking vacuum. Sample heating is carried out by direct current.

III. EXPERIMENTS

A. GaAs on a S-terminated GaAs (001) surface

A GaAs buffer layer was grown on a GaAs(001) substrate in another MBE system to grow a uniform buffer layer. The surface showed a (2×4) reconstruction after the growth. Subsequently sulfur molecules were supplied onto the surface at 400 °C to form a S-terminated (2×6) reconstructed surface.[10] The surface was then covered by As at room temperature to form a protection layer for carrying the sample from the MBE chamber to the LE-FIB system. In the UHV chamber of the LE-FIB system, the As protection layer was

[a]Electronic mail: tchikyo@momokusa.nrim.go.jp

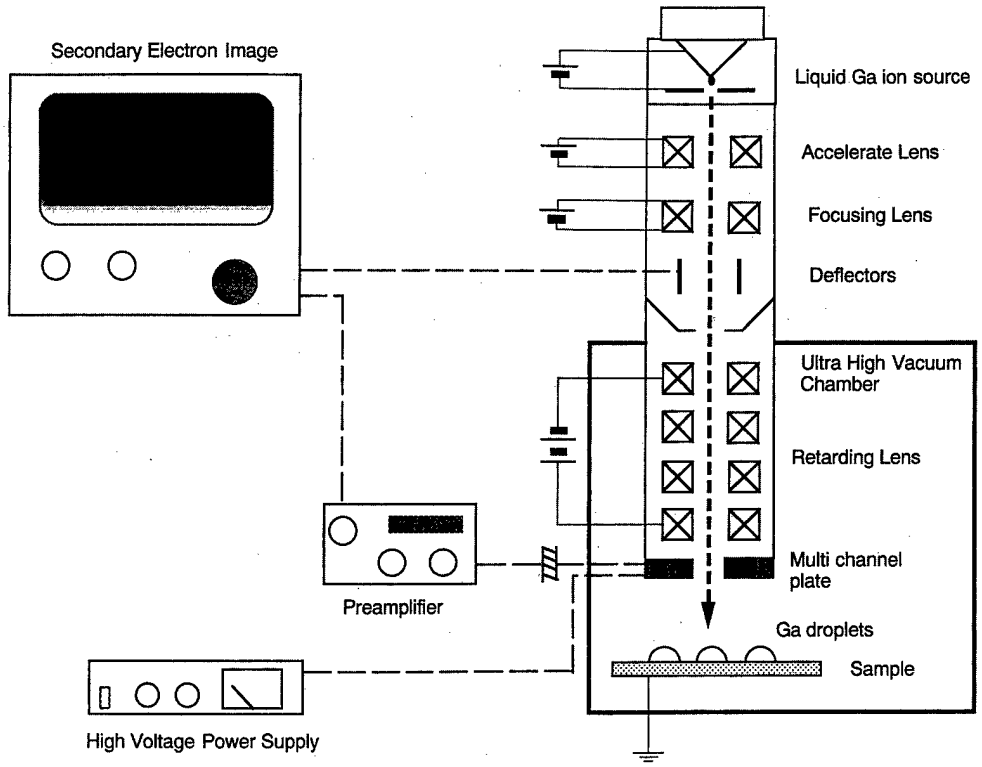


FIG. 1. Schematic illustration of the proposed low energy focused ion beam system.

removed thermally by direct current. The surface structure was monitored by reflection of high energy electron diffraction (RHEED).

Ga ions were supplied to the surface at 30–100 eV at room temperature in a repeated line scan mode. The ion beam current was 1.0 nA. The scanning speed was 25 cm/s and the total scanning time per line was 7.0 min. Ga ions were also supplied to the surface to form a 10×10 matrix of Ga droplets at 30 eV for 100 s (1.0 s Ga supply at each point). Surface morphologies were observed by SEM and the surface states were characterized by scanning tunneling spectroscopy (STS).

B. GaAs on a As-terminated Si (001) surface

After a conventional cleaning process, a rectangle shaped Si (001) sample was dipped into a HF solution ($HF:H_2O = 1:9$) and loaded into the MBE system. The sample was heated up to 1200 °C and flashed for 3 s by direct current in an As molecule atmosphere (As pressure was approximately 10^{-4} Pa). Subsequently the temperature of the sample was reduced to 400 °C to achieve an As termination of the Si surface.

After the partial pressure of As molecules became less than 2×10^{-6} Pa, the sample was transferred to the LE-FIB chamber and the sample surface was irradiated by Ga ions with 100 eV at room temperature to remove some As dimers from the surface. The scanned region was 200 μm$\times$200 μm. On the surface, Ga molecules were supplied at 500 °C using an effusion cell. Then As molecules were supplied again at the same temperature to grow GaAs microcrystals. The surface morphologies were observed by SEM.

IV. RESULTS AND DISCUSSION

When the Ga ions were supplied to the S-terminated GaAs surface below 100 eV, a series of Ga droplets 200–300 nm in size were observed. Lines of Ga droplets formed at 100 eV were shown in Fig. 2. The ion beam diameter was about 5.0 μm. The droplets seemed to locate in the center of the scanned ion beam. The space between the Ga droplets was 2.3 μm on average.

When Ga ions were supplied at 30 eV, Ga droplets in line were also observed in the beam center region. However, the space between the droplets was about 6.3 μm, which was wider than that at 100 eV. The droplets are shown in Fig. 3. In this case, the beam diameter was 10 μm because of the ion repulsion in the column.

The Ga ion beam has an energy distribution when it is extracted from the source. This distribution remains after retarding. The value is estimated to be ±10 eV for the liquid Ga ion source.[11] If the higher energy Ga ions are focused to the beam center to hit the S-terminated GaAs surface, defects such as surface vacancies or sulfur removed from the area will be formed in the center by ion bombardment. These defects will be nucleation sites for Ga to form droplets. Other Ga ions, which have a lower kinetic energy, land on the surface in an over focused condition and lose their charge immediately. These Ga atoms migrate on the surface after

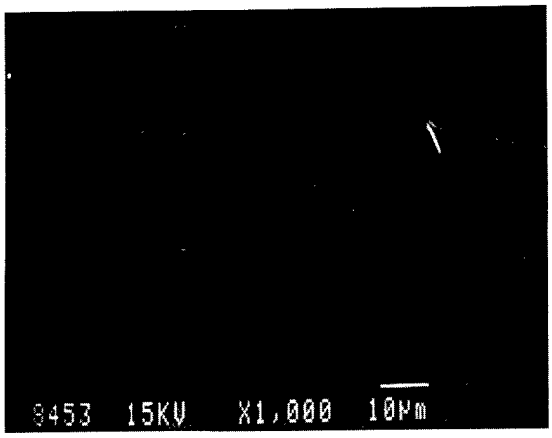


FIG. 2. Ga droplets in line formed by 100 eV ion irradiation. Ga droplets 200–300 nm in size were observed. The space between droplets is 2.3 μm on average.

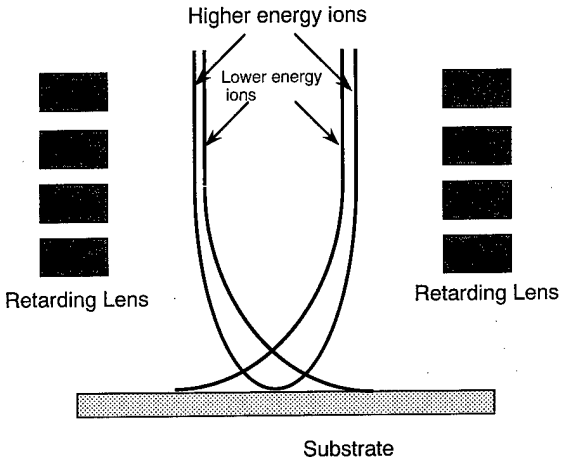


FIG. 4. Speculated trajectory of the higher energy Ga ions and the lower energy Ga ions. The higher energy ions are focused to the center region and create nucleation sites. The lower energy ions land on the surrounding region to supply Ga atoms.

landing. They may be trapped by the defects and subsequently form Ga droplets. As a result, a series of Ga droplets are thought to be formed in the center of the scanned ion beam. Some of the landing Ga atoms may contribute to make an intermixing region. Observed white lines in Fig. 3 are speculated to be regions where the intermixing has occurred. However the intermixing region is not likely to affect the Ga surface diffusion seriously because Ga droplets are formed by surface diffusion on micron scale.

When the Ga ions are supplied at the lower kinetic energy of 30 eV, the number of surface defects is reduced, leading to fewer nucleation sites. For this reason, fewer Ga droplets are formed, as observed in Fig. 3. This model is schematically illustrated in Figs. 4 and 5.

Defect density at the surface may affect the Ga diffusivity in the scanning direction. If the scanned region has greater defect density, the diffusion length of Ga becomes shorter. In this case, spacing of the Ga droplets becomes narrow. This may be the reason why the Ga droplets were arranged every 2.3 μm when the ions were supplied at 100 eV. On the other hand, a lower density of defects, caused by the 30 eV ion supply, leads to a longer diffusion length, resulting in the relatively larger separation between Ga droplets as observed in Figs. 2 and 3.

An advantage of this method is that fine structures can be formed even though the beam diameter is greater than 5.0 μm. Basically, the droplet size is thought to be controlled by the amount of Ga ions supplied. The LE-FIB combined with "droplet epitaxy" seems to be useful for fabricating fine structures.

To estimate surface damage caused by the ion bombardment, a STS measurement was carried out. The initial S-terminated GaAs surface was found to have few surface states because significant band bending was not observed. However, when the Ga ions were supplied at 30 eV to cover 20% of the surface with Ga atoms, the Fermi level began to move to the center of the GaAs band gap. This means that some of the Ga ions made surface states, resulting in Fermi level pinning. Only a 1% surface defect density for surface sites is sufficient for the Fermi level to be pinned. From these

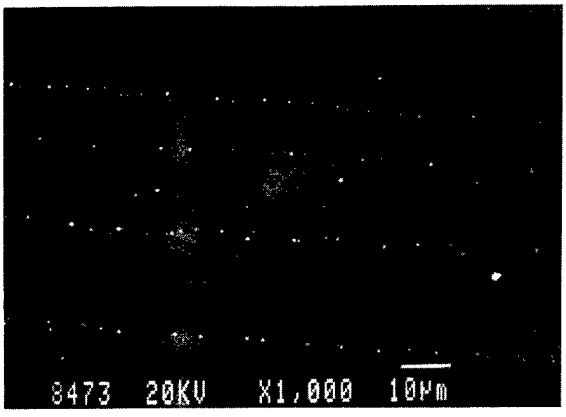


FIG. 3. Ga droplets in line formed by 30 eV ion irradiation, Ga droplets 200–300 nm in size were also observed. However the space between the droplets is 6.3 μm on average.

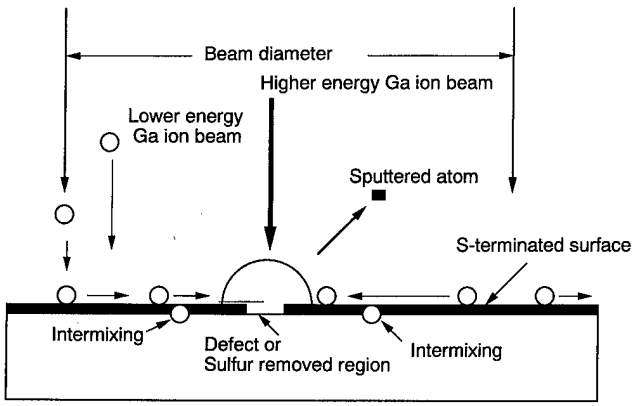


FIG. 5. Schematic illustration of defect formation by ion bombardment and growth of Ga droplets by surface diffusion.

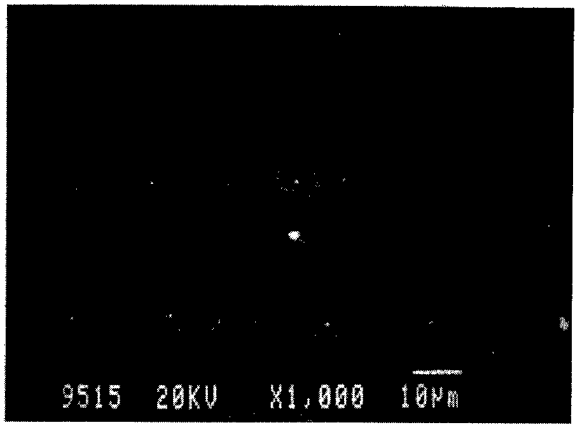


FIG. 6. GaAs microcrystals in line after As molecule supply. GaAs microcrystals of 500 nm were observed after the As molecule supply to the Ga droplets. The white spots are the GaAs microcrystals. The GaAs microcrystals were polycrystalline.

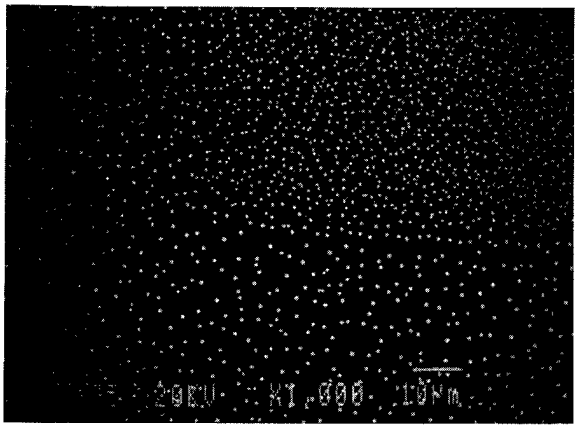


FIG. 7. Ga droplets on the As-terminated Si(001) surface. A smaller size of Ga droplets with high density were observed in a region where Ga ion beams with 100 eV were irradiated (right upper region). On the remaining As-terminated region, larger size of Ga droplets with lower density was observed.

results, most of the Ga ions were thought to land softly on the surface and cause little damage. These results agree with the proposed model that higher energy ions make surface defects for nucleation and lower energy ions land softly on the surface to grow Ga droplets.

After the As molecule supply to the Ga droplets at 200 °C, GaAs microcrystals in line were observed as in Fig. 6. A three-dimensional growth of GaAs occurred and this result meant that a lateral growth of GaAs, which was observed on an As-stabilized GaAs surface,[4] did not occur on the surface, although some parts of the surface had an intermixing region between the irradiated Ga- and S-terminated GaAs.

The microcrystals were polycrystalline due to the low growth temperature. However, if the size of the Ga droplets was smaller than 50 nm, the microcrystals should be single crystal as already confirmed in previous reports.[4,5]

When the Ga ion with 30 eV was supplied to the S-terminated GaAs surface to form a Ga droplet matrix, Ga droplets of 100 nm in size were observed every 10 μm. The droplets were located in the center of the irradiated area. This result means that the higher energy ions make surface defects or the S-removed region in the center of the irradiated area and lower energy ions land softly on the surrounding region. In the same manner with the Ga droplets formation by repeated line scan mode, Ga atoms were trapped in the center region and formed Ga droplets.

The As-terminated Si (001) surface has As–As dimers on a Si surface and has a fully coordinated surface with (2 $\times$1) symmetry, where all the dangling bonds are occupied. This surface structure creates a stable surface with a low surface free energy.[12] Actually the surface free energy of the As-terminated Si (001) surface is estimated to be 0.38 eV/atom, while the Si (2$\times$1) surface is estimated to be 1.45 eV/atom.[13] A Ga surface has almost the same value of the surface free energy as the Si (001) surface. From these facts, a partially As-removed Si surface is expected to give preferential nucleation sites for semi-spherical Ga droplet formation. As the As-terminated Si surface has fully occupied bonds, additional As molecules do not adhere to the surface. This will create another surface which is similar to the S-terminated GaAs surface.

As shown in Fig. 7, in the region where Ga ions were scanned, Ga droplets of 300 nm in size with a high density were apparently observed, while on the As-terminated region, larger Ga droplets of 500 nm were found in low density.

After the As molecule supply, a three-dimensional growth was observed as shown in Fig. 8. If the proposed As dimers were removed periodically by finer Ga ion beams, position control or arranged GaAs microcrystals would be expected, leading to a new method of a position control of microcrystals.

In summary, GaAs microcrystal growth in combination with low energy focused ion beams and the "droplet epitaxy" was attempted on the S-terminated GaAs (001) surface

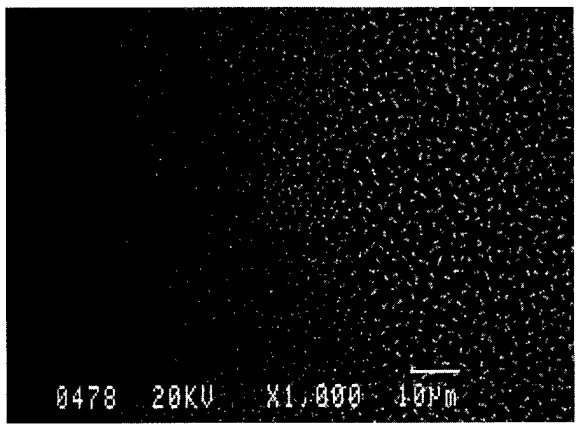


FIG. 8. GaAs microcrystals on the As-terminated Si(001). Smaller and high density GaAs microcrystals were observed in the region where Ga ions were irradiated (left region). On the remaining region, larger and lower density GaAs microcrystals were observed.

and As-terminated Si (001) surface. On both surfaces, the low energy focused ion beam played an important role in position control of microcrystals and nucleation site creation for the microcrystal growth. From the results obtained, a possibility of fine structure fabrication by low energy focused ion beam was demonstrated.

ACKNOWLEDGMENT

The authors are grateful to Dr. Tsukamoto for his fruitful discussions on the S-terminated GaAs surface.

[1]See for example, L. M. Terman, Appl. Surf. Sci. **117**, 1 (1977).
[2]H. Sakaki, Surf. Sci. **267**, 623 (1992).
[3]Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).
[4]T. Chikyow and N. Koguchi, Appl. Phys. Lett. **61**, 2431 (1992).
[5]N. Koguchi and K. Ishige, Jpn. J. Appl. Phys., Part 1 **32**, 2052 (1993).
[6]R. Aihara, H. Kasahara, and H. Sawaraki, J. Vac. Sci. Technol. B **7**, 79 (1989).
[7]K. Gamo, in *Materials Research Society Proceedings: Beam Solid Interactions: Fundamentals and Applications*, edited by M. Nastasi, L. R. Herbots, and R. S. Averback (Materials Research Society, Pittsburgh, PA, 1993), p. 577.
[8]S. Nagamachi, Y. Yamakage, H. Maruno, M. Ueda, S. Sugimoto, and M. Asari, Appl. Phys. Lett. **62**, 2143 (1993).
[9]K. Pak, I. Saitoh, N. Ohshima, and H. Yonezu, J. Cryst. Growth **140**, 244 (1994).
[10]S. Tsukamoto and N. Koguchi, Jpn. J. Appl. Phys., Part 2 **33**, L1185 (1994).
[11]L. W. Swanson, G. A. Schwind, and A. E. Bell, J. Appl. Phys. **51**, 3453 (1980).
[12]R. D. Bringans and M. A. Olmstead, J. Vac. Sci. Technol. B **7**, 1232 (1989).
[13]D. K. Biegelsen, R. D. Bringans, J. E. Northrup, M. C. Schabel, and L. E. Swartz, Phys. Rev. B **47**, 9589 (1993).

AlGaAs/GaAs heterostructures grown on a focused-Be-ion-beam written backgate

Y. Hirayama[a] and T. Saku
NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa 243-01, Japan

(Received 17 November 1997; accepted 15 April 1998)

Novel structures, in which an AlGaAs/GaAs modulation doped structure is overgrown on an underlying Be-implanted *p*-type region, are successfully fabricated using a system in which focused-ion-beam (FIB) implantation and molecular-beam epitaxy chambers are connected through a high vacuum tunnel. The two-dimensional electron gas (2DEG) at the heterointerface is well controlled by a voltage applied to the Be-FIB written backgate. Though Be out diffusion into the overgrown layer is observed, the sharp front of the out diffusion enables us to fabricate devices with a small separation between the 2DEG and *p*-type backgate. The three-dimensional hole gas (3DHG) formed by the Be-FIB implantation is used not only as a backgate but also for measuring the interaction between 2DEG and 3DHG. © *1998 American Vacuum Society.*
[S0734-211X(98)04304-2]

I. INTRODUCTION

A combination of focused-ion-beam (FIB) selective doping and molecular-beam epitaxy (MBE) overgrowth has been used to fabricate several kinds of three-dimensional structures.[1–5] Among them, localized buried backgates formed by FIB implantation are widely applied to control the double layer structures.[4–7] Linfield *et al.*[4] and the authors[7] formed a localized backgate using a Ga-FIB insulating technology. The Si-doped *n*-type layer is cut into several segments by Ga-FIB scanning,[8] and it is possible to grow a modulation-doped structure on it. In these structures both the Si-doped backgate and the two-dimensional electron gas (2DEG) at the heterointerface are *n*-type, which means an AlGaAs barrier layer is necessary in order to reduce the gate leakage current. In general, it is more difficult to grow a modulation-doped AlGaAs/GaAs structure on AlGaAs than on GaAs.

In this article we demonstrate a different type of backgate structure where the back-gate is formed from a Be-FIB implanted *p*-type region. The *pn*-junction-type structures enable us to grow a AlGaAs/GaAs modulation-doped structure directly on the Be-implanted GaAs without inserting an AlGaAs barrier layer. The electron density of the 2DEG is well controlled by applying a backgate bias less than the band gap energy value where a forward current starts to flow through the *pn* junction.

The Be-FIB written backgate (three-dimensional hole gas; 3DHG) and the 2DEG have independent Ohmic contacts, so it is also possible to measure the drag effect[9–11] between 2DEG and 3DHG using these structures.[12]

II. FABRICATION

A schematic diagram of the fabricated structure is shown in Fig. 1. First, an undoped GaAs layer was grown on a semi-insulating GaAs substrate by MBE. Then, the growth was interrupted and the sample was transferred from the MBE chamber to the FIB chamber. All systems including the tunnel between the two chambers were kept in a vacuum better than 10^{-9} Torr. The Be ion beam extracted from the Au–Si–Be liquid-metal ion source was implanted in the designed place at 50 kV. The ion dose was changed from 10^{13} to 6×10^{14} cm^{-2} by adjusting the number of scans. To ensure the good contact between Ohmic metals and Be-FIB written *p*-type regions, the Be ion dose was increased to more than 1.5×10^{14} cm^{-2} in the contact regions. After FIB scanning, the sample was returned to the MBE chamber and the modulation-doped structure was overgrown. In the initial stage of the second growth, the sample was heated to 670 °C. Then, the overgrowth was carried out at the growth temperature of 660 °C. The modulation-doped structure consists of an undoped 250 nm thick GaAs layer, an undoped 15 nm $Al_{0.33}Ga_{0.67}As$ spacer layer, a Si-doped (2×10^{18} cm^{-3}) 25 nm $Al_{0.33}Ga_{0.67}As$ layer, an undoped 15 nm $Al_{0.33}Ga_{0.67}As$ layer and a Si-doped (10^{18} cm^{-3}) 15 nm GaAs cap layer. The undoped $Al_{0.33}$ $Ga_{0.67}$ As spacer layer was 25 nm in the structure used for the secondary ion mass spectroscopy (SIMS) profile measurements. The implanted Be was activated during the second growth process and no additional annealing was carried out after the MBE growth.

After completion of the growth, a Hall-bar pattern including five terminals was defined by chemical etching. The Au–Ge–Ni ohmic metals were formed at the ends of the five terminals for *n*-type Ohmic contacts. On the other hand, Au/Au–Zn metals were evaporated at the ends of the Be-implanted region as *p*-type Ohmic contacts. A Ti/Au Schottky gate was evaporated in the center of the Hall bar. The Ti/Au film was then used as a self aligned mask for a mesa etch to a depth greater than the penetration depth of the 50 kV Be ions. Thus the structure whose cross section is shown schematically in Fig. 1(b) was completed. The final structure was a 10-μm-wide Hall-bar with 5-μm-wide voltage terminals attached at intervals of 30 μm. The Be ions were selectively implanted so that there was no *p*-type region

[a]Electronic mail: hirayama@will.brl.ntt.co.jp

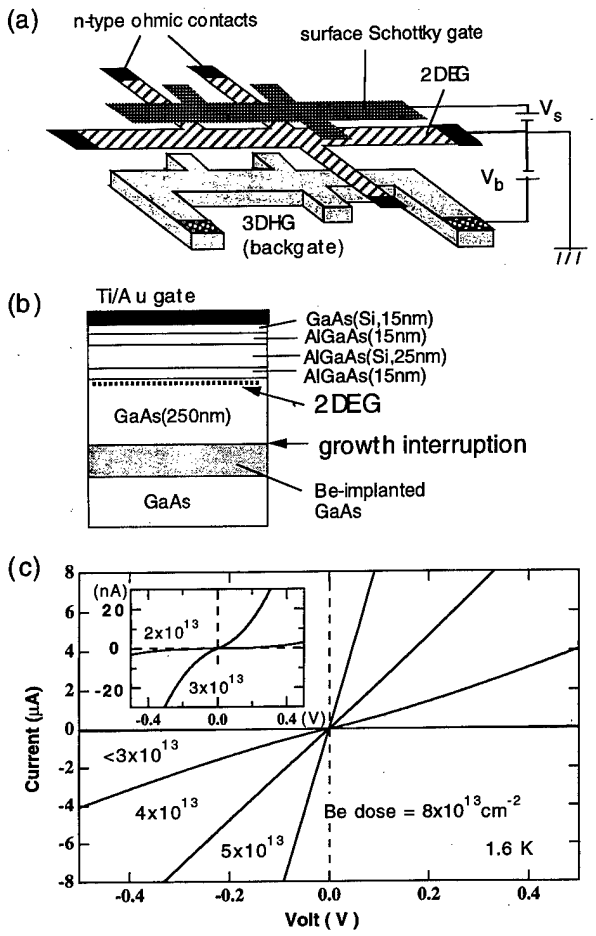


FIG. 1. (a) Schematic diagram of the fabricated backgate devices and (b) the cross section at the center of the Hall bar. In the center region of the Hall bar, the modulation-doped heterostructure and Be-implanted region are simultaneously etched using a Ti/Au surface gate as a self-alignment mask. (c) Low-temperature $I-V$ characteristics of the Be-FIB written p-type region with various Be ion doses.

under the n-type Ohmic contacts. In addition, the self-alignment process protects the Ti/Au gate directly overlapping the Be-implanted region. These features suppress the gate leakage current for both back and surface gates and enable us to control the characteristics of the 2DEG in a rather wide range.

The current–voltage characteristics measured between two p-type ohmic contacts through the Be-implanted region are shown in Fig. 1(c). Though the Be-implanted region exhibits good Ohmic behavior at room temperature, conductance at low temperature drastically decreases when the Be dose becomes less than 10^{14} cm^{-2}. However, linear Ohmic behavior is obtained at 1.6 K when the Be dose is beyond 4×10^{13} cm^{-2} as shown in Fig. 1(c). For a smaller Be dose the Ohmic behavior is not perfect. In spite of this poor Ohmic behavior, the Be-implanted region still acts as a backgate down to the ion dose of 10^{13} cm^{-2} because of the low leakage current between the Be-implanted backgate and the 2DEG.

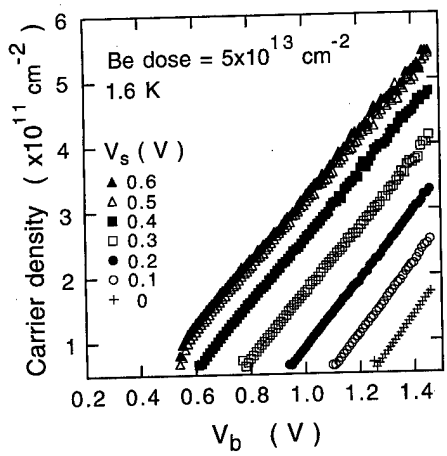


FIG. 2. Electron density change as a function of backgate bias (V_b) at 1.6 K for the backgate device with a Be dose of 5×10^{13} cm^{-2}. Carrier densities of the 2DEG are plotted for several surface-gate bias (V_s) values.

III. TRANSPORT CHARACTERISTICS OF 2DEG

The mobility and electron density of the 2DEG at the heterointerface were estimated from the transport characteristics in a low-magnetic-field region ($B\leqslant0.15$ T) with a dc current of 500 nA. All transport measurements were carried out at 1.6 K. The backgate leakage current was less than 100 pA for the backgate bias (V_b) of less than 1.3 V. The leakage current increased when the bias reached 1.4 V but did not exceed 10 nA in this measurement. The gate leakage current was also limited in the range less than 10 nA for the surface gate. For all the devices with a Be-implanted backgate, the 2DEG is depleted from the back side when $V_b=0$ V. As V_b increases, the 2DEG is formed at the heterointerface. The threshold backgate voltage, where the 2DEG starts to populate, increases with Be ion dose.

Examples of carrier density change as a function of V_b are shown in Fig. 2 for a backgate device with a Be dose of 5×10^{13} cm^{-2}. The voltage applied to the surface gate (V_s) is a parameter. The carrier density increases linearly with increasing V_b and the slope $\partial n/\partial V_b$ is equal for all V_s values. This indicates that the effective distance between the backgate and the 2DEG is constant and independent of V_s and V_b. The slope $\partial n/\partial V_b=4.86\times10^{11}$ cm^{-2}/V corresponds to the distance of 142 nm. Figure 3 shows electron density as a function of V_b for devices with different Be doses. For all Be doses used in this experiment, the slope ∂n/∂Vb is constant and independent of V_s. However, this slope clearly increases with Be dose as shown in Fig. 3. This indicates a reduction in the effective distance between the backgate and the 2DEG and is related to the Be out diffusion as discussed later.

The measured mobility (μ) is ranged from 4×10^4 cm^2/V s (at $n=10^{11}$ cm^{-2}) to 4×10^5 cm^2/V s (at 5×10^{11} cm^{-2}) and μ is proportioned to $n^{1.5}$. This relation suggests a dominant role of remote ionized-impurity scattering.[13] Though the mobility characteristics are discussed in detail in a separate paper,[14] it should be noted here that the mobility values are almost equal to those for the devices without a Be-implanted backgate. For some combi-

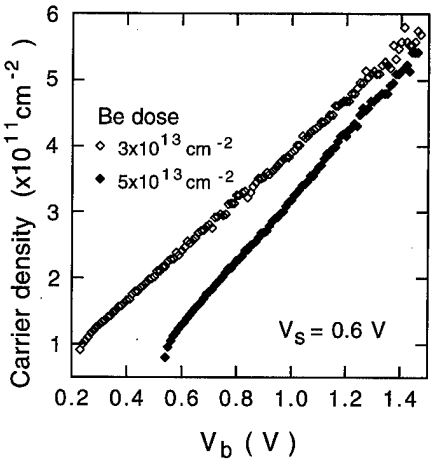


FIG. 3. Comparison of carrier density change as a function of V_b between two devices with different Be doses. These data are obtained at the same V_s value of 0.6 V. The slope $\partial n/\partial V_b$ corresponds to the effective distance between the 2DEG and the backgate.

nations of V_s and V_b the mobility is rather enhanced due to the modification of the electron wave function distribution. These results further suggest that remote ionized-impurity scattering due to the dopants in the Si-doped AlGaAs layer and residual impurities in the AlGaAs spacer layer dominate the mobility and that the trace of implantation-induced damage and the out diffusion of Be in the overgrown layer do not degrade the quality of the 2DEG. This situation is confirmed for the Be dose up to 8×10^{13} cm^{-2}.

IV. SIMS PROFILE MEASUREMENTS

For further investigation of the structure fabricated by the FIB+MBE process, SIMS depth profiles were measured for Al, Si, and Be, and the result is shown in Fig. 4. The large Be ion dose of 1.8×10^{14} cm^{-2} was used to clarify the depth

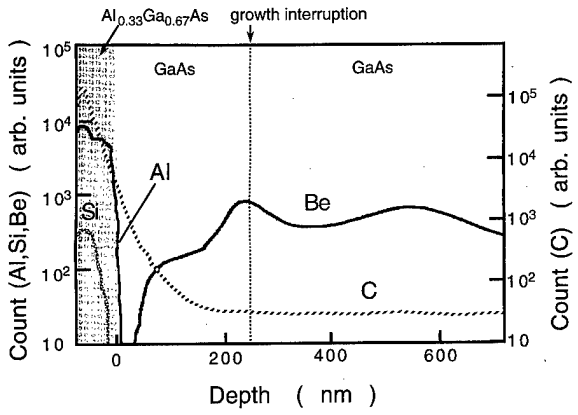


FIG. 4. SIMS profiles of Al, Si, Be and C obtained for the backgate device with Be dose of 1.8×10^{14} cm^{-2}. The dotted line indicates the growth-interrupted and Be-implanted position. The density of Be drops abruptly at the out diffusion front. The depth-profile measurements were started after a few tens of seconds initial sputtering for checking the position. Therefore, the depth of a horizontal axis is calibrated from the interface between AlGaAs and GaAs using the sputter rate of GaAs. The profiles of Al, Si and Be were measured using a positive ion analysis and that of C was measured using a negative ion analysis with a different run.

distribution of Be. The Al and Si distributions clearly indicate the position of the AlGaAs/GaAs heterointerface and Si modulation doping. The Be profile has two peaks. The deep-side peak approximately corresponds to the penetration depth of Be at 50 kV, which indicates it comes from the implanted Be itself. The shallow-side peak agrees well with the growth interrupted position. The Be ions probably pile up at the growth-interrupted surface during the thermal treatment before starting the overgrowth. The out diffusion of Be into the overgrown layer is clearly seen in the depth profile. However, it is important that the edge of the Be out diffusion is rather sharp and the Be signal becomes smaller than the detection limit at the heterointerface in Fig. 4.

The abrupt decrease in Be density enables us to estimate the front of the out diffusion from SIMS data. This out diffusion is enhanced by increasing Be dose. The Be out diffusion length can also be estimated from the transport measurements discussed in the previous section. The $\partial n/\partial V_b$ slope corresponds to the effective distance between the 2DEG and the backgate and the Be out diffusion length can be estimated by subtracting this value from the overgrown GaAs-layer thickness (250 nm). The out diffusion lengths estimated by both methods fall on the same line, which increases linearly with the logarithm of Be ion dose when the Be dose is larger than 10^{13} cm^{-2}. For example, the estimated out diffusion lengths are 76, 108 and 120 nm for the Be dose of 3×10^{13}, 5×10^{13} and 8×10^{13} cm^{-2}, respectively.[14] These values correspond to the distance between the 2DEG and the backgate, or in other words between the 2DEG and the Be out-diffusion edge, of 174, 142 and 130 nm. As discussed previously, good 2DEG properties are maintained even for the Be dose of 8×10^{13} cm^{-2} in spite of the small distance (130 nm) between the 2DEG and the backgate. This electrical property strongly supports the sharp drop of Be density at the front of the out diffusion and is consistent with the SIMS profile measured in Fig. 4.

In addition, the profile of C was also measured and the result is shown again in Fig. 4. In a previous report,[15] C contamination was clearly observed when the growth was interrupted in air but not when interrupted in the high vacuum. The C detected near the surface in Fig. 4 indicates the C contamination of the air exposed surface. This surface contamination makes a tail down to 150 nm in this SIMS profile measurement. The important point is that the measured C profile is flat with the value determined by the background contamination of the SIMS measurement equipment at the growth interrupted position. There is no pile of the C contamination at the growth interrupted interface where the hump of Be is clearly seen. This indicates that a contamination-free process is realized in the all-vacuum FIB +MBE system used in this experiment.

V. DRAG EXPERIMENTS

The Be-FIB written backgate is a 3DHG and the distance between this backgate and the 2DEG becomes as narrow as 130 nm in the fabricated structure. In addition, the Be-FIB is selectively scanned, so that separate contacts are realized for

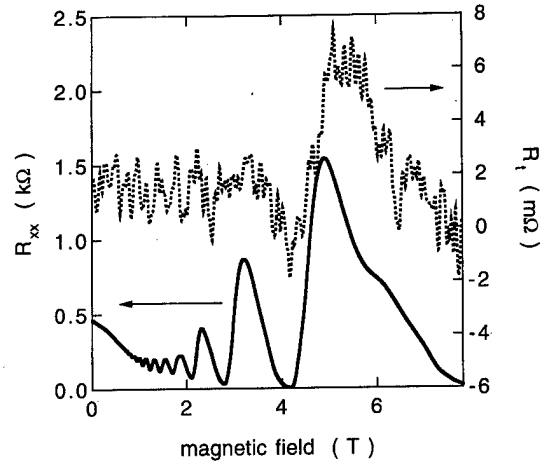


FIG. 5. Longitudinal resistance (R_{xx}) of the 2DEG and the mutual transresistance (R_t) between the 2DEG and the 3DHG measured for the device with the Be-ion dose of 8×10^{13} cm^{-2}. R_t is determined by dividing the drag voltage induced in the 3DHG by the driving current (1 μA). The measurements were carried out at 1.6 K.

the 2DEG and the 3DHG. Therefore, it is possible to apply this structure for the study of the interaction between 2DEG and 3DHG. We measured the drag signal that appeared in the 3DHG when a driven current was applied to the 2DEG, and the result is shown in Fig. 5. The device with Be dose of 8×10^{13} cm^{-2} was used in this measurement because of the good linear Ohmic behavior of the 3DHG at low temperature. The ac lock-in technique with a frequency of 13 Hz was used and the driven current was about 1 μA. The SdH oscillation of the 2DEG was also measured using the voltage terminals schematically shown in Fig. 1(a).

Though the drag signal is very weak and it is difficult to distinguish it from noise in the low magnetic field region, an oscillatory structure similar to the SdH oscillation of the 2DEG is seen in the high magnetic field region. The resistance of the 3DHG varies monotonically as a function of the magnetic field (about 10% decrease by a magnetic field change from 0 to 8 T) and there is no fine structure corresponding to the SdH oscillation of the 2DEG, so that the oscillatory signal induced in the 3DHG probably reflects the interaction between two layers. Our results are similar to those reported for drag experiments between 2DEG and 2DHG (or 2DEGs).[9–11] The sign of the drag signal suggests that a momentum transfer occurs between electrons in the 2DEG and holes in the 3DHG. Though further study is necessary, a phonon mediated drag effect may occur between the 2DEG and the 3DHG.

VI. CONCLUSIONS

A modulation-doped AlGaAs/GaAs heterostructure has been grown on selectively Be-implanted GaAs using all-vacuum FIB+MBE system. The *pn*-junction-type backgate devices can be obtained by growing a heterointerface after the rather thin GaAs layer. The carrier density of the 2DEG is well modified in the range of $(1-6)\times10^{11}$ cm^{-2} for applying a small voltage less than 1.5 V to the gate. The SIMS profile of C indicates contamination-free overgrowth. However, a build up of Be at the growth-interrupted plane and Be out diffusion are observed. Nevertheless, the sharp front of the out diffusion enables us to fabricate the devices without a reduction of mobility even when the effective distance between the 2DEG and the backgate is less than 150 nm. In addition, the drag signal between the 2DEG and the 3DHG can also be measured in this structure.

ACKNOWLEDGMENTS

The authors would like to thank K. Muraki and S. Tarucha for their valuable discussions and N. Uesugi for his encouragement throughout this work. This work was supported in part by the Japan Society for the Promotion of Science ("Research for the Future" Program JSPS-RFTF96P00103).

[1]E. Miyauchi and H. Hashimoto, Nucl. Instrum. Methods Phys. Res. B **7/8**, 851 (1985).
[2]H. Arimoto, A. Kawano, H. Kitada, A. Endoh, and T. Fujii, J. Vac. Sci. Technol. B **9**, 2675 (1991).
[3]F. Wakaya, K. Umeda, J. Yanagisawa, Y. Yuba, S. Takaoka, K. Murase, and K. Gamo, Jpn. J. Appl. Phys., Part 1 **33**, 7223 (1994).
[4]E. H. Linfield, G. A. C. Jones, D. A. Ritchie, and J. H. Thompson, Semicond. Sci. Technol. **8**, 415 (1993).
[5]K. M. Brown, N. Turner, J. T. Nicholls, E. H. Linfield, M. Pepper, D. A. Ritchie, and G. A. C. Jones, Phys. Rev. B **50**, 15 465 (1994).
[6]A. R. Hamilton, E. H. Linfield, M. J. Kelly, D. A. Ritchie, G. A. C. Jones, and M. Pepper, Phys. Rev. B **51**, 17 600 (1995).
[7]K.-J. Friedland, Y. Hirayama, T. Fujisawa, T. Saku, and S. Tarucha, Physica B **227**, 31 (1996).
[8]Y. Hirayama and H. Okamoto, Jpn. J. Appl. Phys., Part 2 **24**, L965 (1985); K. Nakamura, T. Nozaki, T. Shiokawa, K. Toyoda, and S. Namba, *ibid.* **24**, L903 (1985).
[9]N. P. R. Hill, J. T. Nicholls, E. H. Linfield, M. Pepper, D. A. Ritchie, A. R. Hamilton, and G. A. C. Jones, J. Phys.: Condens. Matter **8**, L557 (1996).
[10]H. Rubel, A. Fischer, W. Dietsche, K. von Klitzing, and K. Ebert, Phys. Rev. Lett. **78**, 1763 (1997).
[11]H. Rubel, A. Fischer, W. Dietsche, C. Jorger, K. von Klitzing, and K. Eberl, Proceedings of the 12th International Conference on the Electronic Properties of Two-Dimensional Systems, Tokyo, 1997 (unpublished), p. 525, EP2DS-12.
[12]Although the drag experiments generally have been reported for closely-placed 2DEG double layers (Refs. 9 and 10), the drag effect was clearly observed for a coupling between 2D and 3D for electrons [P. M. Solomon, P. J. Proce, D. J. Frank, and D. C. La Tulipe, Phys. Rev. Lett. **63**, 2508 (1989)]. In case of electron-hole coupling, a combination of 2DEG and 2DHG using modulation doping needs a large layer separation beyond 190 nm (Ref. 11). By replacing 2DHG to 3DHG, the separation can be reduced as shown in this article. Therefore, the present 2DEG-3DHG system is also interesting to study the drag effect.
[13]T. Ando, J. Phys. Soc. Jpn. **51**, 3900 (1982).
[14]Y. Hirayama and T. Saku, Semicond. Sci. Technol. **12**, 1465 (1997).
[15]A. Takamori, E. Miyauchi, H. Arimoto, Y. Bamba, T. Morita, and H. Hashimoto, Jpn. J. Appl. Phys., Part 2 **24**, L414 (1985).

Formation of GaAs/AlGaAs constricted-channel field-effect transistor structures by focused Ga implantation and transport of electrons via focused ion beam induced localized states

H. Kim,[a] T. Noda, and H. Sakaki[b]
Research Center for Advanced Science and Technology, University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153, Japan

(Received 19 November 1997; accepted 1 May 1998)

We report on the fabrication of *n*-AlGaAs/GaAs constricted-channel field-effect transistor (FET) structures by focused ion beam (FIB) implantation, and study transport properties of low-dimensional electrons via FIB induced localized states. In this work, quantum wire FET samples with in-plane gates and top gate were formed by using a focused Ga ion beam by which Ga ions of a high dose density ($>10^{15}$ cm^{-2}) are introduced in the vicinity of a conductive channel at room temperature. The drain conductance G_d was studied in constricted channels with various nominal widths, W. It is found that G_d of a narrow channel ($W\sim 1$ μm) becomes quite small and nonlinear at low temperatures. When G_d is measured as a function of gate voltage, V_g, periodic oscillations are clearly observed at 15 K with a typical spacing of $V_g\sim 60$ mV, whereas the oscillation weakens and transforms to steplike structures at higher temperature (250 K). This oscillatory characteristic may possibly result from single electron transport through quantum dots associated with FIB-induced random potentials in the channel. © *1998 American Vacuum Society.* [S0734-211X(98)04504-1]

I. INTRODUCTION

The implantation of Ga ions using focused ion beam (FIB) can be used to fabricate two-dimensional electron gas (2DEG) devices with very fine feature sizes.[1–4] For example, tunnel junctions with very small capacitance can be prepared, enabling the formation of single electron tunneling devices[5–8] because excellent insulating behaviors[2] are obtained in the implanted regions by such a technique. Note that the FIB process on AlGaAs/GaAs single heterojunctions[9,10] is a convenient single maskless step, for the fabrication of single electron devices as electrons can be laterally constricted by implanted regions. In making use of FIB and other implantation techniques, one must, of course, deal with the damage caused by implanted ions and unintentional incorporation of implanted ions in the conductive channel, particularly when its width is comparable with the scattering range of ions.

In structures defined by FIB implantation, defect states may be formed not only in the implanted region, but also in the neighboring channel.[11] These states may capture electrons, and if the density of states is high enough, the Fermi level in the channel may be influenced. The electron transport in this system is affected by the defects' distribution and also by the random potentials in the constricted channel. To investigate these effects, the conductance of constricted channels was measured with special attentions to the role of localized states in FIB-induced quantum wire channels. As the dose of implants increases, electrons in the channel will be influenced by localized states more effectively than those in lightly doped wires,[12–14] and the role of Coulomb blockade will be enhanced. Hence we employ here the FIB process with very high ($>10^{15}$ cm^{-2}) dose while the typical dose is 10^{12} cm^{-2} for the FIB insulation writing.

This article describes our study of electron transport in constricted channels of various width to evaluate the role of the laterally spread damages locally introduced by Ga ion beam. We show that the transport and field-effect transistor (FET) characteristics of constricted *n*-AlGaAs/GaAs 2DEG channels with the nominal width of 1 μm are influenced by such localized states at low temperatures, as manifested in the conductance oscillation and nonlinear current–voltage characteristics. We point out possible roles of single electron tunneling in these device structures.

II. DEVICE FABRICATION

A modulation-doped *n*-AlGaAs/GaAs single heterojunction was first grown by molecular beam epitaxy by depositing onto a (001) oriented Cr-doped semi-insulating GaAs substrate, successively a 600-nm-thick undoped GaAs buffer layer, a 20-nm-thick undoped AlGaAs spacer layer, a 40-nm-thick Si-doped *n*-AlGaAs layer ($N_{Si}=1\times10^{18}$ cm^{-3}) and a 10-nm-thick GaAs cap layer. A 2DEG was formed at 70 nm underneath the surface. This wafer had an electron mobility μ of 1.4×10^5 cm^2/V s and the sheet electron concentration N_s of 4.3×10^{11} cm^{-2} at 77 K after illumination. A cross-shaped mesa structure of Figs. 1 and 2 was formed on this wafer by optical lithography and wet chemical etching using H_3PO_4 : H_2O_2 : H_2O=1 : 1 : 20 solution. A typical etching rate at room temperature was 300 nm/min. Then, Ohmic contacts were made by depositing Au/Ge, followed by the metal lift-off process and thermal annealing at

[a]Author to whom correspondence should be addressed; electronic mail: kim@quanta.rcast.u-tokyo.ac.jp
[b]Also at Quantum Transition Project, JST, 4-7-6 Komaba, Meguro-ku, Tokyo 153, Japan.

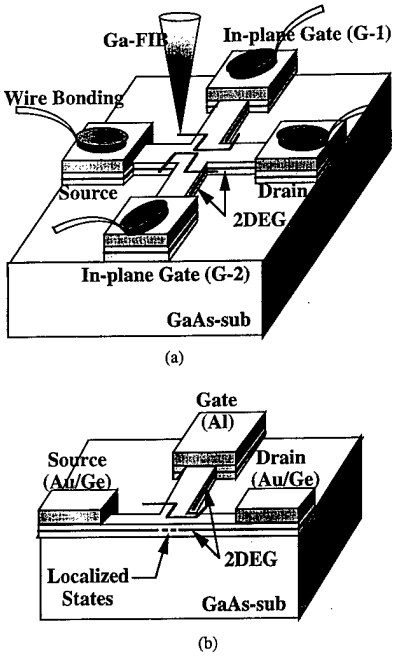


FIG. 1. (a) Schematic view of the device structure used in the experiment. (b) A cross-sectional view of the constriction pattern written by Ga FIB.

400 °C for 60 s. Then, Ga-FIB implantation was performed as illustrated by thin solid lines in Fig. 1 with which two in-plane gates and a constricted channel were formed. For this step, a 100-nm-diam 80 kV Ga-FIB of 12 pA was scanned on the wafer surface to the dose of 1×10^{15} cm^{-2} to form highly resistive regions. Three different samples with constrictions of 1, 5, and 10 μm in the nominal width and 5 μm in length were fabricated. A narrow channel was formed as illustrated in Fig. 1 with no thermal annealing. Finally, the electrical contacts were made by bonding wires on the electrodes at 150 °C.

A scanning electron microscope (SEM) image of the device is shown in Fig. 2. Here the bright lines schematically indicate the Ga-implanted region, along which the resistivity was locally enhanced to constrict 2DEG channel. All the FIB patterns were prepared with single-pass mode, where the beam writes only once over the pattern. The device consists of the source (left), the drain (right), and two in-plane gates (G-1 and G-2). In addition, a 20 μm$\times$20 μm square-shaped top gate (G-3) was formed, as shown by dotted lines in Fig. 2. The electrical characteristics were measured in the temperature range from 13 to 250 K.

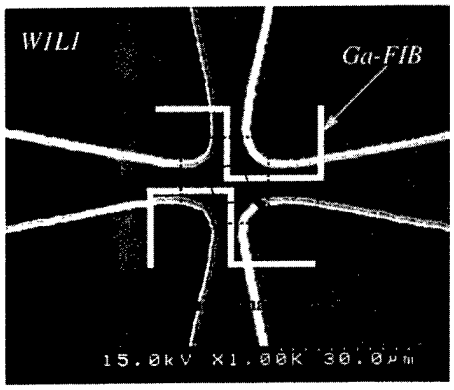


FIG. 2. SEM micrograph of our device structure. The nominal width and length are both 1 μm. Ga^+ ions focused to a spot size of 0.1 μm are implanted into the area schematically shown by white lines.

III. DEVICE CHARACTERIZATION

We first investigated $I-V$ characteristics of three samples A, B, and C for the current flowing between the source and the drain. The constricted channels of these samples had the common length L of 5 μm but different nominal widths with $W=10$, 5, and 1 μm for A, B, and C, respectively. Their drain current versus drain voltage characteristics measured at various temperatures are shown in Figs. 3(a), 3(b), and 3(c). For sample A with 10 μm-wide channel, the conductance increased from 50 to 280 μS as the temperature is lowered from 300 to 4.2 K as shown in Fig. 3(a). Though the increase of conductance at low temperatures suggests the increase of mobility, the measured conductance is far lower than that expected from the electron mobility and concentration of an unprocessed wafer. This implies that the constricted channel of this sample is either substantially narrowed or has a reduced electron concentration.

For sample B with $W=5$ μm, the conductance is substantially lower and less sensitive to temperature than that of sample A and gets nonlinear at low temperatures. This suggests that the entire region of constricted channel is influenced by the FIB process. In addition, the conductance of sample C with the nominal width of 1 μm is Ohmic ($\sim$15 μS) only at 300 K but it gets highly nonlinear as the temperature decreases. The actual channel width is determined by the lateral extension of the FIB-induced damage outside the implanted region. The strong nonlinearity at low temperatures suggests that the conduction in sample C with W ~1 μm is dominated by the random potential fluctuation in the constricted region.

To investigate this transport more in detail, we measured the drain current at $V_d=5$ mV as a function of the top gate (G-3) voltage (V_g) at 15 K. The result is shown in Fig. 4 for sample D whose channel is 1 μm both in length and width. Here, no post-implantation annealing was performed and the localization of electrons may play an important role on electrical characteristics. Even when the lateral spreading of scattered ions induces relatively weak and random modulation of potentials, implants can modify the conduction band and partially create a region of higher resistance. Under an appropriate condition, the FIB-induced modulations of potentionals are large enough to form locally insulating barriers in the channel but low enough to keep the channel mostly conducting.

As shown in Fig. 4, the drain current at 15 K oscillates as a function of gate voltage with the period of about 60 mV. At higher temperatures, the oscillation weakens but clear steplike structures are observed as shown in Fig. 5 with the step spacing of about 40 mV. Assuming these periodic structures result from the charging energy of a small capacitor, C_g, we

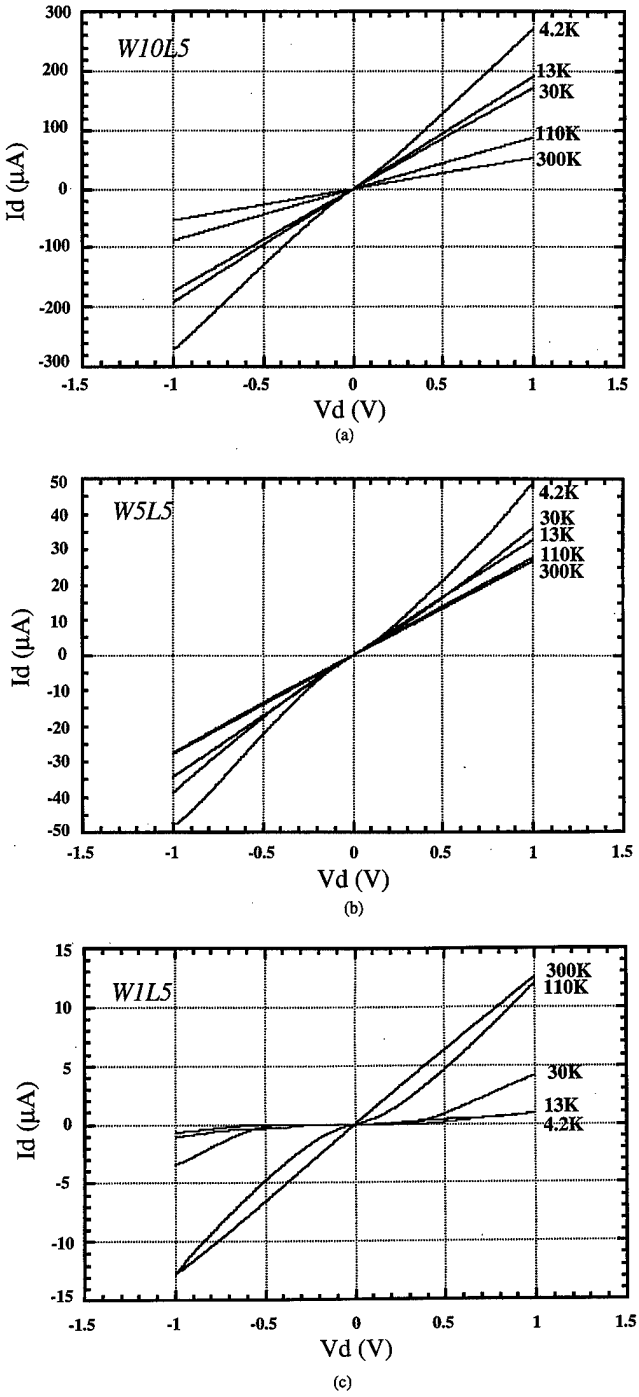


FIG. 3. $I-V$ characteristic of three samples (A), (B), and (C) are shown in (a), (b), and (c), respectively. The channel width (W) and length (L) are 10 μm and 5 μm for sample A, 5 μm and 5 μm, for B, and 1 μm and 5 μm for C.

find that $C_g(e/\Delta V_g)$ is 2.8 aF for ΔV_g of 40–60 mV. There is a possibility that these oscillations result from the single electron tunneling through quantum dots which are formed by the FIB-induced potential fluctuations in the channel region, since coupled dots may be formed in the channel when V_g is near or below the threshold voltage. Note that the FIB-induced lateral confinement potential may not be completely straight along the direction of current flow, as defects induced by the FIB damage are laterally spread and usually distributed around 0.3–0.5 eV below the conduction band.[2]

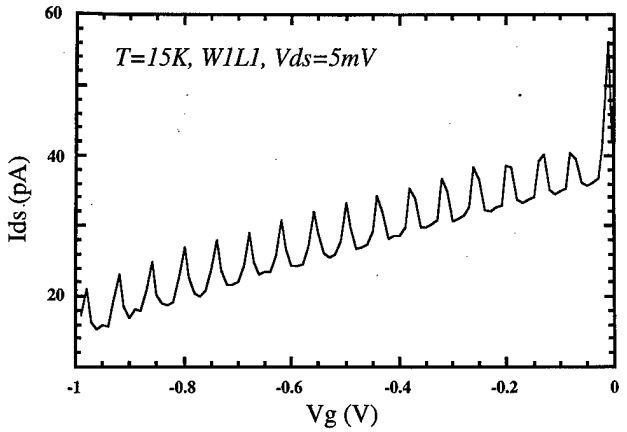


FIG. 4. Gate voltage (V_g) dependence of the two-terminal current between the source (S) and the drain (D) in a sample D with $W=1$ μm and $L=1$ μm. Current oscillations were observed at 15 K.

As the temperature is increased, steplike I_d-V_g characteristics are observed as well as the shift of the threshold voltage. The observed structure may be due to the discreteness of electron charge in the electron transport through a modulated channel where the Coulomb repulsion of electrons in quantum dots plays important roles. A possible explanation for these peaks might be the electron transport through a series of quantum dots formed by the bombardment as the FIB implantation. In our experiments, the spatial extension l of the localized electronic state is estimated from the gate voltage range between the channel definition and the pinch-off, $\delta V_{\text{gate}} \sim e n_s W_{\text{eff}} l / C_{\text{gate}}$,[15] where W_{eff} is an effective channel width wider than 500 nm with which the channel region to be insulating and $n_s = 2.5\times 10^{11}$ cm^{-2} is the electron density in the channel at definition. By using the estimate value of C_{gate} and $\delta V_{\text{gate}} \sim 1$ V, we find it to be ~14 nm. From the junction capacitance and the spatial extension of the localized electron state, the size of multiquantum dots can be estimated.

IV. SUMMARY

We have investigated electrical properties of constricted channels defined by high dose Ga FIB ($>10^{15}$ cm^{-2}) and

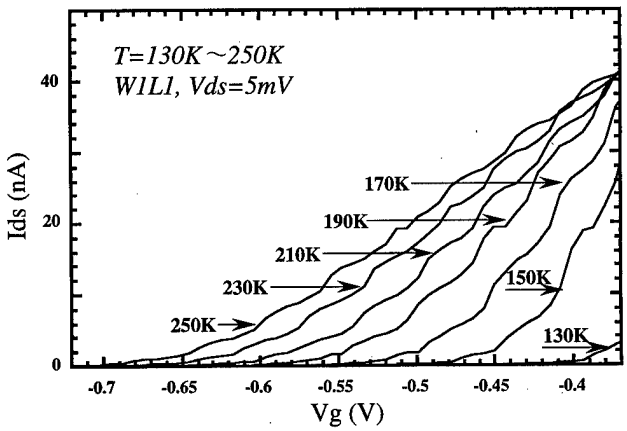


FIG. 5. Temperature dependence of the drain current at drain voltage of 5.0 mV measured as a function of gate voltage at various temperatures between 130 and 250 K.

evaluated the laterally spread FIB-induced defects which seem to cause single electron tunneling via quantum dots in the channel. In particular, we have examined the transport of low-dimensional electrons via FIB-induced localized states in modulation-doped, GaAs/AlGaAs 2DEG constricted-channel FET structures. The current–voltage characteristics in a narrow channel was found to be highly nonlinear, indicative of lateral spreading of scattered ions. The operation of constricted-channel FET structures has been confirmed by conductance measurements. Conductance oscillations are observed and may be attributed to the single electron tunneling via microsegments within the constriction. We speculate that each oscillation corresponds to the addition of a single electron between the self-formed quantum dots in the constrictive channel.

ACKNOWLEDGMENTS

This work was partly supported by a Grant-in-Aid from the Ministry of Education, Science, Sports, and Culture, Japan.

[1]Y. Hirayama, T. Saku, and Y. Horikoshi, Phys. Rev. B **39**, 5535 (1989).
[2]Y. A. Soh and G. L. Snider, J. Vac. Sci. Technol. B **11**, 2629 (1993).
[3]M. Itoh, T. Saku, and S. Tarucha, Jpn. J. Appl. Phys., Part 1 **31**, 4487 (1992).
[4]Y. Hirayama, A. D. Wieck, and K. Ploog, J. Appl. Phys. **72**, 3022 (1992).
[5]Y. Hirayama and H. Okamoto, Jpn. J. Appl. Phys., Part 2 **24**, L965 (1985).
[6]C. J. Goddings, H. Mizuta, J. R. A. Cleaver, and H. Ahmed, J. Appl. Phys. **76**, 1276 (1994).
[7]C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).
[8]Y. Nagamune, H. Sakaki, L. P. Kouwenhoven, L. C. Mur, and C. J. P. M. Harmans, Appl. Phys. Lett. **64**, 2379 (1994).
[9]T. Fujisawa, Y. Hirayama, and S. Tarucha, Appl. Phys. Lett. **64**, 2250 (1994).
[10]T. Fujisawa and S. Tarucha, Appl. Phys. Lett. **68**, 526 (1996).
[11]T. Bever, G. Jager-Waldau, M. Eckberg, E. T. Heyen, H. Lage, A. D. Wieck, and K. Ploog, J. Appl. Phys. **72**, 1858 (1992).
[12]J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62**, 583 (1989).
[13]R. J. Blaikie and K. Nakazato, Appl. Phys. Lett. **64**, 118 (1994).
[14]J. A. Nixon and J. H. Davies, Phys. Rev. B **41**, 7929 (1990).
[15]C. W. J. Beenakker, Phys. Rev. B **45**, 9222 (1992).

Hopping conduction and localized states in p-Si wires formed by focused ion beam implantations

Hirotaka Iwano,[a] Shigeaki Zaima,[b] and Yukio Yasuda
Department of Crystalline Materials Science, School of Engineering, Nagoya University, Furo-cho Chikusa-Ku, Nagoya 464-8603, Japan

(Received 20 November 1997; accepted 14 April 1998)

Localized states in p-Si wires formed in n-Si(100) substrates by selective ion implantation using a focused Ga^+ ion beam have been investigated. The electrical conductance has the temperature dependence of one-dimensional (1D) variable-range-hopping conduction in the temperature range below 50 K, $\sigma=\sigma_0 \exp[-(T_0/T)^{1/2}]$. The magnetoresistance $R(H)$ of p-Si wires at 4.2 K shows the negative and positive magnetoresistances at weak and strong magnetic fields, which have the relation that $R(H)/R_0 \propto \exp(-\beta H)$ and $R(H)/R_0 \propto \exp(\alpha H^2)$, respectively, where R_0 is the resistance without the magnetic field H. These characteristics can be explained by the 1D hopping conduction mechanism under the magnetic field. By expanding a three-dimensional model, we derive an equation of positive magnetoresistance in 1D hopping conduction. From the temperature dependence of conductance and positive magnetoresistance, the localization length, hopping distance, and density of localized states are estimated to be about 2 nm, 8–9 nm and about 10^9 cm^{-1} eV^{-1}, respectively. This result indicates that the carrier transport is confined with the region of a few tens of nanometers. © *1998 American Vacuum Society.* [S0734-211X(98)01304-3]

I. INTRODUCTION

The size of individual devices in ultralarge-scale integrated circuits (ULSIs) is at the submicron scale and, within the next decade, the device dimensions will be reduced to below 100 nm.[1] In such small devices, the device functions will be affected by quantum transport properties. By harnessing the quantum effects such as quantum interference effects, single-electron phenomena and so forth, the development of new devices is also expected.[1–5] Therefore, both the development of semiconductor microfabrication technology and the elucidation of carrier transport phenomena in the low-dimensional system are important.

In our previous work,[6,7] we have developed a successful method to fabricate p-Si wires using the focused ion beam (FIB) doping and investigated the electrical properties of wires. Selective ion implantation by FIB is considered to be one of the simplest methods to realize low-dimensional carrier transport systems. The conducting area of low-doped p-Si wires has been completely controlled by changing the depletion region width in pn junctions with the reverse bias.[6] It has also been found that one-dimensional variable-range-hopping (1D-VRH) conduction is dominant in highly doped wires at low temperatures.[7] However, details of the 1D-VRH systems, such as localization length and so forth, were not clear. In this work, the localization length and hopping distance in the 1D-VRH regime are discussed by measuring magnetoresistance.

[a]Electronic mail: iwano@alice.xtal.nagoya-u.ac.jp
[b]Also with Center for Cooperative Research in Advanced Science and Technology, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan.

II. EXPERIMENTS

Focused Ga^+ ion beams with a diameter of less than 0.1 μm were used to form p-type conductive wires by scanning the surface of n-type Si(100) substrates with a resistivity of 2–3 Ω cm. The length of p-type wires was 50 μm. The accelerating voltage and beam current of Ga^+ ions was 100 keV and 50–130 pA, respectively. The annealing was performed in order to activate the implanted Ga atoms electrically and to recover implantation-induced damages. The following three samples were prepared: The ion dose and annealing temperature were: (a) 5.6×10^9 cm^{-1} and 600 °C, (b) 3.5×10^9 cm^{-1} and 600 °C, and (c) 5.2×10^9 cm^{-1} and 690 °C, respectively. The annealing was performed in N_2 gas for 30 min. The fabrication conditions are summarized in Table I. Before the FIB implantation, heavily doped p-type regions were formed as electrodes by the conventional ion implantation of BF^+ with a dose of 2×10^{15} cm^{-2} at an accelerating voltage of 30 keV and samples were annealed at 900 °C for 30 min.

The electrical conductance was measured in the temperature range from 4.2 to 300 K by using a liquid He cryostat. In the present experiment, no substrate bias was applied. The magnetoresistance was also measured in the range from 0 to 4 T at 4.2 K by applying magnetic field perpendicular to the direction of current flow.

III. RESULTS AND DISCUSSIONS

A. Temperature dependence

Figure 1(a) shows the temperature dependence of electrical conductivity in a 1D system. The data show the presence of two parallel conduction mechanisms. One is the carrier conduction in the valence band, which has an activation-type temperature dependence due to the carrier activation process

TABLE I. Fabrication condition of p-Si wires.

Sample	A	B	C
Line dose (cm^{-1})	5.6×10^9	3.5×10^9	5.2×10^9
Annealing temp. (°C)	600	600	690

TABLE II. Estimated values of $a\cdot\rho_F$.

Sample	A	B	C
$a\rho_F(eV^{-1})$	487	199	284

to the valence band in the temperature range above 50 K, and the other is 1D-VRH below 50 K, as reported previously,[6,7] in which filamentary quasi-1D paths are considered to be formed.[6,7] For all samples prepared in this experiment, the temperature dependence of electrical conductivity is well fitted by the following relationship below 50 K:[6,7]

$$\sigma=\sigma_0\exp[-(T_0/T)^{1/2}], \tag{1}$$

where T_0 is a characteristic temperature, and T the absolute temperature. This fact is a characteristic feature of the 1D-VRH conduction and filamentary quasi-1D paths are considered to be formed in the ion-implanted wire region.[7,8] The characteristic temperature T_0 includes information about the localized states in the conductive p-Si region. According to the percolation approach by Butcher and McInnes,[8] the T_0/T of 1D-VRH systems is given by

$$T_0/T=2N_{p1}(T_0'/T), \tag{2}$$

where $k_BT_0'=(a\rho_F)^{-1}$. Here, N_{p1} is the mean number of percolation paths per site in the 1D case, ρ_F the density of localized states at the Fermi level and a the localization length. The value of N_{p_1} is evaluated to be 4.5 by simulation.[8,9] By using these relationships, the values of $a\rho_F$ for each sample can be estimated from the experimental data and are indicated in Table II.

In the VRH system, the electrical conductivity can also be given by the following expression in a percolation approach,[10,11]

$$\sigma\propto\exp(-2r_c/a), \tag{3}$$

where r_c is the critical hopping distance. The temperature dependence of conductivity in VRH systems results from hopping from the fact that the hopping distance has the temperature dependence of $r\propto T^{-p}$, where $p=1/(d+1)$ and d the dimension of the system.[12] Since the whole conductivity is considered to be dominated by the conductivity at a critical hopping distance r_c,[10] the r_c is used to estimate the hopping distance r in the following discussion.

B. Magnetoresistance

Figure 2 shows the magnetoresistance, $R(H)/R_0$, of samples at 4.2 K, where R_0 is the resistance without the magnetic field. The electric field of 200 V/cm was applied for all samples. Negative magnetoresistance (NMR) clearly appears for sample B below 1 T. In the VRH regime, the probability of each hop is affected by the other localized states due to scattering and interference effects in tunneling processes, because the hopping distance exceeds the average interimpurity distance. The magnetic field affects the phase factor of the hopping probability and the negative magnetoresistance is also expected.[13] The magnetic field dependence of NMR in hopping conduction is described by $R(H)/R_0\propto\exp(-\beta H)$ at weak magnetic fields.[13] The experimental results below 1 T in Fig. 2 are explained by this relationship. As discussed later, a hopping distance in sample B is largest in samples, which is considered to correspond to the appearance of NMR in sample B. The $R(H)/R_0$ changes from negative to positive around 2 T for sample B. Above 2 T, the positive magnetoresistance has a relationship of $R(H)/R_0\propto\exp(\alpha H^2)$ for all samples.

The positive magnetoresistance can be explained by the squeeze of the envelope wave function of localized states in magnetic field, which result in the decrease in localization

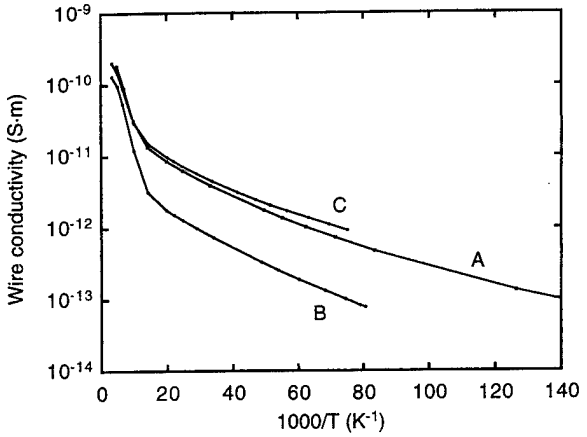


FIG. 1. Temperature dependence of conductivity for samples A, B, and C.

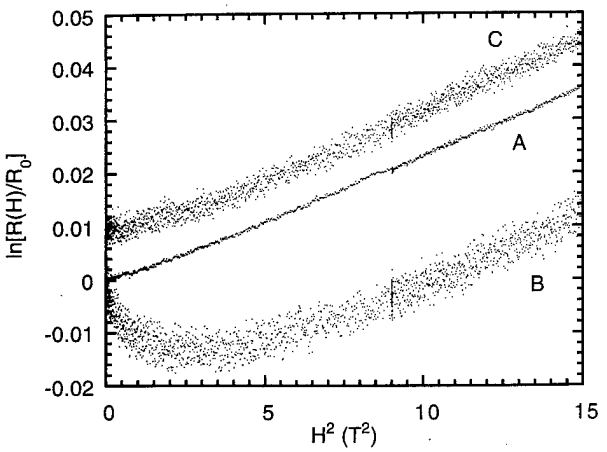


FIG. 2. Normalized magnetoresistance at 4.2 K for samples A, B, and C. The curve of sample C is shifted by +0.01 in order to be distinguished from that of sample A.

length and hopping probability. By a percolation approach in the three-dimensional (3D) case, the positive magnetoresistance is expressed by[12]

$$\ln[R(H)/R_0]=t_3(e^2a/\hbar^2)r^3H^2, \tag{4}$$

where $R(H)$ is the resistance at a magnetic field of H, e is the elementary charge, h the Planck constant and t_3 the parameter. The parameter t_3 is estimated to be $N_{p3}/24\pi=0.036$ ($N_{p3}=2.7$), where N_{p3} is the threshold value of the mean number of bonds per site in the 3D case.[12] It is necessary to derive the expression of magnetoresistance in 1D cases. The same relationship can be obtained in the case of 1D hopping conduction by expanding the 3D model,[12] as mentioned below.

Under the magnetic field applied to the z direction, the envelope wave function of a localized state $F(r)$ in the Coulomb field is given by[12]

$$F(r)\propto\exp[-r/a_B-l^2ra/(24\lambda^4)], \tag{5a}$$

$$\lambda=(\hbar/eH)^{1/2}, \tag{5b}$$

where $l^2=x^2+y^2$, and a_B is the effective Bohr radius which is consistent with the localization length a in Sec. III A. In the following discussion, symbol a is used instead of a_B. The resistance between two localized states $R'(H)$ is given by:

$$R'(H)=R'_0\exp(\xi), \tag{6a}$$

$$\xi=-2r/a-l^2ra/(12\lambda^4). \tag{6b}$$

In the following discussion, the condition $a\ll l\ll\lambda^2/a$ is assumed, which is consistent with estimated values in the following analysis. Since the second term of Eq. (6b) is as small as $\delta=(ra/\lambda^2)\ll1$, squaring both sides of Eq. (6b) and discarding the term δ^2, we obtain the equation of an ellipsoid:

$$z^2/A^2+l^2/C^2=1, \tag{7a}$$

where

$$A=a\xi/2$$

and

$$C=a\xi/2[1-a^4\xi^2/(96\lambda^4)]. \tag{7b}$$

In order to obtain Eqs. (7a) and (7b), r^2 in the term of first order in δ is replaced with $(a\xi)/2$.[12] In the percolation approach, the percolation threshold ξ_c to form conductive networks can be obtained by solving the equation:

$$N_p=V_\xi\cdot N, \tag{8}$$

where V_ξ is the volume of the ellipsoid of a site, and N the density of sites.[12] In the 3D case, V_ξ and N_p are $(4\pi/3)AC^2$, and N_{p3}, respectively. In order to obtain the form of the 1D case, a 1D array of sites is assumed and N is the 1D density of sites. In addition, V_ξ and N_p are replaced to $2C$ and N_{p1}, respectively. By solving Eq. (8) for the 1D case, ξ_c shows magnetic field dependence of $N_{p1}^3a/(96\lambda^4N^3)$. By substituting ξ_c into Eq. (6b), the following equation is obtained for the 1D case:

TABLE III. Estimated values of a, r and ρ_F at 4.2 K.

Sample	A	B	C
a (nm)	2.2	1.6	1.9
r (nm)	8.0	9.2	8.7
ρ_F (cm^{-1} eV^{-1})	2.2×10^9	1.2×10^9	1.6×10^9

$$\ln[R(H)/R_0]=t_1(e^2a/\hbar^2)r^3H^2,$$
$$t_1=N_{p1}^3/96, \tag{9}$$

which has same form as Eq. (4) except for t.

By using Eqs. (3) and (9), the values of r/a and ar^3 are obtained experimentally and the localization length a and hopping distance r can be estimated. The density of localized states ρ_F is also obtained by using Eq. (2). The estimated values of a, r and ρ_F are summarized in Table III. The results show that the localization length and hopping distance are about 2 and 8–9 nm, respectively. There have been several reports on the localized states in the amorphous semiconductor systems.[14,15] Those reports have indicated that the localization length is about 0.1–1 nm and the hopping distance is about 10 nm. The localization lengths obtained in this work are very close to the effective Bohr radii of light and heavy holes bounded by Ga impurities in Si, 1.8 and 1.0 nm, respectively. The hopping distance is comparable to a distance between activated Ga atoms which is estimated from the ion implantation and annealing condition to be on the order of 10 nm. The appearance of 1D-VRH conduction suggests that the effective diameter of the conductive region is comparable to the hopping distance, which means that the effective wire diameter is expected to be less than a few tens of nanometers. Therefore, it can be concluded that the carrier confinement in the nano-size region is achieved for these samples.

The estimated density of localized states, ρ_F, at the Fermi level are about 10^9 cm^{-1} eV^{-1}. A ratio of ρ_F between samples A and B is close to that of the implanted ion doses, which also supports the validity of the analysis in this work. It is difficult to estimate the number density of localized states, because the energy distribution of the density-of-states cannot be evaluated. However, if the effective bandwidth is assumed to be on the order of thermal energy, the number density of localized states is estimated to be about 10^7 cm^{-1}. This value is about ten times smaller than the electrically activated Ga impurity atoms, since the activation efficiency of Ga atoms in Si substrates has been about 5% at an annealing temperature of 690 °C.[6] This fact also suggests that the electrical conductive region is a part of the ion implanted region. Further investigation is necessary to understand the density-of-states of localized states formed in p-Si wires.

IV. CONCLUSIONS

The 1D-VRH conduction in p-Si wires fabricated by selective ion implantation using FIB in n-Si(100) substrates has been investigated. The ion implantation doses and an-

nealing temperatures are changed in the range of 3.5×10^{9}–5.6×10^{9} cm^{-1} at 600–690 °C, respectively. The temperature dependence of wire conductance has the relation of $\sigma\propto\exp[-(T_0/T)^{1/2}]$ below 50 K. The negative and positive magnetoresistance are observed at 4.2 K for weak and strong magnetic fields, which showed the relationships of $R(H)/R_0\propto\exp(-\beta H)$ and $R(H)/R_0\propto\exp(\alpha H^2)$, respectively. Both magnetoresistances are considered to be the characteristic feature of hopping conduction. In this work, the positive magnetoresistance of 1D hopping systems is investigated by expanding a 3D model. From the analysis of 1D-VRH conduction, the localization length and hopping distance are estimated to be about 2 and 8–9 nm, respectively. The results suggest that the diameter of the conductive regions in wires is less than a few tens of nanometers. The densities of localized states are 1.2×10^{9}–2.2×10^{9} cm^{-1} eV^{-1}.

[1]M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, and H. Iwai, Tech. Dig. Int. Electron Devices Meet. **XXX**, 119 (1993).

[2]H. Abe, K. Hamasaki, and T. Ishiguro, Jpn. J. Appl. Phys., Part 1 **33**, 7210 (1994).

[3]K. Yano, T. Ishii, and K. Seki, IEEE Trans. Electron Devices **41**, 1628 (1994).

[4]A. Ohta, H. Niyama, T. Shibata, K. Nakajima, and A. Toriumi, Jpn. J Appl. Phys., Part 1 **34**, 4485 (1995).

[5]P. Castrillo, D. Hessman, M. E. Pistol, J. A. Prieto, C. Pryor, and L. Samuelson, Jpn. J. Appl. Phys., Part 1 **36**, 4188 (1997).

[6]H. Iwamo, S. Zaima, Y. Koide, and Y. Yasuda, J. Vac. Sci. Technol. B **11**, 61 (1993).

[7]H. Iwano, S. Zaima, T. Kimura, K. Matsuo, and Y. Yasuda, Jpn. J. Appl. Phys., Part 1 **33**, 7190 (1994).

[8]P. N. Butcher and J. A. McInnes, Surf. Sci. **196**, 139 (1988).

[9]G. E. Pike and C. H. Seager, Phys. Rev. B **10**, 1421 (1974).

[10]V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. **4**, 2612 (1971).

[11]R. A. Serota, R. K. Kalia, and P. A. Lee, Phys. Rev. B **33**, 8441 (1986).

[12]B. L. Shklovskii and J. A. Efros, *Electric Properties of Doped Semiconductor* (Springer, Berlin, 1984).

[13]B. I. Shklovskii and B. Z. Spivak, *Hopping Transport in Solid* (North-Holland, Amsterdam, 1991).

[14]A. Lewis, Phys. Rev. Lett. **29**, 1555 (1972).

[15]M. L. Knotek, M. Pollak, and T. M. Donovan, Phys. Rev. Lett. **30**, 853 (1973).

End-point detection using focused ion beam-excited photoemissions in milling deep small holes in large scale integrated circuit structures

Hiroshi Yamaguchi,[a] Keiya Saitoh, and Michinobu Mizumura
Production Engineering Research Laboratory, Hitachi, Ltd., 292 Yoshida-cho, Totsuka-ku, Yokohama 244-0817, Japan

(Received 13 April 1998; accepted 14 May 1998)

An experimental instrument was made for the end-point detection of a deep small hole milled in large scale integrated circuit using focused ion beam (FIB). The SCANIIR method using photon-counting technique is adopted for detecting a very weak photoemission. The objective lens and imaging lens made by a quartz single lens are arranged in a parallel beam configuration. Observing optics, *XY* stage and *Z* controlling mechanism for the optics are also used. To avoid the noise from the wall of the hole, a blanking circuit for detected signals was used. Experiments using a wide angle objective optical system showed an aluminum layer lower than 10 μm and the SiO_2 layer below were able to be clearly discriminated. © *1998 American Vacuum Society.* [S0734-211X(98)06604-9]

I. INTRODUCTION

Recently, milling of conductors and insulators in large scale integrated circuit (LSI) by focused ion beam (FIB) has been widely used in the area of developing, manufacturing, and testing of LSIs, especially for failure analysis and on-chip modification.[1–4] As a tool for the end-point detection in milling, FIB-excited photoemission detection has been used.[5,6] This method is an application of surface composition by analysis of neutral ion impact radiation (SCANIIR), an analysis technique using photoemission from the atoms excited by ion beam impact.[7,8] In the FIB-excited photoemission detection method as shown in Figs. 1(a) and 1(b), the sputtered atom from the hole is excited by the ion beam and emits a characteristic spectrum for the elements, and the emission is detected by a photodetector through wavelength-selection filters. The change of the material is detected by the change of the emission signal. This method has a great advantage when compared with the other technique, detecting the change of the secondary charged particle or absorbed current of the target, because it is hardly affected by the secondary electrons from the target and the charge neutralizer.

However, LSI devices are becoming more and more dense, with multilayer structures and fine conductive lines, and it is becoming more and more difficult to detect the end point of the deep, small hole milled by FIB, since the photoemission from the bottom of the hole becomes very weak and noisy. So it is necessary to develop a photodetecting technique for such a deep small hole milling. In this article, we report on how we have amended the method to improve its effectiveness for deep small holes.

II. PARAMETERS DEFINING PHOTOEMISSION QUANTITY

As shown in Fig. 2, holes in LSI now need to be milled deeper and narrower than previously required. Previously, holes were typically 8 μm deep, 5 μm wide at the top, and 3.6 μm wide at the bottom resulting in an emissive solid angle of 0.285 sr; a 1.5 nA focused ion beam with 1 μm diameter was used.[4] Now, typical holes can be 10 μm deep, 3.8 μm wide at the top and 2 μm wide at the bottom resulting in an emissive solid angle of 0.111 sr; a focused 40 pA ion beam with diameter 40 nm from a FIB column with dual-lens optics is used for milling.[9]

The parameters defining the quantity of photons detected are shown in Fig. 3. When a focused ion beam with current i bombards the bottom of the hole, atoms are sputtered with sputtering yield η. These atoms emit light with excitation probability q. The light emitted at emissive solid angle Ω_1 and at incident solid angle Ω_2 passes through the detection optics, which has transparency T, and is detected by a photon counter with detection efficiency μ. The detected photon count N is expressed as

$$N = k\, i\, \eta\, q\, \Omega_1\, \Omega_2\, T\, \mu,$$

where k is a constant.

Ward reported that N is of the order of 10^4 cps,[5] which is a bit larger than the noise level of normal analog detection

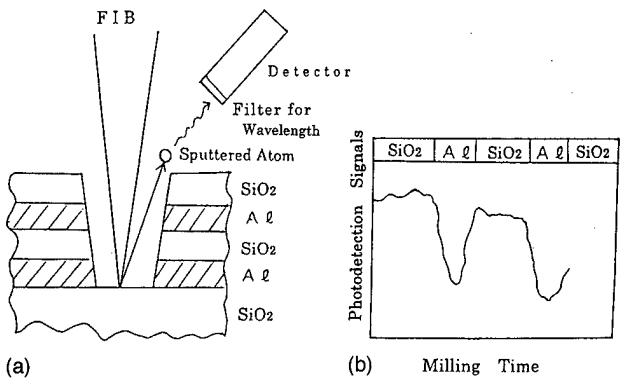


FIG. 1. (a) Depth monitoring with FIB-excited photoemission for LSI milling. (b) Results of the detection of photoemission.

[a]Electronic mail: hyamag@perl.hitachi.co.jp

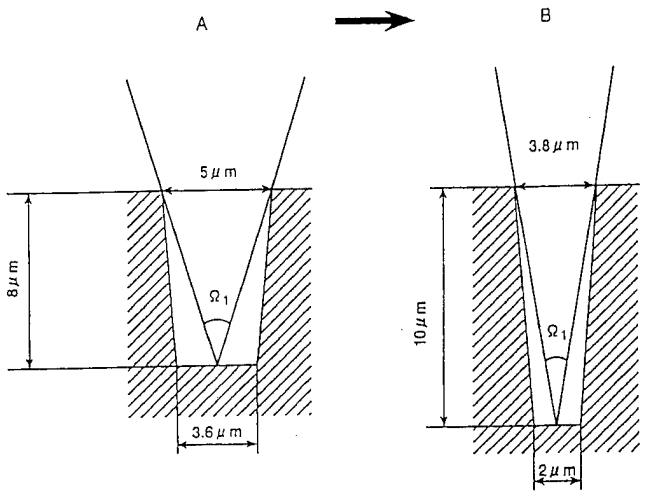


FIG. 2. Photoemission angles when deep small holes are milled. (A) Previous typical hole structure: Photoemission angles: 0.285 sr, FIB current: 1.5 nA. (B) Current typical hole structure: Photoemission angles: 0.111 sr, FIB current: 40 pA.

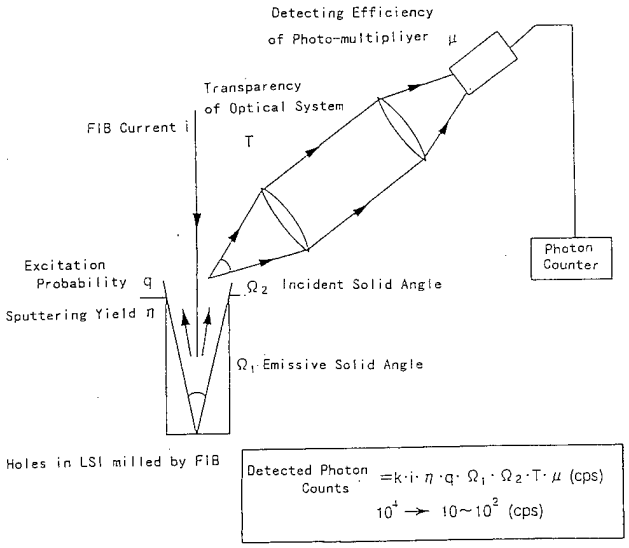


FIG. 3. Parameters defining photoemission quantity.

(10^{-10}–10^{-9} A, which is about 10^3 cps). In this case, N is estimated to be 10–10^2 cps (assuming the sputtering yield of aluminum to be 4.3 atoms/ion, with excitation probability 2 $\times 10^{-4}$ photon/atom). Therefore, a photon-counting technique in which the photons are counted one-by-one is needed instead of an analog-measurement one.

III. OPTICAL SYSTEM FOR PHOTODETECTION

The experimental optical system for photodetection we developed is shown in Fig. 4. The optical axis is set at an angle of 80° from the incidental FIB axis, so that the top of the optical measurement system does not interfere with the ion beam optics. One quartz window (W_1) is placed in front of the objective lens to protect it from the sputtered atoms. A second quartz window (W_2) is placed at the end of the vacuum chamber port to seal it. An XY stage with a Z micron, a tilting mechanism and bellows are used for aligning the optical axis. The main part of the optical measurement system is located behind the second quartz window, therefore it is out of the vacuum. It consists of a half-mirror (M_1), an imaging lens (L_2), a second half-mirror (M_2), an observation lens (L_3), a filter for wavelength selection (F), and a photomultiplier (PM). The illumination light comes from a halogen lamp, passes through an optical-fiber light guide, is collimated by a lens (L_4), and reflected by the first half-mirror. It then passes through the second window, is focused by the objective lens, and illuminates the sample. The light reflected from the sample is observed through the objective, imaging, and observation lenses, and used for aligning and positioning the system. The excited light from the sample passes through the objective and imaging lenses, then the filter selects the desired wavelength and passes that light into the photomultiplier. Finally, the light-pulse signals are counted by the photon counter.

The experimental focused ion-beam milling system[9] we used for evaluation is shown in Fig. 5. A dual-lens ion beam optical column is set on the vacuum chamber. The sample is

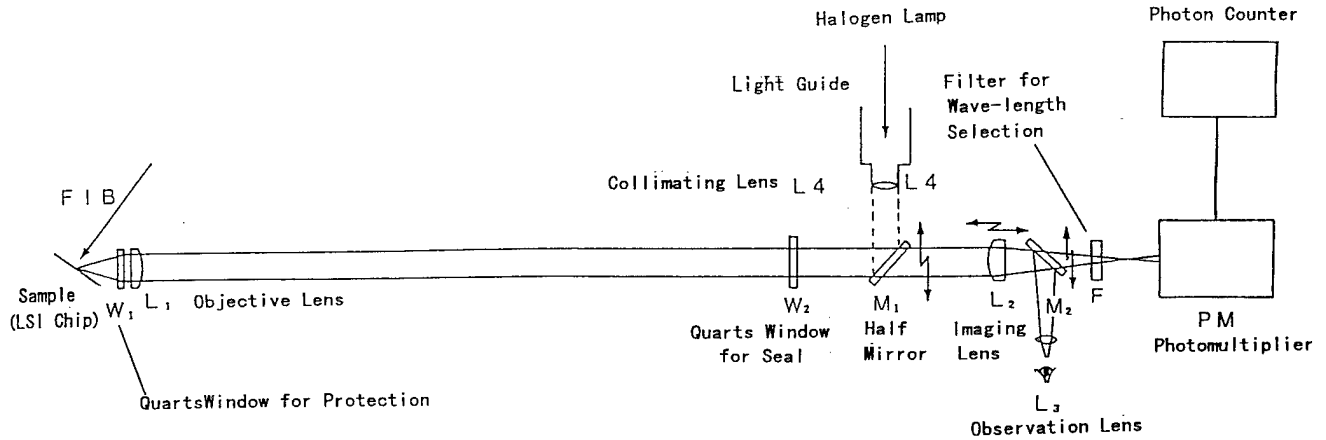


FIG. 4. Optical system for photodetection.

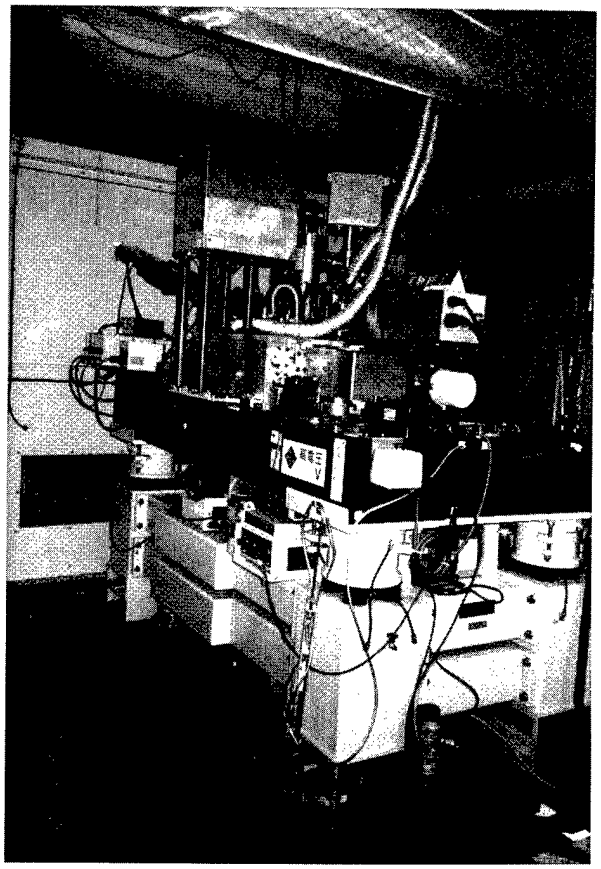


FIG. 5. Experimental focused ion beam milling system.

set on the stage and introduced into the chamber through load locks. The detection optical system is set on the left-hand side of the chamber. The whole system is set on an air servomount to protect it against vibration. A photo of the experimental photodetection system is shown in Fig. 6. It clearly shows the *XY* stage, the optics with observation lens, the photomultiplier, the signal amplifier, and the photon counter.

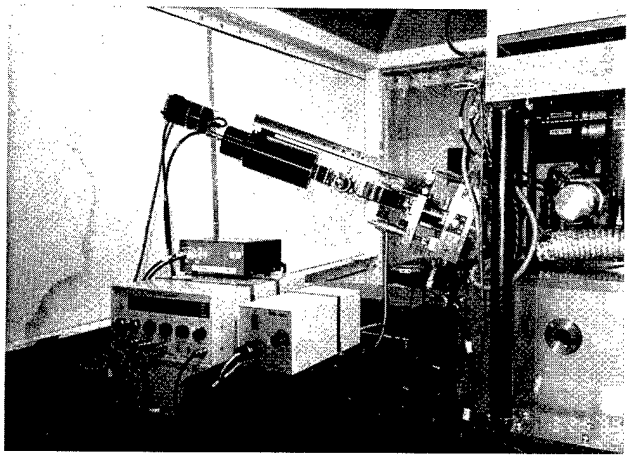


FIG. 6. Experimental photodetection system.

IV. RESULTS AND DISCUSSION

A sample LSI chip having the cross section shown in Fig. 7 was milled using a focused ion beam with a diameter of 40 nm and a beam current 40 pA. The detected photon count versus the milling time is also shown in Fig. 7. The depths of the SiO_2 and aluminum layers were estimated from the dose and sputtering yield, as shown in the figure. In this case, the wavelength-selection filter was not used because the photoemission signal was too weak for the selected wavelengths. Consequently, the total photon counts are shown. The photon emission yield of the SiO_2 layer was twice that of the aluminum one. The graph shows that the changes in the signal when the layer changed became more noisy and less clear with the milling time, i.e., as the drilling reached the lower layers. While the first SiO_2 layer, the first aluminum layer, the second SiO_2 layer, and the second aluminum layer could be discriminated, the lower layers could not.

The possible reasons for this loss of clarity in the signal changes and measures we took to overcome this problem are as follows.

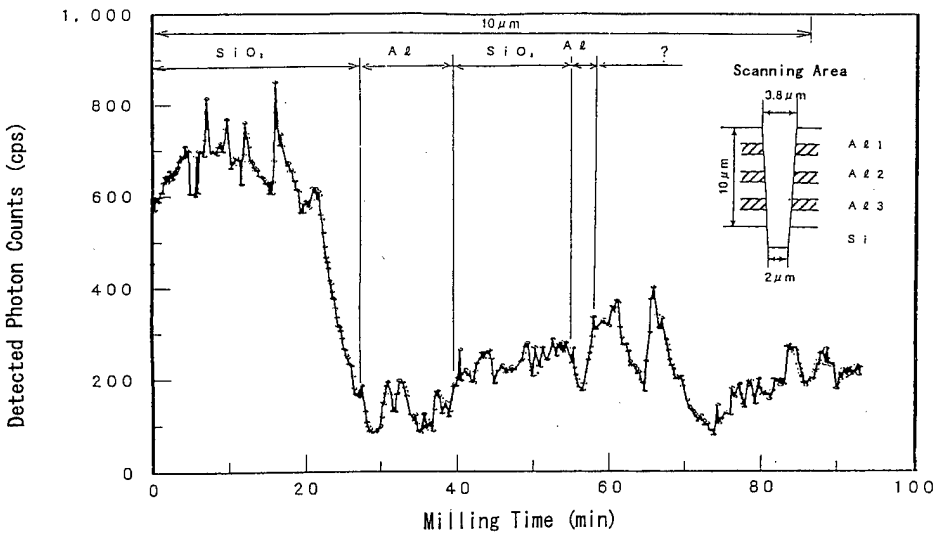


FIG. 7. Relation between milling time and detected photon counts.

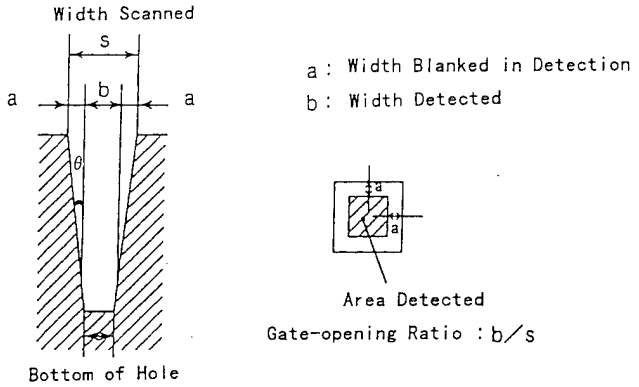


FIG. 8. Partial scan and partial detection of a FIB-milled hole. (To remove effect of the influence of photoemission signal coming from the wall of the hole.) θ=10°–20°.

(1) The S/N ratio is too low, especially at lower layers. This is because noise from the wall of the hole becomes mixed with the signal from the bottom. We have added a blanking circuit to remove the signals coming from the wall.

(2) The signal itself is too weak. We have changed the design of the objective optical system to increase the incident solid angle. We also changed the aperture of the FIB optics to increase the FIB current.

V. PARTIAL SCAN AND PARTIAL DETECTION OF FIB-MILLED HOLE

The wall of a FIB-milled hole becomes slanted by 80°–90° due to the sputtered atoms that are deposited on it (Fig. 8). Even if the scanning area is constant, the bottom becomes narrower as the hole becomes deeper, so the signals from the atoms in the wall become mixed with the desired signal. The wall of the hole is made of the upper aluminum layer, the upper SiO_2 layer, and redeposited atoms, resulting in the low S/N ratio of the photodetection.

To solve this problem, we added a circuit to blank the signals from the wall. This circuit closes the gate between the photomultiplier and the photoncounter when the fringe area, namely the wall of the hole, is being scanned, so that the signals from the wall do not reach the detection circuit. It opens the gate when the bottom of the hole is being scanned, so that the signals from the bottom of the hole do reach the detection circuit. As a result, only the photoemission signals from the bottom of the hole are detected. We call the ratio of the detection width to the scanning width the "gate-opening ratio."

A diagram of our signal-blanking circuit is shown in Fig. 9. Signals from a scan generator pass through two timing circuits (A and B), which determine the width and timing of the required gate signal. An AND circuit then combines the signals to form the gate signal, which is applied between point P and point Q, so that the signals from the photomultiplier are cut at the fringe of the scan area.

Figure 10 shows how the blanking circuit works. The lower gate signal is generated at point R and a sine wave is

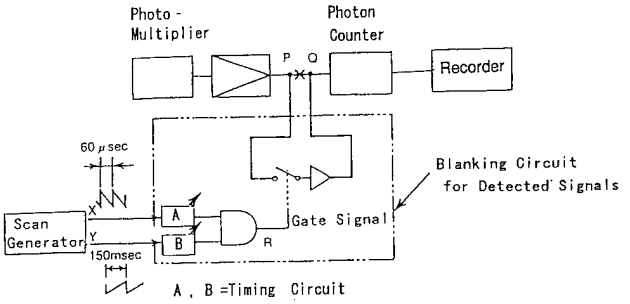


FIG. 9. Blanking circuit for detected signals.

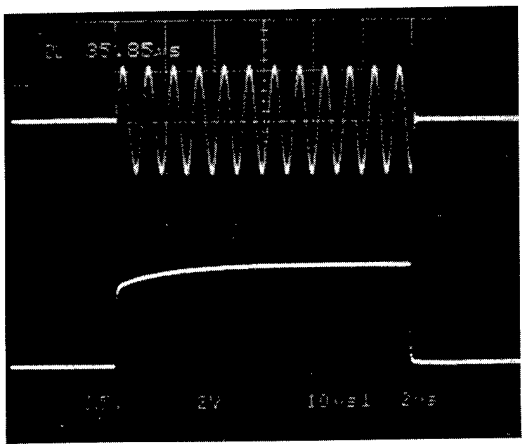


FIG. 10. Outputs from blanking circuit for detected signals. Lower: gated signal generated at point R with gating time 12 μs. Upper: gated sine wave observed at point Q when a continuous sine wave is applied at point P.

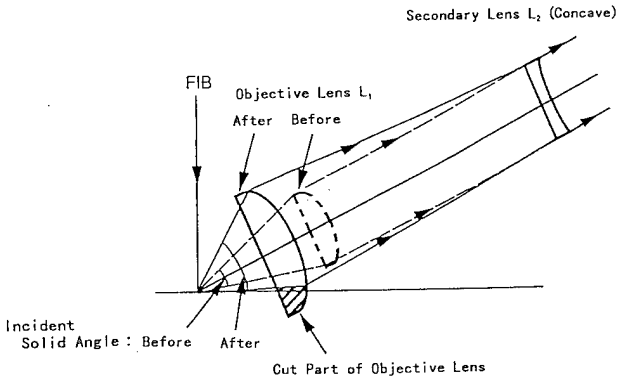


FIG. 11. Increasing of the incident solid angle of objective lens.

TABLE I. Specification of improvements in detection optics.

	Objective lens L_1		Secondary lens L_2		
	Effective diameter	Focal length	Effective diameter	Focal length	Incident solid angle (relative value)
Before	14.0 mm	33.0 mm	···	···	1
After	19.0 mm	23.7 mm	14.0 mm	−208.3 mm	2.6

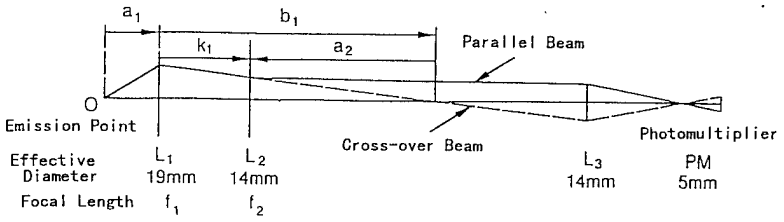


FIG. 12. Optical ray trajectories.

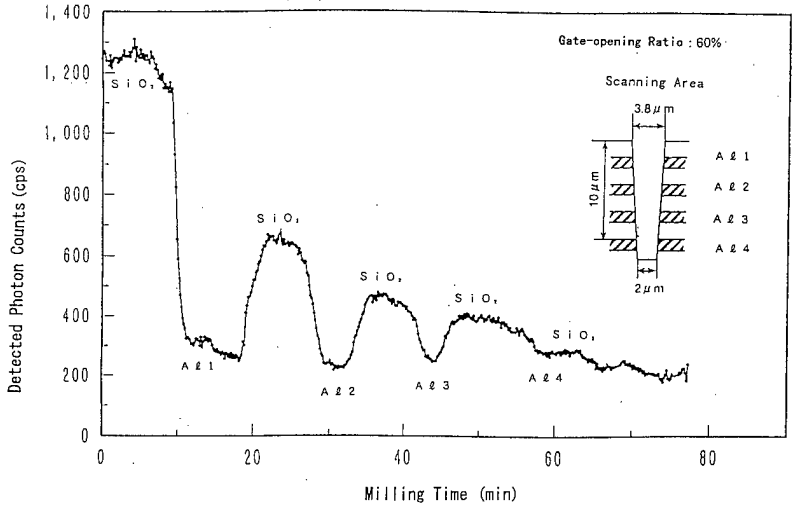


FIG. 13. Relation between milling time and detected photon counts after improvements with gate opening ratio of 60%.

applied at point P. The upper gate signal is the gated sine wave observed at point Q; it shows that the circuit is working satisfactorily.

VI. IMPROVED DETECTION SYSTEM

The photoemission signal itself is not so strong enough because the incident solid angle of the objective lens is too small (0.121 sr). By changing the design of the objective optical system, we have increased the incident solid angle. Figure 11 shows the detection optics close to the sample. The dotted line shows the previous optics when the incident solid angle was 0.121 sr. To enlarge the optics as much as possible, the objective lens was redesigned to have larger diameter and lie closer to the sample. The part of the lens interfering with the sample was removed. A concave lens was inserted to make the beam smaller and to collimate it, so that the beam diameter was the same as before passing this lens.

Optical ray calculation was used to determine the lens dimensions so that wavelengths between 250 and 400 nm for silicon and aluminum emissions are easily introduced into the effective diameter of the photomultiplier. To make this determination, both a collimated and a crossover beam system were investigated. As shown in Fig. 12, in the crossover beam system, when the distance between the sample and the position of the lens was varied, the image point changed greatly, while it was stable in the collimated beam system. We have thus adopted a collimated beam system. The effects of making these improvements are shown in Table I. The incident solid angle of the optics was increased 2.6 times.

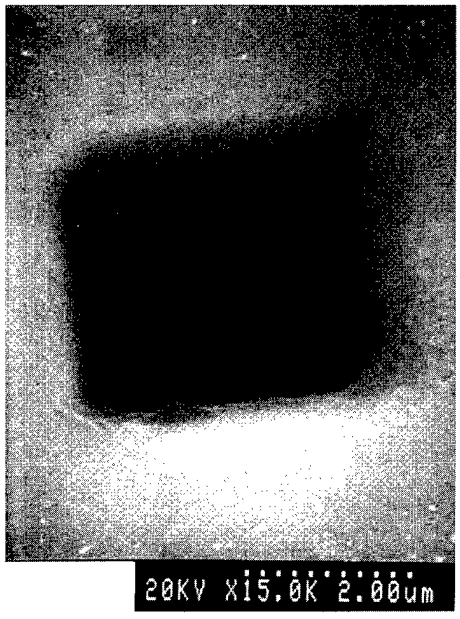


FIG. 14. Scanning electron micrograph of the holes corresponding to Fig. 13 with the multiplication 15 000 times.

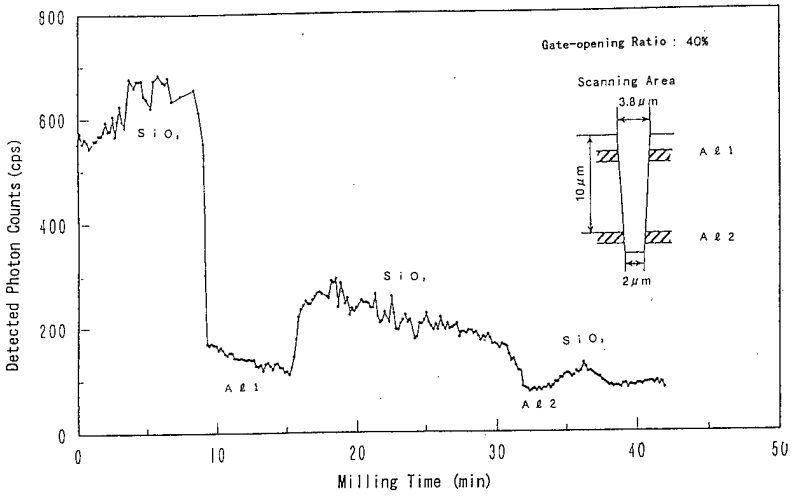


FIG. 15. Relation between milling time and detected photon counts after improvements with gate opening ratio of 40%.

VII. RESULTS OF END-POINT DETECTION

By increasing the beam limiting aperture in the FIB optics, the beam current increased from 40 to 120 pA; the beam size increased from 40 to 80 nm, which is still small enough for milling a small hole in LSI shown in Fig. 2(B). With these improvements, we milled a hole in a LSI chip with four conductive layers having cross-section detection photoemission counts as shown in Fig. 13. We used a current of 120 pA, and a gate-opening ratio of 60%. Compared with the previous measured photon counts data (Fig. 7), the photon count from the SiO_2 and aluminum layers was larger, and we were able to discriminate the fourth aluminum layer, 10 μm deep from the surface. Figure 14 shows a scanning electron micrograph of the holes corresponding to the Fig. 13 with multiplication of 15 thousand times.

However, the signal from the fourth aluminum layer was still a bit obscure when compared with that from the upper layers. This is probably because the gate-opening ratio was too large, causing the photoemission signal from the wall to be mixed with that from the bottom. We thus repeated the experiment using a gate-opening ratio of 40%. As shown in Fig. 15, the fourth aluminum layer was clearly discriminated.

Table II summarizes the parameter changes and maximum detected counts (from Figs. 7 and 15). Here the detected counts means average photon counts from the first SiO_2 layer. And the detected counts without gating equals the detected counts divided by the square of the gate-opening ratio

$$C=A/B^2.$$

TABLE II. Parameter changes with improvements $C=A/B^2$.

	Before	After	Ratio
Detected counts: A	700 cps	650 cps	
Gate opening ratio: B	1	0.4	
Detected counts without gating: C	700 cps	4060 cps	5.8
FIB current	40 pA	120 pA	3
Contribution of improvements in optics			1.93

As a result, the net detection increased about 5.8 times. Because the FIB current increased 3 times, the calculated contribution of the improved detection optics is 1.93. This is 74% of the designed change value 2.6, and shows the improvements realized satisfactory results.

VIII. CONCLUSION

We have developed a method for detecting end points during milling a deep small hole in LSI by FIB. The method uses photon counting detection for photoemission from the atoms excited by the impact of the ion beam. Experiments for LSI chips with alternative four layers of SiO_2, aluminum, and silicon showed that detection is impaired by noise coming from the wall of the hole. This problem was overcome by using a circuit to blank the noise coming from the wall. To make the signal stronger, we increased the incident solid angle 2.6 times by changing the objective optical system. The change in the detected signals coming from the bottom aluminum and SiO_2 layer can now be clearly detected, resulting in effective depth monitoring.

ACKNOWLEDGMENTS

The authors thank Dr. F. Itoh, A. Shimase, Y. Hamamura, J. Azuma, and Dr. T. Ishitani for helpful insights into FIB, Dr. T. Miyauchi and M. Hongo for discussion on optics, Y. Arima and T. Kamimura for fabricating the detection system, T. Yamada (all the above with Hitachi) and K. Otsuka (with Hitachi ULSI Engineering Corp.) for discussion on LSI milling and providing the LSI samples. They are grateful to Professor R. Shimizu (with Osaka University) for explaining the fundamentals of SCANIIR. They also acknowledge Professor J. Orloff (with the University of Maryland), Dr. L. Swanson and Dr. D. Tuggle (with FEI) for helpful insights in FIB.

[1]H. Yamaguchi, A. Shimase, S. Haraichi, M. Usami, and T. Takahashi, Nucl. Instrum. Methods Phys. Res. B **37/38**, 891 (1987).
[2]J. Melngailis, C. R. Musil, E. H. Stevens, M. Utlaut, E. M. Kellog, R. T. Post, M. W. Geis, and R. W. Mountain, J. Vac. Sci. Technol. B **4**, 176 (1986).
[3]L. R. Harriot, Appl. Surf. Sci. **36**, 432 (1989).
[4]F. Itoh, A. Shimase, S. Haraichi, and T. Takahashi, Int. J. Jpn. Soc. Precis. Eng. **27**, 209 (1993).
[5]B. W. Ward, M. Ward, D. Edwards, and N. P. Economou, J. Vac. Sci. Technol. B **6**, 2100 (1988).
[6]F. Itoh, A. Shimase, and S. Haraichi, J. Vac. Sci. Technol. B **9**, 2692 (1991).
[7]C. W. White, Surf. Sci. **49**, 657 (1975).
[8]C. A. Andersen and J. R. Hinthorne, Anal. Chem. **45**, 1421 (1973).
[9]H. Yamaguchi, K. Saitoh, and K. Ishida, Scanning Microscopy Meeting, May 1991, Bethesda, MD, Scanning Microsc. (to be published).

Focused ion beam implantation for opto- and microelectronic devices

H. König, N. Mais, E. Höfling, J. P. Reithmaier,[a] and A. Forchel
Technische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

H. Müssig and H. Brugger
Daimler Benz AG, Research Center, D-89013 Ulm, Germany

(Received 19 November 1997; accepted 23 February 1998)

Focused ion beam implantation is a powerful technology for the fabrication of opto- and microelectronic devices. Optoelectronic devices like gain coupled distributed feedback lasers and nonabsorbing waveguides can be defined in semiconductor heterostructures by the band gap shift due to highly spatially resolved implantation induced thermal intermixing. Single mode emitting devices were fabricated with emission wavelengths of 1 and 1.55 μm in the material systems GaInAs/(Al)GaAs and GaInAsP/InP, respectively. Band gap shifts of more than 65 meV could be reached in GaInAsP quantum film structures which simplifies the integration of nonabsorbing waveguide sections with, e.g., lasers, modulators, and detectors. In highly doped semiconductor layers semi-insulating areas could be defined by focused ion implantation. Depletion lengths down to 50 nm can be controlled and were demonstrated on current injection restricted resonant tunneling devices. By using this technique collector-up heterobipolar transistors were fabricated which exhibit current amplification factors up to 45. © *1998 American Vacuum Society.*
[S0734-211X(98)01204-9]

I. INTRODUCTION

For future optoelectronic devices, e.g., for high speed optical communication systems, an integration of lasers, detectors, waveguides, and possibly driver electronics is necessary. High speed lasers for optical communication are usually based on distributed feedback lasers where a suitable refractive index grating is used to obtain a dynamically stable single mode emission. In order to obtain single longitudinal mode devices, however, a phase shift has to be included into the grating and a precise antireflection coating must be applied. Distributed feedback lasers based on gain coupling in contrast provide single mode emission without phase shifts and coating. The technology used to fabricate these lasers up to now has been mainly based on lithography, etching of the active material, and overgrowth. Here the main problems of the processing include, e.g., the removal of contamination due to the lithography as well as overgrowth[1–4] of faceted samples.

This article gives an overview how focused ion beam technology can be used to simplify the fabrication process for opto- and microelectronic devices. For device processing focused ion beam technologies are promising because they permit maskless patterning and therefore the influence of contamination can be avoided. Furthermore the planarity of the wafer surface is unchanged by the implantation. In combination with molecular beam epitaxy (MBE) ultrahigh vacuum (UHV)-type focused ion beam (FIB) systems may be used for *in situ* patterning processes.[5]

Two main effects can be used in a focused ion beam process to modify the material properties for device applications as it is illustrated in Fig. 1. (I) Implantation induced thermal intermixing in quantum well (QW) structures can be used to increase selectively the band gap [Fig. 1(a)].[6] This effect can be applied to nonabsorbing waveguide areas as well as to lateral carrier confinement[7] and can be used for the integration of waveguide areas with lasers as schematically shown in Fig. 1(b). Due to the high spatial resolution first order gratings can be fabricated which work as gain coupled distributed feedback gratings in single mode emitting lasers.[8–10] (II) With a selective implantation highly resistive areas can be created, as illustrated in Fig. 1(c), for well defined current paths in electronic devices[11] or in-plane gated structures can be defined for transistor and diode functions.[12,13]

II. SYSTEM PERFORMANCE

In the microfabrication facility of Würzburg University two UHV FIB systems fabricated by Eiko Engineering, Nakaminato, Japan, are in use. Both systems are equipped with full lithographic capabilities including correction facilities like, e.g., a laser interferometer with a resolution of 2.5 nm which allows an accurate stitching of several writing fields. Two field sizes, 50×50 and 200×200 μm^2, are available. By using the retarding function of the systems, very low beam diameters with full width at half maximum (FWHM) values between 20 and 30 nm can be achieved and maintained down to ion energies of 10 keV. One FIB system is additionally equipped with a postacceleration setup which allows a maximum energy for single charged ions of 150 keV. The second system is integrated in an UHV processing system including a molecular beam epitaxy system, processing chambers, and a high resolution electron beam patterning system. The maximum acceleration voltage of this system is limited to 70 kV but allows *in situ* gas injection during ion implantation. Both systems show a well-defined beam profile with a Gaussian shape over more than four orders of

[a]Electronic mail: jpreith@physik.uni-wuerzburg.de

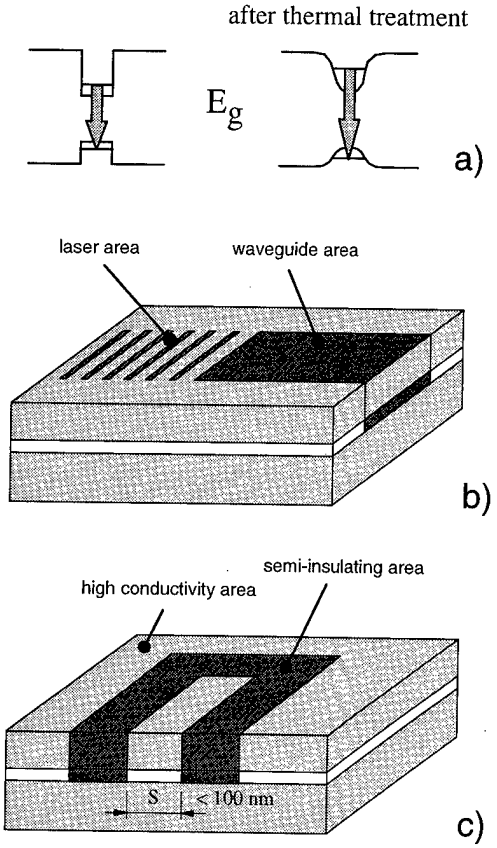


FIG. 1. (a) Principle of band gap shift by implantation induced thermal intermixing in quantum film structures. (b) Schematic design for the integration of optoelectronic devices of, e.g., lasers and waveguide areas. (c) Current path definition in highly doped semiconductors for microelectronic devices.

magnitude[14] and a weak exponential tailing below as background. Due to the low background the spatial resolution is mainly determined by the beam diameter and allows the definition of short period gratings also by maskless patterning. In both systems Ga^+ ions were used which can penetrate more than 300 nm in depth in single crystal semiconductors for ion energies of around 100 keV. This is more than three times larger than expected for penetration depths in amorphous materials and is caused by channeling effects in the [100] oriented semiconductors.[11,15]

III. HIGHLY SPATIALLY RESOLVED LATERAL BAND GAP ENGINEERING

By using implantation induced intermixing a shift in the emission energy to larger band gaps can be achieved. In Fig. 2 a sequence of spectra is shown from a Ga^+ FIB implanted GaInAsP/InP sample. Each photoluminescence (PL) spectrum belongs to an area on the sample implanted with a different ion dose. The sample was annealed at 750 °C by a rapid thermal annealing step for 60 s. Up to a dose of about 2×10^{13} cm^{-2} there is a strong logarithmic dependence between dose and energy shift while the line shape does not change drastically and the intensity remains nearly constant in comparison to the nonimplanted reference spectrum (bottom spectrum in Fig. 2). Within a dose range of about two orders of magnitude the band gap can be controlled very reproducibly over an energy range of more than 65 meV. To demonstrate the high spatial resolution of this maskless patterning technology the influence of ion bombardment in the surrounding of an unimplanted area was investigated by photoluminescence as it is schematically shown in the inset of Fig. 3. In Fig. 3 the PL energy shift is plotted against the width W of the unimplanted area. By changing the width the

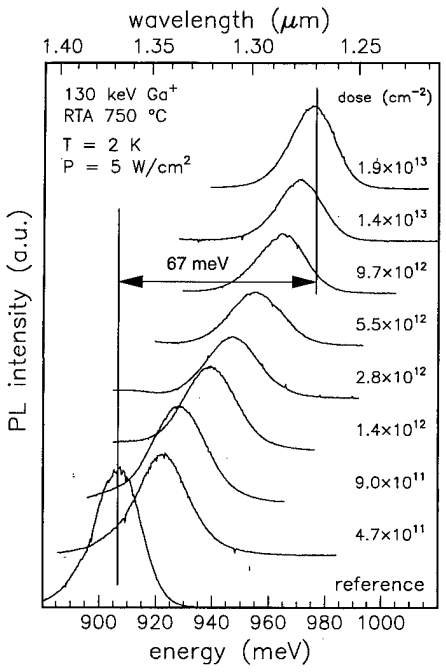


FIG. 2. PL spectra of a FIB implanted GaInAsP quantum well structure for different implantation doses. The bottom spectrum belongs to an unimplanted area.

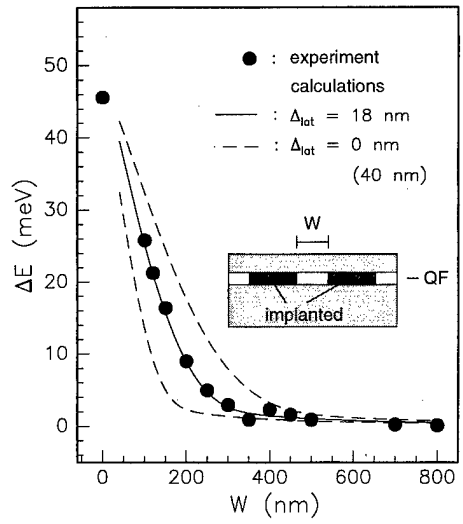


FIG. 3. Shift of the PL energy of a GaAs/AlGaAs quantum well from an unimplanted area surrounded by implanted mesa structures as function of the width W of the unimplanted area. The solid line is calculated for a lateral straggling length of 18 nm while the dashed lines correspond to a straggling length of 40 nm and without straggling, respectively.

PL energy remains the same down to about 200 nm. Only below that value a significant shift can be detected to straggled ions and due to the beginning of an overlap of the beam profiles. The solid line in Fig. 3 is the theoretically expected behavior by assuming a straggling length of 18 nm in a depth of 50 nm, i.e., location of the quantum well, in comparison to 0 and 40 nm straggling length (dashed lines in Fig. 3). Due to this high spatial resolution the definition of band gap modulated gratings with periods below 100 nm are possible and were already demonstrated in blue emitting devices.[16]

IV. GAIN COUPLED DFB LASERS

By a periodic line implantation a gain coupling effect can be produced by a modulation of the band gap. These gain coupled distributed feedback (DFB) lasers have several advantages in comparison to index coupled devices. They show single mode behavior without any phase shift in the grating and without the need for sophisticated antireflection coatings.[17–19] Also a reduced frequency chirp is expected[20] as well as a higher side mode suppression and a reduced spatial hole burning effect[21,22] in comparison to an index coupled structure. Very important for optical communication systems is the back reflection sensitivity. Gain coupled structures are insensitive to back reflection even without any antireflection coating and are therefore very suitable for device integration.[19]

We have fabricated gain coupled GaInAs/AlGaAs DFB lasers emitting at a wavelength of about 1 μm by maskless focused ion beam implantation. The grating was defined after a growth stop just after finishing the active waveguide region of a separate confinement heterostructure (SCH). Due to the preservation of the surface planarity and avoiding any contamination during the grating patterning process the necessary epitaxial overgrowth could be easily done. After an integrated *in situ* thermal annealing step of 2 min at 750 °C prior to the growth process the upper AlGaAs cladding layer was grown. In Fig. 4 room temperature single mode emission spectra from broad area lasers (15 μm stripe width and 800 μm resonator length) were shown for different grating periods varying from 136 to 144 nm.[8] All these devices were defined on the same sample and demonstrate the proposed high single mode yield by using gain coupled gratings.[23]

For future high speed wavelength division multiplexing optical communications systems lasers with well controlled emission wavelengths are important. By using a sampled grating technique the grating periods can be tuned very precisely. In the top part of Fig. 5 this sampled grating technique is illustrated which is based on the combination of at least two subgratings with slightly different periods. This technique makes it possible to overcome the limitations of the addressing grid resolution of the positioning system which in our case is 1 nm. By varying the repetition rate for each subgrating much finer steps for the grating period can be realized. In Fig. 5 the emission wavelengths of gain coupled lasers are plotted as a function of the averaged grating period. The averaged grating period was varied in steps

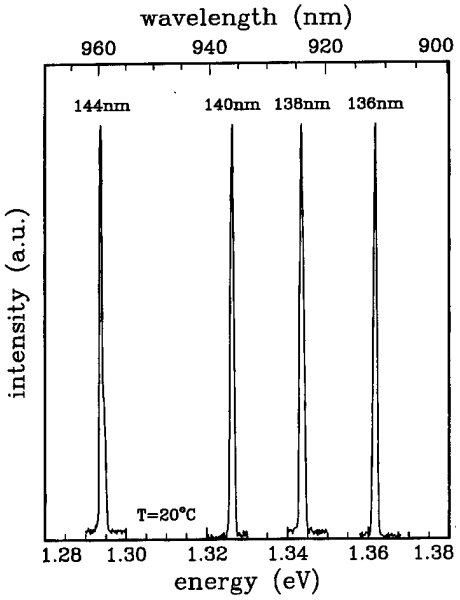


FIG. 4. Emission spectra from GaInAs/AlGaAs gain coupled DFB laser devices for different grating periods defined on the same sample.

of 0.1 nm which tunes the emission wavelength in steps of only 0.4 nm. Due to the linearity between the grating period and emission wavelength the resolution can be even further improved. This accuracy is high enough to fulfill the demands for future high dense multiwavelength communication systems.

A disadvantage of direct implanted gratings into the active region of lasers is residual nonradiative recombination centers due to insufficient annealing of created defects. This

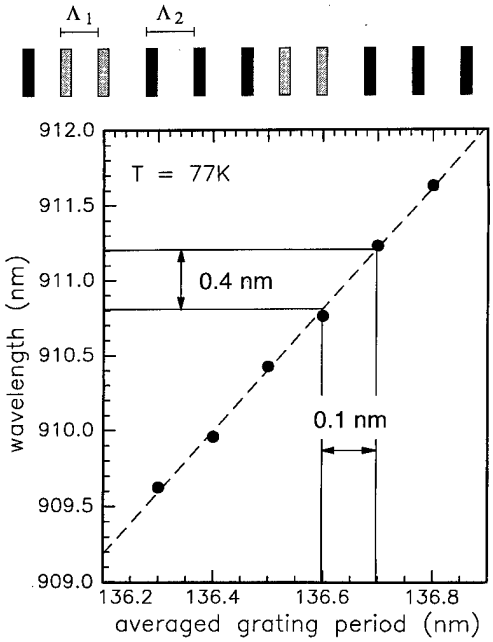


FIG. 5. Correlation between emission wavelength and averaged grating period (solid dots) by using the sampled grating technique as illustrated on top by the combination of two different grating periods Λ_1 and Λ_2. The dashed line is a linear fit to the data points.

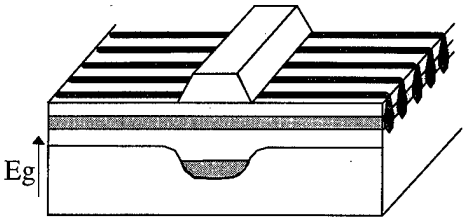


FIG. 6. Schematic drawing for the definition of lateral gain coupled gratings after the fabrication of a ridge waveguide structure. The band gap shift across the ridge geometry and the carrier localization underneath the ridge is indicated.

effect may increase the losses and causes higher threshold currents. To avoid this problem a modified approach was used to define gratings after processing a ridge waveguide structure. In Fig. 6 the process is schematically illustrated. After etching the ridge waveguide structure by dry etching the grating is implanted with Ga ions (100 keV) which have a penetration length of 200–300 nm. The active core region of the laser is protected by the ridge (>1.5 μm in height). The intermixing takes place only lateral to the ridge and forms a grating due to a reduction of the band edge absorption at the emission wavelength. Laterally gain coupled devices were fabricated by this technique on GaInAsP/InP ridge waveguide laser structures. The devices show single mode emission at room temperature at a wavelength of 1.54 μm according to a grating period of 236 nm.[24] Additionally to the reduction of the band edge absorption an improved lateral carrier confinement is formed by the band gap enlargement. The variation of the conduction band energy E_c across the ridge is indicated in Fig. 6 below the QW layer. The band gap is higher on both sides of the ridge and localizes the carrier underneath the ridge. Nonradiative carrier losses outside the injection area can be avoided and the quantum efficiency should be improved. In Fig. 7 the laser output power of different laser devices (solid lines), which are laterally implanted, are plotted in comparison to an unimplanted reference laser. The laterally implanted devices show a strongly improved quantum efficiency and a reduced threshold current and allow continuous mode operation at room temperature.

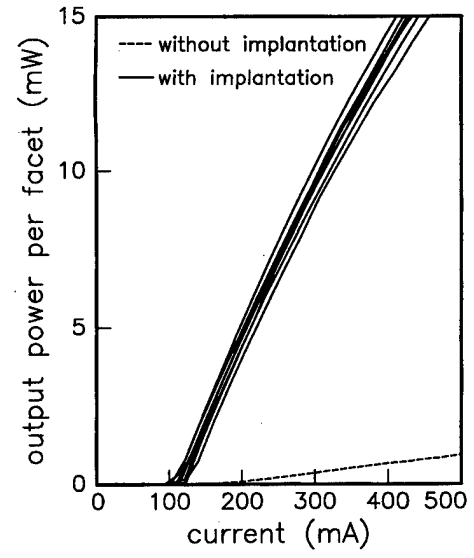


FIG. 7. Light output characteristics of laterally implanted GaInAsP/InP ridge waveguide lasers (solid lines) measured in pulsed mode at room temperature in comparison to an unimplanted laser device from the same sample.

V. BURIED CURRENT PATH RESTRICTION

In highly doped semiconductor material the conductivity can be drastically reduced by ion implantation. At implantation doses of 1×10^{13} cm^{-2} (Ga^+ ions, 140 keV) semi-insulating properties can be achieved with a sheet resistivity of 3×10^8 $\Omega/\square$ in a 200 μm thick 5×10^{18} cm^{-3} Si-doped GaAs layer. Regrowth experiments show that implantation induced isolation is sustained following a MBE growth process at 550 °C without significant decrease of the sheet resistivity.[11] The spatial resolution was tested by restricting the injection current path in a resonant tunneling diode (RTD). For lateral current restriction a mesa structure was implanted with an unimplanted area in slit geometry at the center. By varying the slit width the depletion length can be determined at the current cutoff width. For a dose of 5×10^{12} cm^{-2} a depletion length of 50 nm was determined which allows the fabrication of well-defined buried current paths in the sub-μm regime.

The application of this technology was demonstrated in a collector-up heterobipolar transistor (C-up HBT). In comparison to a conventional emitter-up HBT the modulation frequency limiting capacity between collector and base is much lower in a C-up HBT due to the small lateral dimensions.[25] Unfortunately in these devices a buried current path restriction is necessary to avoid electron losses outside the collector area. In Fig. 8 a schematic drawing of the used structure is shown. The current restricting isolation structure was defined by FIB implantation after MBE growth of the emitter part. After the maskless patterning process the wafer was transferred back to the MBE system for subsequent overgrowth of the base and collector part. All transfers were done in ultrahigh vacuum conditions at a pressure of better than 2×10^{-8} mbar. In Fig. 8 the output characteristic of a device with a current path width of 1.5 μm and a length of 25 μm is plotted. The evaluated current amplification factor of about 45 at a current density of 5×10^4 A/cm^2 is the highest value ever reported for high frequency designed C-up HBTs.[26]

VI. CONCLUSIONS

The capability of maskless patterning by focused ion beam implantation for opto- and microelectronic device applications was demonstrated. The technique of spatially resolved implantation induced thermal intermixing can be used for defining waveguide structures as well as for gain coupled gratings for DFB lasers in a large variety of semiconductor materials. Single mode emitting laser devices were fabricated in GaInAs/AlGaAs and GaInAsP/InP material systems with emission wavelengths of 1 and 1.55 μm, respectively.

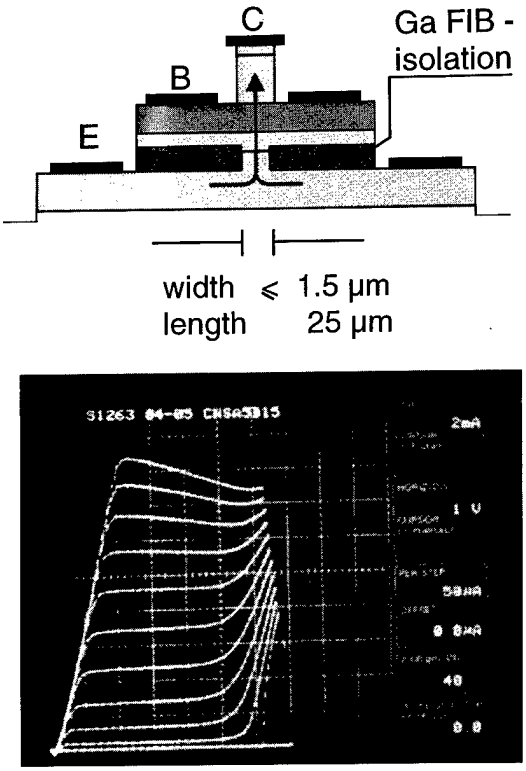


FIG. 8. On top of the schematic drawing of a C-up HBT with a FIB implanted current restriction layer in the emitter part of the device. At the bottom is the device output characteristic of a device with a current path width of 1.5 μm and a length of 25 μm.

Due to the contamination free and planarity conserving patterning technique *in situ* overgrowth processes can be easily attached.

By implantation induced defect creation semi-insulating properties can be created in highly doped semiconductors. A control of the depletion length in the range of 50 nm was achieved which allows the definition of buried current paths in microelectronic devices in the sub-μm range. This technique was applied to a collector-up heterobipolar transistor fabrication process. For high frequency designed devices with a buried current path width of 1.5 μm and a length of 25 μm record current amplification factors of 45 are shown.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support of this work by the European Community through the NANOLASE Project, the BMBF through the Photonik II project and the "Bayerische Forschungsstiftung." The supply of epitaxial material from Alcatel SEL, Stuttgart, and Alcatel Alsthom Recherche, Marcoussis, is gratefully acknowledged.

[1]W. T. Tsang, F. S. Choa, M. C. Wu, K. Chen, A. Logan, S. N. G. Chu, A. M. Sergent, and C. A. Burrus, Appl. Phys. Lett. **60**, 2580 (1992).
[2]Y. Nakano, H. Cao, K. Tada, M. Dobashi, and H. Hosomatsu, Jpn. J. Appl. Phys., Part 1 **32**, 825 (1993).
[3]H. Lu, G. P. Li, and T. Makino, Appl. Phys. Lett. **63**, 589 (1993).
[4]Y. Luo, Y. Nakano, K. Tada, H. Hosomatsu, and H. Iwaoka, Appl. Phys. Lett. **56**, 1620 (1990).
[5]A. Orth, J. P. Reithmaier, F. Faller, and A. Forchel, J. Vac. Sci. Technol. B **13**, 2714 (1995).
[6]J. Oshinowo, J. Dreybrodt, A. Forchel, N. Mestres, J. M. Calleja, I. Gyuro, P. Speier, and E. Zielinski, J. Appl. Phys. **74**, 1983 (1993).
[7]K. Ishida, T. Takamori, K. Matsui, T. Fukunaga, T. Morita, E. Miyauchi, H. Hashimoto, and H. Nakashima, Jpn. J. Appl. Phys., Part 1 **25**, L783 (1986).
[8]A. Orth, J. P. Reithmaier, R. Zeh, H. Doleschel, and A. Forchel, Appl. Phys. Lett. **69**, 1906 (1996).
[9]A. Orth, E. Höfling, R. Zeh, J. P. Reithmaier, A. Forchel, J. Weber, I. Gyuro, and E. Zielinski, presented at the Indium Phosphide and Related Materials Conference in Schwäbisch Gmünd, April 1996, Conference Proceedings, 1996.
[10]E. Höfling, A. Orth, J. P. Reithmaier, and A. Forchel, International Laser Conference, Haifa, Conference Proceedings, 1996.
[11]H. Müssig, T. Hackbarth, H. Brugger, A. Orth, J. P. Reithmaier, and A. Forchel, Mater. Sci. Eng., B **35**, 208 (1995).
[12]Y. Hirayama, Appl. Phys. Lett. **61**, 1667 (1992).
[13]S. Nakata, Phys. Rev. B **46**, 13326 (1992).
[14]A. Kieslich, J. P. Reithmaier, and A. Forchel, J. Vac. Sci. Technol. B **12**, 3518 (1994).
[15]A. Kieslich, J. Straka, and A. Forchel, J. Appl. Phys. **72**, 6014 (1992).
[16]D. Eisert, G. Bacher, N. Mais, J. P. Reithmaier, A. Forchel, B. Jobst, D. Hommel, and G. Landwehr, Appl. Phys. Lett. **68**, 599 (1996).
[17]H. Kogelnik and C. V. Shank, J. Appl. Phys. **43**, 2327 (1972).
[18]Y. Nakano, Y. Luo, and K. Tada, Appl. Phys. Lett. **55**, 1606 (1989).
[19]Y. Nakano, H. Cao, K. Tada, Y. Luo, M. Dobashi, and H. Hosomatsu, Jpn. J. Appl. Phys., Part 1 **32**, 825 (1993).
[20]Y. Luo, R. Takahashi, Y. Nakano, K. Tada, T. Kamiya, H. Hosomatsu, and H. Iwaoka, Appl. Phys. Lett. **59**, 37 (1991).
[21]T. Sudoh, Y. Nakano, K. Tada, K. Kikuchi, and H. Hosomatsu, IEEE Photonics Technol. Lett. **5**, 1276 (1993).
[22]G. Morthier, P. Vankwikelberge, K. David, and R. Baets, IEEE Photonics Technol. Lett. **2**, 170 (1990).
[23]K. David, G. Morthier, P. Vankwikelberge, R. G. Baets, T. Wolf, and B. Borchert, IEEE J. Quantum Electron. **27**, 1714 (1991).
[24]H. König, J. P. Reithmaier, A. Forchel, J. L. Gentner, and L. Goldstein (unpublished).
[25]H. Kroemer, Proc. IEEE **70**, 13 (1982).
[26]S. Yamahata, Y. Matsuoka, and T. Ishibashi, IEEE Electron Device Lett. **14**, 173 (1993).

Direct writing of active loads by focused ion beams

C. Wiemann, M. Versen, and A. D. Wieck[a)]
Lehrstuhl für Angewandte Festkörperphysik, Ruhr-Universität Bochum, Universitätsstrasse 150, D-44780 Bochum, Germany

(Received 17 April 1998; accepted 27 May 1998)

With a 100 nm focus of a focused Ga^+ ion beam with 100 keV, we write insulating lines in electronic layers of $In_{0.21}Ga_{0.79}As$ quantum wells. In this way, in-plane-gate (IPG) transistors are formed which can be operated at room temperature. In a typical integration application of a common source circuit, the pull-up resistance represents a serious problem due to the high geometric aspect ratio necessary for it. For example, the typical specific sheet resistivity of the $In_{0.19}Ga_{0.79}As$ quantum well of 1.2 kΩ needs to be increased to 100 kΩ by a 1 μm wide, about 83 μm long channel. In order to save this waste of area we introduce active loads in the form of a narrow channel. In this way, the pull-up resistor requires orders of magnitude less area and stabilizes the drain current due to velocity saturation, leading to lower supply voltages. Inverters in this technology are presented and characterized. In finite element simulations these circuits are further investigated. The operation of these systems is based on the lateral depletion of adjacent quantum well areas. The basic differences between depletion within *pn* half spaces and *pn* half planes are discussed analytically, showing a marked dependence on dimensionality. In particular, it is shown that the ruggedness of IPGs can be explained by these phenomena. © *1998 American Vacuum Society.* [S0734-211X(98)11404-X]

I. LATERAL FIELD EFFECT DEVICES

In writing a focused-ion-beam (FIB)[1,2] line with Ga^+ ions of 100 keV and a 100 nm focus across a mesa, which contains the charge carrier layer close to the surface, we overcompensate these carriers with an appropriate ion species. This leads to lateral *pn* junctions which divide the former homogeneous charge carrier layer into two electrostatically separated sheets (inset I of Fig. 1).[3] As one possible basic material for our field effect devices we use laterally homogenous $In_{0.21}Ga_{0.79}As$ quantum wells. Using molecular beam epitaxy (MBE) for these we grow an 80 nm undoped buffer on a semi-insulating GaAs (100) wafer, followed by an AlAs/GaAs superlattice with a total thickness of 72 nm and another buffer of 53 nm. On this we grow a pseudomorphic 13 nm $In_{0.21}Ga_{0.79}As$ layer and a 4 nm thick $Al_{0.20}Ga_{0.80}As$ spacer both undoped. Next is a 57 nm Si-2×10^{18} cm^{-3}-doped layer of $Al_{0.20}Ga_{0.80}As$, covered by a 5 nm cap layer of $Al_{0.20}Ga_{0.80}As$. This leads to a high electron density of $n_e=9\times10^{11}$ cm^{-2} and a mobility of $\mu=5500$ cm^2/V s at room temperature in the $In_{0.21}Ga_{0.79}As$ quantum well.

Next these samples are wet chemically etched to define discrete lateral mesa structures. The current–voltage characteristic of the lateral interface at room temperature in inset I of Fig. 1 exhibits vanishing currents at positive and negative bias of several volts, respectively. At a higher bias, reversible breakdown of the junction occurs. In GaAs, the implanted Ga ions induce deep impurity levels or even *p*-type doping,[4] whereas Ga is a reasonable *p*-type dopant in Si.[5] Due to this introduction of deep levels in GaAs, the implanted Ga serves to deplete the electron density without being thermally activated, although thermal activation should be a standard step in semiconductor processing.

In inset II of Fig. 1, a FIB line with a gap in its middle is shown. This gap may be of a typical width of 1 μm, but also ones as small as 200 nm have shown to be conductive.[6] Around zero bias, this constriction shows a clear ohmic behavior and runs into a saturation current for higher bias, which is due to the maximum drift velocity in the semiconductor.

In combining insets I and II of Fig. 1, we obtain inset III. The constriction is tunable by the adjacent in-plane-gate (IPG) and shows current–voltage characteristics like a field-effect transistor.[7] It is worth noting that this transistor concept does not depend on a particular semiconductor material and has up to now been realized in $Al_xGa_{1-x}As$/GaAs heterostructures,[7] $In_xGa_{1-x}As$/GaAs quantum wells,[6] SiGe heterostructures,[8] nonepitaxial silicon-on-insulator,[5] and on standard bulk silicon.[9] For these experiments, FIB techniques have been used as well as homogeneous ion implantation through opened resist windows, deep mesa etching[10] and laser induced local diffusion.[11]

II. ACTIVE LOADS

The application of this lateral field-effect transistor concept seems to be obvious: Since there is no alignment necessary between gate, source and drain, the fabrication of direct FIB writing is a very reliable and simple single process step, even if the throughput is judged to be too small for industrial processes.[12] Another advantage of the IPG principle is its connecting technique in just omitting FIB lines in between components.

For example, a simple inverter can be formed by a transistor in common source configuration and a pull-up resistor between drain and the positive supply voltage as indicated in Fig. 2. This "pull-up" resistor could be made by a long, ohmic path of the two-dimensional electron gas. The resistance would then be the product of the aspect ratio (length/

a)Electronic mail: andreas.wieck@ruhr-uni-bochum.de

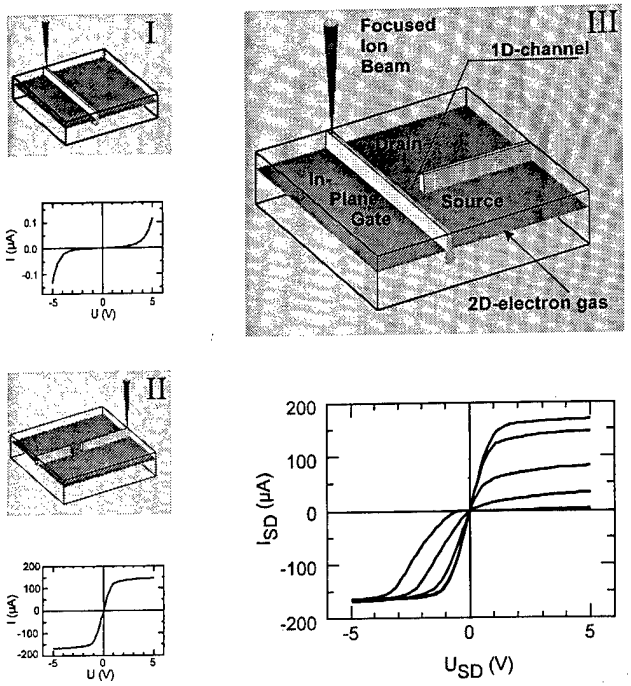


FIG. 1. Evolution of an in-plane-gate transistor's creation: Insulation writing (I), definition of an unipolar channel (II) and buildup of the complete device (III). At each stage, the corresponding current–voltage characteristics are given. In III, the in-plane-gate voltage is a parameter ranging from +1 V (uppermost curve in first quadrant) to −3 V (lowest curve in first quadrant) in steps of 1 V.

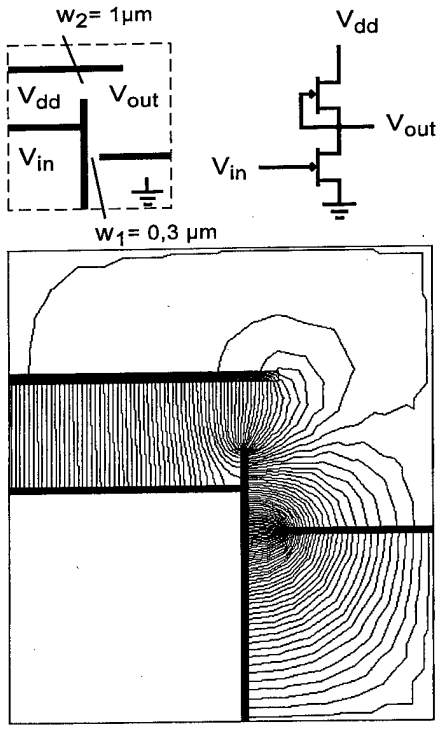


FIG. 2. Inverter circuit, FIB-writing scheme and simulated potential distribution for $V_{dd}=6$ V, $V_{in}=0$ V, $V_{out}=$float, the potential difference between two neighboring lines being $\Delta V=0.05$ V.

width) and the specific two-dimensional resistance which is 1.2 kΩ. This technique of fabricating pull-up resistors is of course very innefficient due to the area waste of such a long channel (typically a 100 kΩ resistor is needed which would need a channel of about 83 μm length and 1 μm width). In order to overcome this problem we adopt the principal of active loads from current integration industry in just forming a second IPG transistor with gate and source connected to each other (circuit diagram of Fig. 2). This has the advantage of very little area needed and moreover the nonohmic behavior due to velocity saturation in this channel gives rise to a constant current behavior which is much better than a simple, ohmic "pull-up" resistor. By the method of finite elements we calculate the potential distribution of such an arrangement which is indicated in Fig. 2. However, this simulation does not include the spatial extent of the depletion zones which will be discussed in detail below. The output characteristic with respect to the input voltage of such a device is shown in Fig. 3.

The active pull-up channel is just current-limiting and therefore always conductive. In the case of depleting the lower transistor the main voltage drop occurs at this device and the output voltage is high. Opening the lower path shifts the main voltage drop to the upper transistor and the output voltage decreases. The switching levels depend on the sheet resistance of our heterostructure, which is directely related to the carrier density and mobility in the quantum well, and the width of the implanted channels. As the implantation times are on the order of milliseconds for each structure, the damage of the conducting areas by an eventual ion beam halo does not affect the electrical parameters of the circuit. Furthermore we have several possibilities to adjust the electrical parameters of the inverter, i.e., by geometrical variation of the writing scheme or shifting the bias. In this way we adjust the different range of input and output voltages of the inverter. Here, the inverter function is clearly demonstrated and gives in combination with logic gates in the IPG technique[13] the base for digital circuits, exclusively fabricated by focused ion beams.

In the preceding discussion, the extent and tunability of depletion zones were not discussed. Since both extent and tunability strongly depend on the dimensionality of the system they are addressed in the next paragraph.

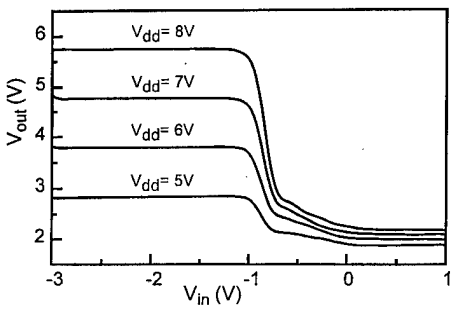


FIG. 3. Drain voltage vs gate voltage of the circuit in Fig. 2.

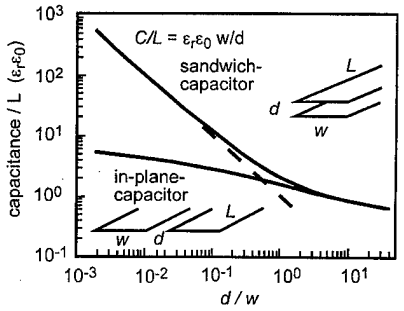


FIG. 4. Capacitance C of a sandwich capacitor and an in-plane capacitor vs d/w as depicted in the corresponding insets normalized to the length L of the stripes and $\epsilon_r\epsilon_0$. The broken line represents the well-known -1 slope for small d/w. For high d/w C/L converges into the coplanar case (solid line).

III. THE EXTENT OF DEPLETION LENGTHS IN IPG TRANSISTORS

In contrast to a sandwich capacitor of a length L, width w and a homogeneous plate spacing d in-plane capacitors of a length L and a lateral spacing d exhibit a substantial lower capacitance C and an inhomogeneous electric field distribution. The dependence $C/\epsilon_0\epsilon_r L$ versus d/w is shown for both cases in a double logarithmic plot in Fig. 4.[14] Of course, the capacitance of a sandwich device is for $w \gg d$ (which means negligible fringing fields) proportional to w/d. But what is the capacitance between two coplanar (in-plane) stripes with widths w and distance d? As it can be seen in the lower curve of Fig. 4 with the corresponding inset, this capacitance decreases by only a factor of 4 if d/w is increased by four orders of magnitude. In other words, the in-plane capacitance is only logarithmically dependent on d/w and yields the approximate value of $C \approx 4\epsilon_0\epsilon_r L$ in typical IPG systems with a mesa size of a few ten μm.[14,15] This indicates that the in-plane capacitance is a fundamental one, given in the unit of the dielectric constant $\epsilon_0\epsilon_r$, which means Farad/meter and it depends just on the length L of an in-plane capacitor. For GaAs, this capacitance amounts to 0.5 fF/μm.

Let us now calculate the width of the depleted gap d first in the case of a sandwich capacitor between two half spaces (charge carrier density N_3) and then of an in-plane capacitor between two half planes (carrier density N_2). Q denotes the accumulated charge on one capacitor plate and V the voltage applied to the capacitor, neglecting the built-in voltage V_{bi} for simplicity. In a more detailed picture, V may be replaced by $V + V_{bi} - 2k_BT/e$. The elementary charge is indicated by e, Boltzmann's constant by k_B, and the absolute temperature by T.

In the case of a sandwich arrangement, $C = Q/V = \epsilon_0\epsilon_r Lw/d$. The half of the charge lacking in the gap must equal the one on a capacitor plate, i.e., $eN_3Lwd/2 = Q$. Combining these two identities, we get

$$d = \sqrt{\frac{2\epsilon_0\epsilon_r V}{eN_3}}$$

for a sandwich capacitor. Analogously, we obtain for the in-plane capacitor $C = Q/V \approx 4\epsilon_0\epsilon_r L$ and $eN_2Ld/2 = Q$, which yields

$$d \approx \frac{8\epsilon_0\epsilon_r V}{eN_2}$$

for an in-plane capacitor in accordance with results of calculations including the magnetic field dependence.[16]

This result is surprising in two ways: First, the extension d of the depleted gap is directly proportional to the applied voltage V and thus IPG channels can be tuned linearly in width by the gate voltage.

Second, the electric field strength E within the capacitor gap is approximately constant, i.e., $E = V/d \approx eN_2/8\epsilon_0\epsilon_r$. For a typical carrier density of $N_2 = 10^{12}$ cm^{-2} and a relative dielectric constant of $\epsilon_r = 13.1$, this yields $E = 1.7 \times 10^4$ V/cm. The inverse of this value amounts to about 580 nm/V which is a realistic depletion coefficient since 1 μm wide IPG channels are depleted at gate voltages of typically -2 V.[17]

The fact that the electric field is almost *independent of the applied voltage* suggests that breakdown voltages are not reached as easily as in sandwich diodes, where the electric field increases with $\sqrt{V}$, which could explain the experimentally experienced ruggedness of the IPGs. This is an important advantage of the IPG technology and may be exploited in different semiconductor materials.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support within the DFG-Graduiertenkolleg 384 and the expert help of Rolf Wernhardt and Georg Kortenbruck.

[1] J. Melngailis, J. Vac. Sci. Technol. B **5**, 469 (1987).
[2] J. Orloff, Rev. Sci. Instrum. **64**, 1105 (1993).
[3] R. D. Tscheuschner and A. D. Wieck, Superlattices Microstruct. **20**, 615 (1996).
[4] Y. Hirayama and H. Okamoto, Jpn. J. Appl. Phys., Part 2 **24**, L985 (1985).
[5] C. Crell, C. Wieczorek, H.-U. Schreiber, and A. D. Wieck, Appl. Phys. Lett. **68**, 2538 (1996).
[6] A. D. Wieck and K. Ploog, Appl. Phys. Lett. **61**, 1048 (1992).
[7] A. D. Wieck and K. Ploog, Appl. Phys. Lett. **56**, 928 (1990).
[8] D. Többen, D. K. de Vries, A. D. Wieck, M. Holzmann, G. Abstreiter, and F. Schäffler, Appl. Phys. Lett. **67**, 1579 (1995).
[9] C. Crell, S. Friedrich, H.-U. Schreiber, and A. D. Wieck, J. Appl. Phys. **82**, 4616 (1997).
[10] J. Nieder, A. D. Wieck, P. Grambow, H. Lage, D. Heitmann, K. v. Klitzing, and K. Ploog, Appl. Phys. Lett. **57**, 2695 (1990).
[11] P. Baumgartner, K. Brunner, G. Abstreiter, G. Böhm, G. Tränkle, and G. Weimann, Appl. Phys. Lett. **64**, 592 (1994).
[12] W. L. Brown and A. Wagner, Proceedings of the International Ion Engineering Congress, ISIAT'83 & IPAT'83, p. 1738 A.
[13] U.S. Patent No. 5,338,692 (1994).
[14] D. K. de Vries, P. Stelmaszyk, and A. D. Wieck, J. Appl. Phys. **79**, 8087 (1996).
[15] D. K. de Vries and A. D. Wieck, J. Vac. Sci. Technol. B **13**, 394 (1995).
[16] D. B. Chklovskii, B. I. Shklovskii, and L. I. Glazman, Phys. Rev. B **46**, 4026 (1992).
[17] M. A. Armstrong, S. Etchin, J. Melngailis, and D. A. Antoniadis, J. Appl. Phys. **78**, 560 (1995).

In situ scanning tunneling microscope studies of high-energy, focused ion implantation of Ga into GaAs: Direct observation of ion beam profiles

G. A. C. Jones,[a] P. D. Rose, and S. Brown
Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom

(Received 17 April 1998; accepted 28 May 1998)

The profile of an ion beam line exposure of Ga^+ ions into GaAs(100) was characterized *in situ* using scanning probe microscopy. Current imaging tunneling spectroscopy was used to characterize the surface defects induced by the high-energy Ga^+ ions. Spatially reproducible features, approximately one per ion and 2–3 nm in diameter, were observed on the irradiated surface. Differential conductance spectra of these features indicated that the image contrast was due to acceptor states induced in the surface band gap. The density of these defects was used to form a profile of the ion beam in the semiconductor surface. The resulting profile was in excellent agreement with the two-Gaussian fit reported in previous work. © *1998 American Vacuum Society.*
[S0734-211X(98)11504-4]

I. INTRODUCTION

The use of high-energy focused ion beams (FIBs) for maskless patterning of semiconductor structures is a well established technique.[1]

Figure 1 shows the effect of 30 keV Ga^+ ions on the electrical resistivity of a 50 nm thick n^+ GaAs layer. When this technology is combined in situ with molecular beam epitaxy (MBE) the layer may be buried within the crystal to produce a patterned back-gated device.[2] Of particular interest is the definition of sub-micron structures using this technique,[3] thus a method to accurately characterize the effective ion beam resolution *in the semiconductor surface* is highly desirable.

Previous work by Ben Assayag *et al.*[4] used transmission electron microscopy (TEM) to image an amorphous regions in GaAs and GaAs/AlGaAs heterostructures. Modeling their results, they showed that the ion beam had a Gaussian profile at low probe current, i.e., with a small beam limiting aperture in the column. At high currents, a bi-Gaussian distribution was observed, which they thought could be correlated with the ion optics. Lateral spreading of the ion beam damage was also examined by Bever *et al.*[5] Spatially resolved cathodoluminescence spectra and electron transport measurements showed that the damage extended considerably outside the implanted regions. They reported a 'tail' dose, which decreased exponentially over a range of a few microns. Yamamoto *et al.*[6] used transport measurements and the electron focusing effect to estimate the ballistic length and mean free path of electrons in a channel defined by FIB. They reported scattering up to 4 μm from the FIB irradiated line, whereas the FIB spot size for their experiments was 0.1 μm.

II. EXPERIMENT

This work was performed using a unique ultrahigh vacuum (UHV) MBE/FIB/scanning tunneling microscope (STM) facility, which allows in situ transfer of samples under UHV. The layer growth was carried out in Vacuum Generators V80H MBE growth chambers. Commercially prepared undoped GaAs(100) substrates were first out-gassed at 450 °C for 1 h in a preparation chamber, prior to being loaded into the growth chamber and a further out-gassing process at 620 °C under an As flux for 20 min. Next, a 1 μm layer of Si-doped GaAs with a nominal carrier concentration of 1×10^{24} m^{-3} was grown on the substrate wafer at a temperature of 600 °C, measured using an optical pyrometer.

The wafer was immediately transferred in situ to the UHV FIB implantation chamber, via UHV transfer tubes. The FIB column used for these experiments was a twin-lens system with a Ga^+ liquid metal ion source and had an ultimate spot size of 50 nm. The source was operated at a potential of 30 kV with an extraction current of 2 μA. A selectable aperture was used to set the sample current at 50 pA. The 30 keV Ga^+ ions impinged normal to the wafer surface and the full width half maximum spot size was estimated from secondary electron images to be approximately 200 nm. Under computer control, the ion beam was scanned across the sample surface to form a variably spaced line pattern with a line dose of 1×10^{10} ions m^{-1}. Assuming a spot size of 200 nm, this is equivalent to a peak areal dose of $\sim1\times10^{17}$ ions m^{-2}, which is typical of the dose required for electrical isolation of 50 nm n^+ layers.

The patterning procedure took less than 20 min, after which the wafer was transferred immediately to the STM chamber. The pressure of the STM, FIB and transfer chambers was $<10^{-7}$ Pa throughout the experiments, thus no significant surface contamination occurred. This was confirmed by atomic resolution STM images on all wafers. Typically, topographic STM images were formed using a constant current of 0.1 nA with a sample bias ranging from +1 to +2 V.

III. RESULTS AND DISCUSSION

A typical low resolution topographic image of a 7 μm×7 μm area of the ion beam irradiated surface is

[a]Electronic mail: gaj1@cam.ac.uk

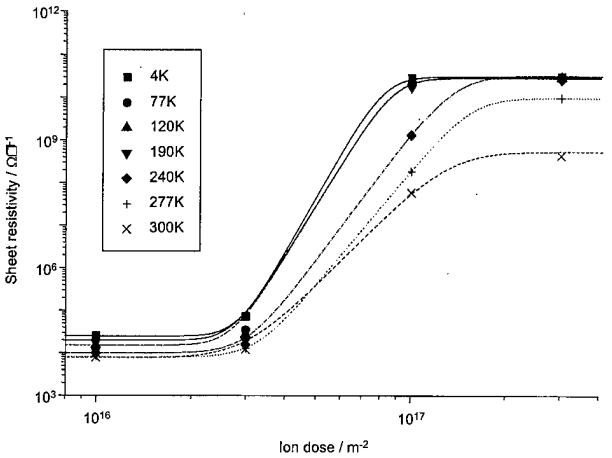


FIG. 1. Electrical resistivity of a 50 nm n^+ GaAs layer vs ion dose.

shown in Fig. 2. This sample was ion beam patterned with a set of parallel lines, with alternate 1 and 2 μm spacing. Faint contrast due to the FIB irradiation is observed in this large-scale image of the sample surface. Note that the ion beam lines do not appear parallel in the STM image, as at large deflections the piezo-electric scan of the STM tip is nonlinear. As expected, at this ion dose, large-scale sputtering did not occur as the growth islands were preserved on the surface.

Figure 3 shows the corresponding "current" image. This is formed by recording the error signal in the STM feedback loop, between the preset and actual tip-sample current, while the topographic image is recorded. Strong contrast is observed between the ion beam irradiated and virgin regions of the wafer surface. This contrast is a function of the tip scanning speed and the gain of the STM tip feedback loop. In addition, the island surface features are not observed in this image. Thus the ion beam irradiation causes a change in the surface tunneling conditions, rather than a gross topographic change.

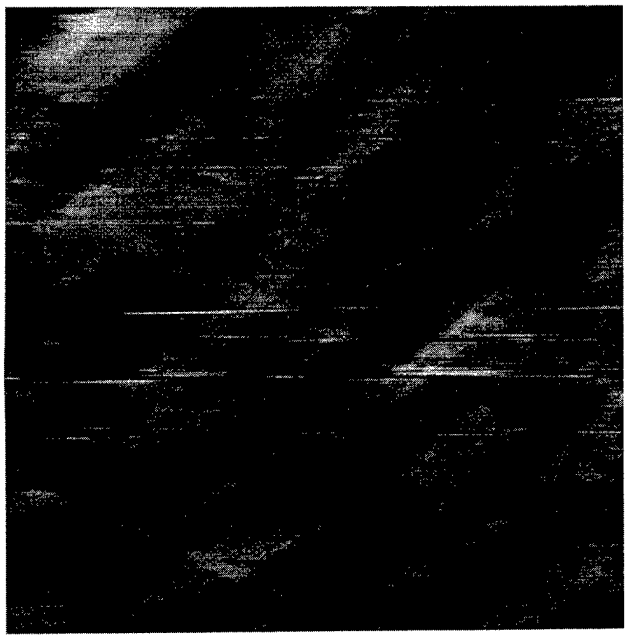


FIG. 2. 7×7 μm STM topographic image of ion beam lines, which can be faintly observed running from the top left to bottom right. Striations running from the top right to the bottom left of the picture are growth islands orientated perpendicular to the [011] direction and typical for MBE grown GaAs(100). Separation of the ion beam lines was alternately 1 and 2 μm.

FIG. 3. STM "current" image, taken concurrently with the topographic image shown in Fig. 2.

Atomic resolution images of the virgin and irradiated surface were taken on all the wafers. An atomic resolution image taken at the center of the ion beam implanted line is shown in Fig. 4. This shows regions, ~2 nm in size, where the top monolayer of atoms has been removed. Further, small clusters of atoms are visible on the surface. Line scans were performed to confirm the monolayer step height of approximately 0.15 nm. Away from the ion beam damaged area, the surface images did not show these defects. The top monolayer was identified by the (2×4) surface reconstruction, which was the surface reconstruction in the unirradiated material. The regions where the underlying monolayer was exposed showed the $c(4\times4)$ reconstruction. This suggests that regions of the top monolayer of atoms were sputtered by the high-energy ions.

The number of these sputtered regions in the image is difficult to quantify, as they may coalesce to form larger areas of damage, however around 50 distinct regions in the image can be observed. Assuming a spot size of ~200 nm, the areal dose in the center of the ion beam line is $\sim 10^{17}$ ions m^{-2}, thus the expected number of ions incident in this 40×40 nm area is ~150. The number density of these sputtered regions, approximately 2 nm in size, is thus of the same order as the ion dose. Larger regions are either formed by co-incident impacts of two or more ions, or formed by a cluster of ions.

Current-imaging tunneling spectroscopy (CITS) was used

FIG. 4. Atomic scale image of the ion beam irradiated region. Image size is 40×40 nm.

to further characterize the defects in the surface, induced by ion beam irradiation. In the CITS mode, the tip is raster scanned across the surface and a topographic image is taken in the normal manner. However, at each point, the feedback loop controlling the tip is broken and the tip position is held constant. The tunnel current is then recorded at a number of preset bias voltages. The result is a set of images, which map out the surface states present at a particular tip-sample bias.

Figure 5 shows a CITS image of the ion beam irradiated surface, with a sample bias of +2.0 V with respect to the tip. The unirradiated region produces a faint signal, since this bias is above the conduction band edge. In the area of ion beam irradiation, strong reproducible features were observed. These features are present in CITS images at all biases from +0.25 to +2.0 V.[7] The features are thought to coincide with the sputtered regions observed in Fig. 4, but this is difficult to confirm, as the resolution of the topographic image taken concurrently with the CITS images is compromised. The line

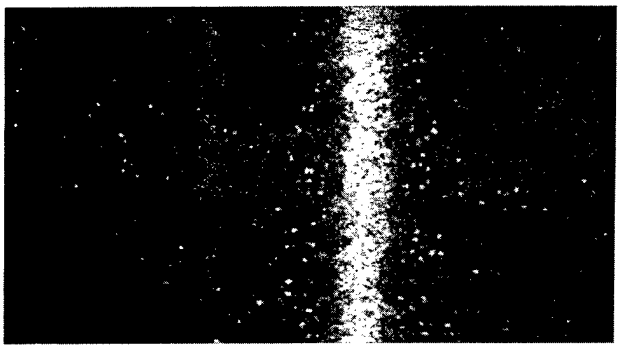


FIG. 5. CITS image of ion beam irradiated line, taken with a sample bias of +2.0 V. Image size is 1.1×2 μm.

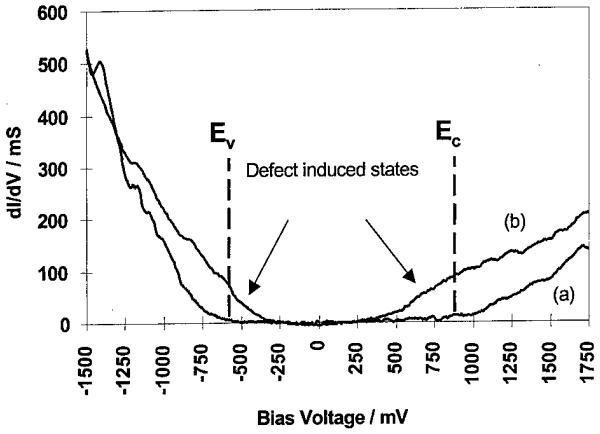


FIG. 6. Differential conductance spectra of (a) un-irradiated material and (b) FIB defect sites. The horizontal axis is the actual bias voltage applied between the sample and STM tip and thus the plots are a measure of the surface state band gap of the material.

density of these features was $\sim 4\times10^{9}$ m^{-1}, almost half that of the ion dose. We believe these features, which were induced by the ion beam irradiation, are due to acceptor states at or just below the surface.

The nature of these surface states was further characterized with differential conductance spectra, formed using a lock-in amplifier technique.[8] Figure 6 shows mean spectra taken at a number of points (a) in the unirradiated surface and (b) at the defect sites observed in CITS images. The horizontal axis is the actual bias voltage applied between the sample and STM tip and thus the plots are a measure of the surface state band gap of the material. The unirradiated surface shows no conductance within the surface state band gap of approximately 1.4 eV. A broad band of states was observed below the conduction band edge of the irradiated material, which corresponds to acceptor states in the surface band gap. The energy resolution of the differential conductance spectra is not high enough to accurately identify specific defect states. Possible acceptor states are Ga antisite

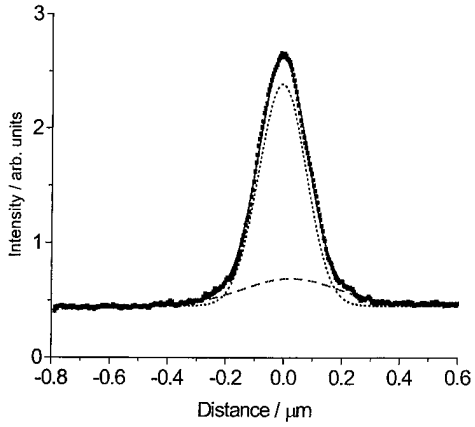


FIG. 7. Profile of the ion beam line formed from the CITS images. The double-Gaussian fit is indicated by the solid line just visible behind the data points. Its two components are indicated by the dotted and dashed lines, respectively.

defects, i.e., Ga atoms sitting on As sites in the lattice. Donor states were also detected on the conductance spectra.

The density of these defects was used to form a profile of the ion beam width. The CITS image shown in Fig. 5, and three others taken at the same sample bias, were integrated along the ion beam line. Assuming the observed defect density is proportional to the ion dose, the mean brightness is roughly proportional to the dose, since the defect size is approximately constant.

Figure 7 shows the beam profile calculated by this method. The beam profile is approximately Gaussian, but the best fit to the data was obtained with a two-Gaussian fit, with full width half maxima of 180 ± 10 and 340 ± 20 nm, respectively.

IV. CONCLUSIONS

The surface electronic states of GaAs are dramatically affected by Ga^+ FIB irradiation. Using the "current" and CITS images, a reliable method for spatially resolving ion beam isolation lines was developed. This opens up the possibility of fabricating devices, with registration between microstructures patterned using FIB lithography and nanometer scale features patterned using the STM tip. CITS images and differential conductance spectra were used to investigate the electrically active defects induced in the ion beam irradiated surface. From the spatial distribution of these defects, the width of the electrically active irradiated line was calculated.

[1]Y. Hirayama and H. Okamoto, J. Vac. Sci. Technol. B **6**, 1019 (1988).
[2]E. H. Linfield, G. A. C. Jones, D. A. Ritchie, A. R. Hamilton, and N. Iredale, J. Cryst. Growth **127**, 41 (1993).
[3]N. Iredale, E. H. Linfield, P. D. Rose, D. A. Ritchie, M. Pepper, and G. A. C. Jones, Semicond. Sci. Technol. **12**, 137 (1997).
[4]G. Ben Assayag, C. Vieu, J. Gierak, P. Sudraud, and A. Corbin, J. Vac. Sci. Technol. B **11**, 2420 (1993).
[5]T. Bever, G. Jäger-Waldau, M. Eckberg, E. Heyen, H. Lage, A. Wieck, and K. Ploog, J. Appl. Phys. **72**, 1858 (1992).
[6]T. Yamamoto, J. Janagisawa, K. Gamo, S. Takaoka, and K. Murase, Jpn. J. Appl. Phys., Part 1 **32**, 6268 (1993).
[7]S. J. Brown, P. D. Rose, E. H. Linfield, D. A. Ritchie, V. Drouot, and G. A. C. Jones, J. Cryst. Growth **175/176**, 346 (1997).
[8]G. Binning, K. H. Frank, H. Fuchs, N. Garcia, B. Reihl, H. Rohrer, F. Salvar, and A. R. Williams, Phys. Rev. Lett. **55**, 991 (1985).

Ion beam synthesis of cobalt disilicide using focused ion beam implantation

J. Teichert,[a] L. Bischoff, and S. Hausmann
Forschungszentrum Rossendorf e.V., Institut für Ionenstrahlphysik und Materialforschung, P.O. Box 510119, D-01314 Dresden, Germany

(Received 17 November 1997; accepted 1 May 1998)

Cobalt disilicide layers were formed by cobalt focused ion beam implantation into silicon. It was found that the $CoSi_2$ layer formation strongly depends on the pixel dwell time. In order to obtain continuous layers, short dwell times of a few μs are needed. Rutherford backscattering and channeling measurements were carried out to understand this effect. The results suggest that the accumulated irradiation damage is larger for longer dwell times. The sputtering yield of cobalt ions was measured and the formation of $CoSi_2$ in noncrystalline silicon investigated. © *1998 American Vacuum Society.* [S0734-211X(98)04404-7]

I. INTRODUCTION

The study of transition metal silicides, in particular of cobalt disilicide, has been a topic of great research interest in recent years because of their promising properties and their compatibility with silicon device technology.[1] $CoSi_2$ is metallic with low resistivity (15 $\mu\Omega$ cm) and high thermal stability. Crystalline $CoSi_2$ has a cubic lattice with a mismatch to silicon of only -1.2%. In recent years, ion beam synthesis (IBS) of $CoSi_2$ has been intensively investigated using conventional broad beam implantation.[2,3] Epitaxial buried or surface layers with perfect interfaces have been produced in (100) and (111) silicon. An elevated substrate temperature (about 400 °C) has been used to prevent the amorphization of the silicon during implantation. Subsequent annealing has usually been performed in two steps at about 600 °C and 1000 °C in order to form the silicide layer (Ostwald ripening) and to reduce crystal defects.

Cobalt focused ion beam (FIB) implantation has been carried out by Aoki *et al.*[4] and Bischoff *et al.*[5] The IBS of cobalt disilicide microstructures with a FIB is advantageous because of the simplification of the process (no implantation mask), the high lateral resolution of the FIB, and the flexibility in pattern design, dose, and ion energy.

II. EXPERIMENTAL EQUIPMENT

The experiments have been performed using the FIB equipment IMSA-100. The ion column contains two electrostatic lenses, an achromatic ExB mass separator and a prelens deflection system. A more detailed description of the FIB system has been given earlier.[6,7] The ion beam is extracted from a liquid-metal ion source of a hair-pin type wetted with an eutectic alloy of cobalt and neodymium. The alloy consists of 36 at. % Co and 64 at. % Nd and has a melting point of 566 °C. For implantation Co^+ and Co^{++} ions have been used at an acceleration voltage of 35 kV. Typical currents are 0.5 nA Co^+ and 0.6 nA Co^{++} corresponding to ion fluxes of 3.1×10^9 s^{-1} and 1.9×10^9 s^{-1}, respectively. The beam spot size was between 300 and 100 nm. Figure 1 presents the beam profile which was measured by sputtering of a thin gold film on a silicon substrate with various doses. For the ion beam synthesis experiments two target heating assemblies have been developed allowing temperatures up to 450 °C (2 in. wafers) and 600 °C (10 $\times$ 10 mm^2 chips).

III. ION BEAM SYNTHESIS OF $CoSi_2$

A. Effects of high current density and pixel dwell time

From broad beam implantation it is known that IBS of $CoSi_2$ requires a careful choice of the implantation parameters. The ion energy determines the depth and the dose the thickness of the formed layer. A minimum dose, called the critical dose, is required to obtain closed layers. The critical dose is determined by the maximum concentration of implanted cobalt atoms in the silicon which must be about 18 at. %. Due to the broadening of the cobalt distribution the critical dose and thus the minimum layer thickness increase with ion energy. The target heating during implantation prevents amorphization due to ion beam induced defects. For the epitaxial growth of the cobalt disilicide the damage of the silicon matrix must be low and therefore the substrate temperature sufficiently high. On the other hand, a too high temperature increases the cobalt diffusion which has a negative effect on the layer formation.

There are two main differences between FIB and broad beam implantation. First, the FIB writes the patterns in a serial manner, i.e., it irradiates the sample in discrete pixels with a given pixel dwell time. This dwell time is an additional parameter which can be varied. It is possible to write a pattern with a certain dose using a short dwell time and a large repetition or vice versa. Second, the current density of the FIB of about 1–10 A cm^{-2} is orders of magnitude higher than that of ion beams in conventional implanters.

In the following, investigations of the influence of the pixel dwell time on $CoSi_2$ layer formation are presented. A series of 40×40 μm^2 squares were implanted in (111) Si. All the parameters were constant (about 5×10^{16}

[a]Electronic mail:j.teichert@fz-rossendorf.de

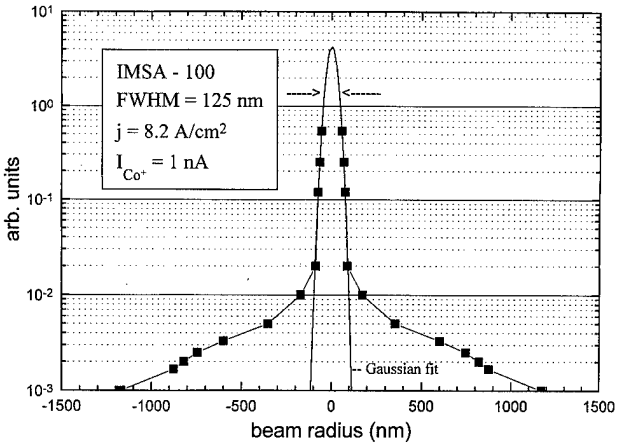


FIG. 1. Measured intensity distribution of the Co^+ ion beam.

Co^+ cm^{-2}, 35 keV, 420 °C) with exception of the pixel dwell time which was varied between 1 and 200 μs and the corresponding repetition number in order to hold the dose constant. After implantation, the sample was annealed (600 °C for 1 h and 1000 °C for 30 min in N_2) and the silicon top layer was removed by reactive ion etching (CF_4 for 4 min). Scanning electron microscope (SEM) micrographs of the produced $CoSi_2$ layers are shown in Fig. 2. These plane view images depict the influence of the pixel dwell time. For the shortest dwell time of 1 μs a closed layer is formed. With increasing dwell times the layer quality deteriorates and the film tears. For the longest dwell times coarsening of the $CoSi_2$ happens and rough layers consisting of isolated $CoSi_2$ grains were obtained.

Another series of implantations with varied pixel dwell time performed at room temperature shows similar microstructures of the $CoSi_2$ layers for all dwell times. The layers are not closed but contain large $CoSi_2$ precipitates. Thus, inaccuracies in the implantation dose connected with dwell time variation or other fluctuations can be excluded. Santamore *et al.*[8] found a dwell time dependence of the sputter yield. But this effect occurs at longer pixel dwell times (>2 ms) and is independent of the target temperature. In contrast to broad beam implantation the FIB does not heat up the target. In spite of the extremely high current density of the FIB there is also no significant local target temperature rise at the beam position. It was shown by Melngailis[9] that due to the small beam spot and the high thermal conductance of Si the heat flow is large and the temperature rise is only some degrees.

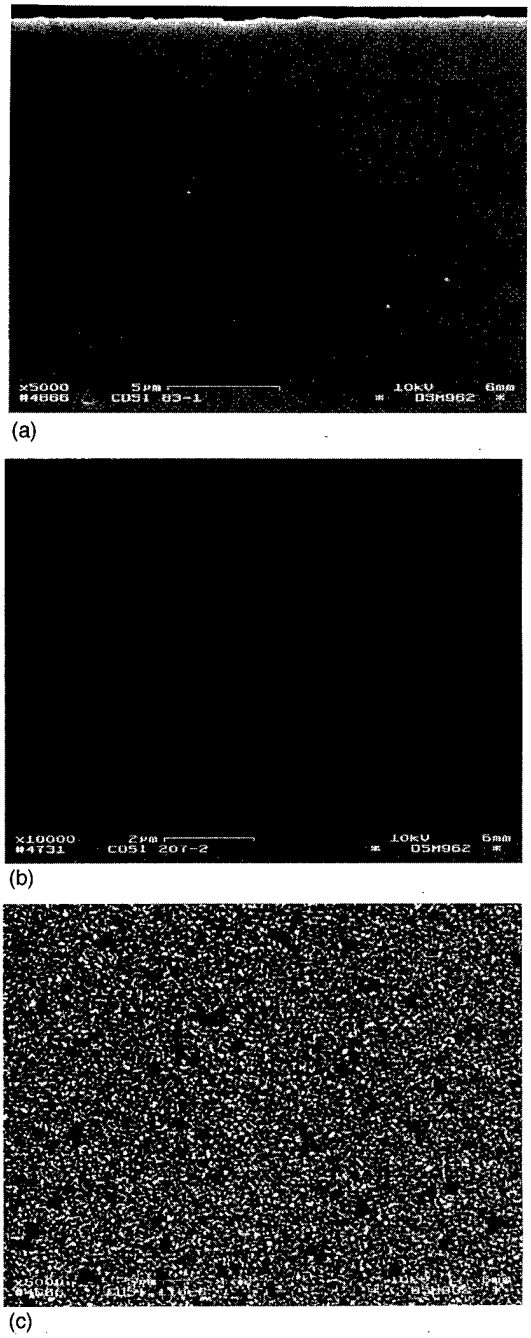


FIG. 2. SEM plane view micrographs of cobalt disilicide layers produced in Si (111) with 35 keV and rather similar implantation conditions with exception of the pixel dwell time shown after annealing and reactive ion etching. (a) 1 μs dwell time, 5.5×10^{16} cm^{-2}, 420 °C, (b) 5 μs dwell time, 6.0×10^{16} cm^{-2}, 430 °C, (c) 200 μs dwell time, 5.3×10^{16} cm^{-2}, 420 °C [the same scale as (b)].

B. Characterization by Rutherford backscattering and channeling

In order to explain our results either the precipitate growth or the ion beam induced damage must be affected by the length of the dwell time. To obtain a deeper understanding of the effect detailed information on the cobalt distribution and damage accumulation after implantation should be obtained. A common technique for that purpose is Rutherford backscattering (RBS) and channeling analysis which is widely used in investigations involving conventional broad beam implantation. RBS and channeling analysis provide information about the cobalt distribution as well as radiation damage of as-implanted and annealed samples. A suitable tool for RBS measurements of small areas is the nuclear microprobe.[10] On (111) Si samples squares of 40×40 μm^2 were implanted with Co^{++} (70 keV) ions with a dose of 5×10^{16} cm^{-2} at 400 °C. These samples have been analyzed

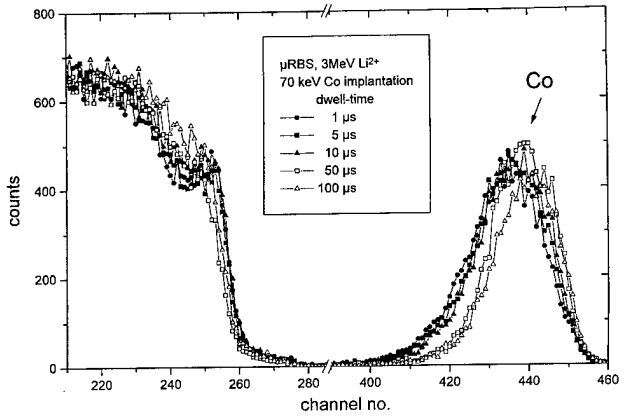


FIG. 3. Nuclear microprobe RBS results measured on as-implanted samples. The implantation parameters are 5×10^{16} Co^{++} cm^{-2}, 70 keV, 430 °C. The spectra for implantation with different dwell times are shown.

in the as-implanted state using an analyzing beam of about 5 μm diameter. Figure 3 shows the measured random RBS spectra for five different pixel dwell times. It is obvious that for the three shortest dwell times of 1, 5, and 10 μs the deposition of the cobalt atoms is deeper in the silicon. In the figure the cobalt peak is marked and the peak shift towards smaller channel numbers corresponds to a larger depth in the substrate. For these three samples continuous silicide layers were obtained after annealing. Since the FIB implantation was carried out at 0° tilt angle, the deeper cobalt distributions may be due to channeling of the cobalt ions. Thus, the decreasing depth at increasing dwell time might be explained by dechanneling due to the accumulated damage, which increases at longer dwell time.

For channeling analysis the standard RBS setup at a Van de Graaff accelerator with a beam diameter of about 1 mm combined with a special preparation method of the samples have been applied.[11] In this method the nonimplanted part of the silicon is covered by a thick layer containing only light elements. Circular areas of 300 μm diameter were implanted which takes about 45 min. The implantations were carried out with pixel dwell times of 1 and 250 μs at room temperature (RT) and 430 °C. The dose was 2×10^{16} cm^{-2} for both dwell times, i.e., for 250 μs dwell time one writing cycle and for 1 μs, 250 writing cycles were needed. Since the total area of the 300 μm spot was written in one cycle the repetition times were about 11 s for 1 μs dwell time. The dose of 2×10^{16} cm^{-2} is lower than the critical dose for closed layer formation but it is sufficient to study the damage accumulation after implantation in the as-implanted samples. Figure 4 shows the interesting part of the measured RBS/channeling analysis of the sample implanted at 430 °C. For 250 μs dwell time [Fig. 4(b)] the damage peak of the silicon in the aligned spectrum indicates that the implantation layer is amorphous. However, in the case of 1 μs there is significantly lower damage which does not reach the random level. The damage peak position corresponds to the depth where the Co ions are deposited. Between this peak and the surface there exists a region of low damage indicating that the Si top layer is crystalline.

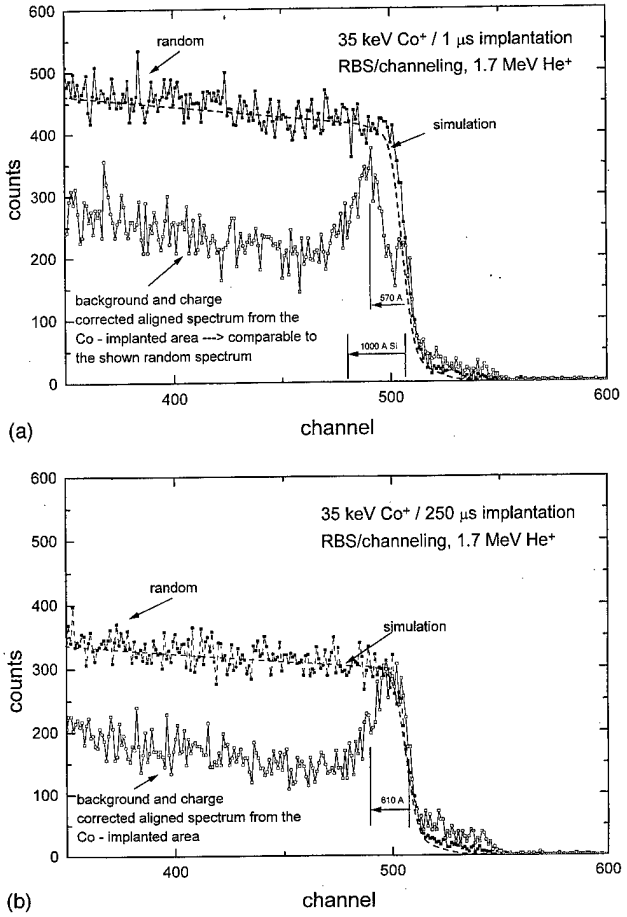


FIG. 4. Random and aligned spectra of as-implanted samples. Implantation with a dose of 1.5×10^{16} Co^{+} cm^{-2} at 35 keV and 430 °C for two different dwell times of 1 μs (a) and 250 μs (b).

In summary, both measurements established that in FIB implantation the crystal defect accumulation depends on the pixel dwell time and is larger for longer dwell times. Similar results were recently found by Musil *et al.*[12] who studied the damage production in GaAs induced by gallium FIB implantation. In the $CoSi_2$ ion beam synthesis, as mentioned above, a different degree of damage of the silicon matrix has a strong influence on the layer formation during the process of Ostwald ripening. Therefore, the heavier damage in the silicon crystal may be the reason for the insufficient quality of the $CoSi_2$ layers produced with long pixel dwell times. Using short dwell times of about 1 μs the results are similar to that of conventional broad beams where the thermally induced defect annihilation prevents an amorphization of silicon during implantation.

C. Sputtering

High dose implantation at ion energies of a few ten kV is strongly affected by target sputtering. For this reason sputtering yields of cobalt ions on silicon have been determined using the volume loss method. Square holes of 20 $\times$20 μm^2 were produced by irradiation with doses varying from 5×10^{15} cm^{-2} to 8×10^{17} cm^{-2}. This method is precise for large doses where the steady-state concentration is

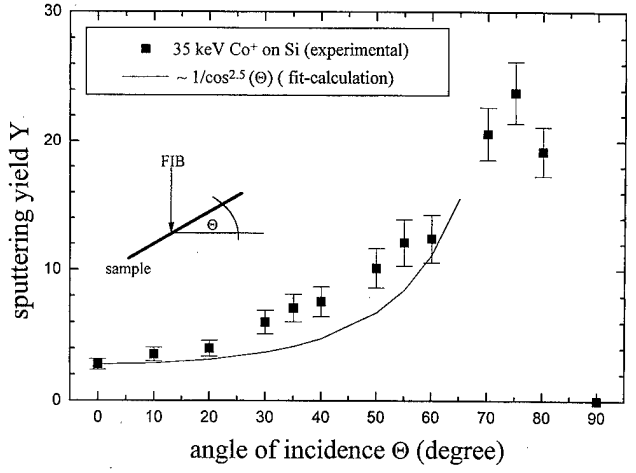


FIG. 5. Sputtering yield as a function of angle of incidence for 35 keV Co ions on Si.

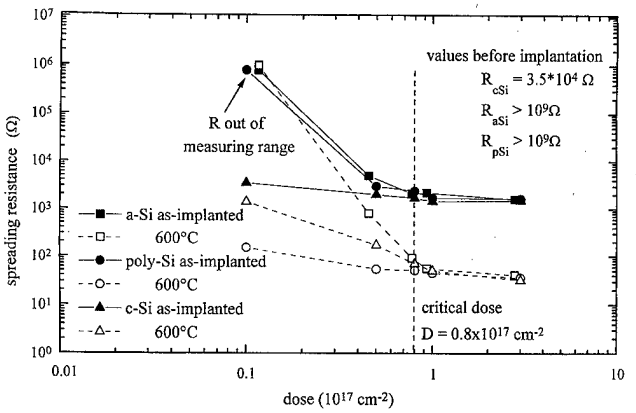


FIG. 6. Spreading resistance as a function of implantation dose for the different Si substrates measured before and after annealing.

reached at the bottom of the sputtering holes. Then the number of target atoms lost can be calculated from the sputter hole volume and the atomic density of the target material. The only inaccuracy is that the sputtering coefficient is measured at a target which is enriched with projectile atoms. Figure 5 presents the sputtering yield for 35 keV Co^+ ions as a function of angle of incidence. Compared with normal incidence the sputtering coefficient at 54° [angle of the ⟨111⟩ planes in (100) Si] is about five times larger. Increasing the target temperature results in a nearly linear decrease of the sputtering coefficient from 2.5 (RT) to 1.4 (400 °C).

D. Implantation in noncrystalline silicon

In micromechanical applications the surfaces often consist of polycrystalline silicon (poly-Si) or amorphous silicon (*a*-Si). Therefore these substrates and, for comparison, also crystalline silicon (*c*-Si) were implanted with 35 keV cobalt ions at RT. Stripes of 150 μm length and 10 μm width were written with different doses. The subsequent annealing was performed at 600 °C for 1 h in N_2. The samples were electrically analyzed using spreading resistance measurements. In Fig. 6 the results are shown as a function of implantation dose before and after annealing. For doses above the critical dose for closed layer formation the data of the different materials are equal and correspond to a resistivity of about 50 $\mu\Omega$ cm. The annealing at 1000 °C had not been carried out because a disintegration of the $CoSi_2$ structures takes place. In a polycrystalline or amorphous matrix the thermal stability of the $CoSi_2$ is up to about 700 °C.[13] Above this temperature diffusion of the Co occurs, mainly towards the surface, and large isolated $CoSi_2$ grains grow. This effect is rather similar to that happens at implantations in crystalline silicon when the crystal is strongly damaged (low temperature or long pixel dwell times). It should be noted that in contrast to the high-temperature implantation discussed earlier here polycrystalline $CoSi_2$ is formed. The difference is not only a higher resistivity and a lower thermal stability but also another type of interface to silicon. Here the $CoSi_2$/Si junction is Ohmic, whereas in the case of single-crystalline silicon the $CoSi_2$/Si junction is a Schottky contact.

ACKNOWLEDGMENTS

The Deutsche Forschungsgemeinschaft for their financial support under contract Te 250/1-1 and Te 250/3-1 are gratefully acknowledged. The authors would like to thank E. Christalle and Dr. R. Müller for SEM imaging, Dr. D. Grambole, F. Herrmann, and Dr. M. Voelskow for RBS/C measurements, and I. Beatus for technical support.

[1]K. Maex, Mater. Sci. Eng., R. **11**, 53 (1993).
[2]A. E. White, K. T. Short, R. C. Dynes, J. P. Garno, and J. M. Gibson, Appl. Phys. Lett. **50**, 95 (1987).
[3]S. Mantl, Mater. Sci. Eng., R. **8**, 1 (1993).
[4]T. Aoki, K. Gamo, S. Namba, T. Shiokawa, K. Toyoda, H. Okabayashi, H. Mori, and H. Fuijita, Nucl. Instrum. Methods Phys. Res. B **39**, 291 (1989).
[5]L. Bischoff, J. Teichert, E. Hesse, D. Panknin, and W. Skorupa, J. Vac. Sci. Technol. B **12**, 3523 (1994).
[6]L. Bischoff, E. Hesse, D. Janssen, F. K. Naehring, F. Nötzold, G. Schmidt, and J. Teichert, Microelectron. Eng. **13**, 637 (1991).
[7]L. Bischoff, E. Hesse, D. Panknin, W. Skorupa, and J. Teichert, Microelectron. Eng. **23**, 115 (1994).
[8]D. Santamore, K. Edinger, J. Orloff, and J. Melngailis, J. Vac. Sci. Technol. B **15**, 2346 (1997).
[9]J. Melngailis, J. Vac. Sci. Technol. B **5**, 469 (1987).
[10]F. Herrmann and D. Grambole, Nucl. Instrum. Methods Phys. Res. B **104**, 26 (1995).
[11]J. Teichert, M. Voelskow, L. Bischoff, and S. Hausmann, Vaccum (to be published).
[12]C. R. Musil, J. Melngailis, S. Etchin, and T. E. Haynes, J. Appl. Phys. **80**, 3727 (1996).
[13]J. Teichert, L. Bischoff, E. Hesse, P. Schneider, D. Panknin, T. Gessner, B. Löbner, and N. Zichner, J. Micromech. Microeng. **6**, 272 (1996).

AUTHOR INDEX

To papers from the Japan/U.S. Workshop on Formation of Ion Nanobeams and Applications to Materials Processing

Aihara, R.–(4) 2484

Berry, I. L.–(4) 2444, 2469
Bischoff, L.–(4) 2574
Brown, S.–(4) 2570
Brugger, H.–(4) 2562

Cheng, Zhaohui–(4) 2473
Chikyow, Toyohiro–(4) 2538

Forchel, A.–(4) 2484, 2562
Fujita, Jun-Ichi–(4) 2439
Fukai, T.–(4) 2479

Gamo, Kenji–(4) 2511
Griffis, D. P.–(4) 2494

Hausmann, S.–(4) 2574
Hirayama, Y.–(4) 2543
Höfling, E.–(4) 2562

Isakozawa, S.–(4) 2532
Ishikawa, Junzo–(4) 2515
Ishitani, T.–(4) 2532
Iwano, Hirotaka–(4) 2551

Jarausch, K. F.–(4) 2494
Jones, G. A. C.–(4) 2570
Joyce, R. J.–(4) 2449

Kalbitzer, S.–(4) 2455
Kamino, T.–(4) 2532
Kim, H.–(4) 2547
Kinokuni, M.–(4) 2484
Knoblauch, A.–(4) 2455
Koguchi, Nobuyuki–(4) 2538
Koike, H.–(4) 2532
Komuro, Masanori–(4) 2506
König, H.–(4) 2562
Kubena, R. L.–(4) 2449
Kumura, Yoshinori–(4) 2489
Kuramoto, Yasuyuki–(4) 2473

Lee, Hyung-Ik–(4) 2528

Mais, N.–(4) 2562
Matsui, Shinji–(4) 2439
Matsukawa, T.–(4) 2479
Matsukawa, Takashi–(4) 2489
Matsumoto, H.–(4) 2532
McNulty, H. H.–(4) 2449
Mimura, R.–(4) 2484
Miyamoto, Iwao–(4) 2506
Mizumura, Michinobu–(4) 2555
Mogren, S.–(4) 2469
Müssig, H.–(4) 2562

Nagamachi, Shinji–(4) 2515
Nakayama, Masayoshi–(4) 2511
Nassar, Raja–(4) 2499
Nihei, Yoshimasa–(4) 2473
Noda, T.–(4) 2547

Ohdomari, I.–(4) 2479
Ohdomari, Iwao–(4) 2489
Ohno, Naoto–(4) 2506
Okabe, Jun–(4) 2489
Owari, Masanori–(4) 2473

Phillips, J. R.–(4) 2494

Reithmaier, J. P.–(4) 2562
Rose, P. D.–(4) 2570
Russell, P. E.–(4) 2494

Saitoh, Keiya–(4) 2555
Saka, H.–(4) 2522
Sakaguchi, Kiyoshi–(4) 2462
Sakaki, H.–(4) 2547
Sakamoto, Tetsuo–(4) 2473
Saku, T.–(4) 2543
Sawaragi, H.–(4) 2484
Sekine, Tetsu–(4) 2462
Shimizu, Ryuichi–(4) 2528
Shinada, T.–(4) 2479
Shinada, Takahiro–(4) 2489
Stark, T. J.–(4) 2494
Stratton, F. P.–(4) 2449

Takahashi, Masanori–(4) 2473
Takeda, Shuuichi–(4) 2506
Taniguchi, Jun–(4) 2506
Taniguchi, Y.–(4) 2532
Teichert, J.–(4) 2574

Ueda, Masahiro–(4) 2515

Vajo, J.–(4) 2449
Vasile, Michael J.–(4) 2499
Versen, M.–(4) 2567

Wakaya, Fujio–(4) 2511
Wieck, A. D.–(4) 2567
Wiemann, C.–(4) 2567

Xie, Jushan–(4) 2499

Yaguchi, T.–(4) 2532
Yamaguchi, Hiroshi–(4) 2555
Yanagisawa, Junichi–(4) 2511
Yasuda, Yukio–(4) 2551

Zaima, Shigeaki–(4) 2551

CUMULATIVE AUTHOR INDEX

All authors published to date in Volume 16 are listed alphabetically with the issue (in parentheses) and page numbers following the dash. An (E) after the page number indicates Erratum.

Abe, Y.–(2) 803
Abe, Yoshiaki–(2) 818
Abrams, R. H.–(2) 749
Abu-Ageel, N.–(1) 142
Achiq, A.–(4) 1851
Adachi, Hiroshi–(2) 875
Adamson, Douglas H.–(2) 544
Agarwala, Sambhu–(2) 511
Ager, R.–(1) 142
Ahmed, M. M.–(4) 2034
Ahn, C. C.–(4) 1937
Ahn, Sook–(1) 192
Aihara, R.–(4) 2484
Aitchison, J. S.–(4) 1860
Aizawa, Shin-ichi–(1) 38
Albrecht, M.–(1) 19
Alkemade, P. F. A.–(1) 373; (4) 1971
Alperin, J.–(3) 1275, 1401, 1489
Altmeyer, S.–(4) 1983
Alvis, R.–(1) 339; (3) 1467
Ambacher, O.–(4) 2224
Ambrose, V.–(1) 420
Andoh, Yasuko–(1) 38
Anselm, K. A.–(3) 1426
Antipov, V. G.–(3) 1289
Aqariden, F.–(3) 1309
Archer, L. A.–(4) 2121
Arita, Kiyoshi–(2) 519, 670
Arlett, J.–(2) 578
Arnoldbik, W. M.–(1) 373
Arthur, J. R.–(3) 1356
Asada, M.–(2) 851
Asakawa, K.–(1) 1
Asano, Tanemasa–(2) 519, 651, 670
Asenjo, A.–(2) 654
Ashihara, Kazuto–(2) 875
Aslam, D. M.–(1) 142
Aslam, Dean M.–(2) 712
Aspnes, D. E.–(4) 2355, 2367
Auret, F. D.–(4) 1873

Bachmann, T.–(2) 540
Badzian, A.–(2) 681
Badzian, T.–(2) 681
Baek, Junho–(1) 430
Bahierathan, B.–(2) 599
Baidakova, M. V.–(3) 1456
Baillargeon, J. N.–(3) 1422
Baker, J.–(1) 420
Baklanov, M. R.–(1) 164
Balasubramaniam, H.–(3) 1467
Balestra, F.–(4) 1812
Bandhu, R.–(4) 2177
Baptist, R.–(2) 841
Barnett, Scott A.–(4) 1885
Bashir, R.–(4) 2118
Baskin, L. M.–(1) 232
Beaudoin, R. M.–(3) 1408
Belk, J. G.–(4) 2373
Bell, K. A.–(4) 2367
Bell, L. D.–(4) 2286
Ben Assayag, G.–(4) 1919
Bennett, B. R.–(4) 2381
Bennett, Joe–(1) 447
Benninghoven, A.–(1) 292, 298
Bensaoula, A.–(3) 1270
Bensaoula, A. H.–(4) 2280
Ber, B. Ya.–(1) 426
Beresford, R.–(3) 1293
Berishev, I.–(3) 1270
Berning, D. W.–(1) 339
Berry, I. L.–(4) 2444, 2469
Bertness, K. A.–(3) 1492
Bhat, R.–(3) 1417
Bhattacharya, P.–(3) 1343, 1417
Bhave, Tejashree M.–(4) 2073
Bhoraskar, S. V.–(4) 2073
Bi, W. G.–(4) 2395
Bischoff, L.–(4) 2574
Black, Andrew J.–(1) 98
Blant, A. V.–(4) 2237
Block, T. R.–(3) 1475
Bonanni, A.–(4) 2334
Bonanni, B.–(4) 2334
Bonar, J. R.–(4) 1860
Bonard, J.-M.–(4) 2334
Boonzaayer, M.–(3) 1484
Boonzaayer, Martin–(3) 1502
Booske, J.–(2) 532
Booske, John H.–(1) 415
Bormatova, L. V.–(2) 678
Böttcher, J.–(1) 210
Bottomley, Lawrence–(3) 1326
Boudart, B.–(2) 561
Boyd, Wendell, Jr.–(1) 447
Boyer, Lionel–(4) 2006
Bracken, C.–(4) 2118
Brandt, Oliver–(4) 2229
Braun, W.–(3) 1507; (4) 2317, 2404
Braun, Wolfgang–(3) 1502
Braunstein, J.–(4) 1864
Breiland, W. G.–(3) 1498
Briddon, P. R.–(4) 1794
Brillson, L. J.–(4) 2177
Brodie, I.–(2) 758
Brongersma, S. H.–(4) 2188
Brown, April S.–(3) 1300, 1326; (4) 2308
Brown, J. D.–(3) 1282
Brown, S.–(4) 2570
Bruce, D. M.–(4) 1773
Brugger, H.–(4) 2562
Bruschi, P.–(2) 665
Bryce, A. C.–(4) 1818
Buchanan, M.–(4) 1790
Buh, G. H.–(2) 826
Bukesov, S. A.–(4) 2082
Burm, J.–(4) 2110
Busygina, L. A.–(1) 426
Buyanova, I. A.–(4) 1928
Byun, Jae-Dong–(2) 858

Cabibil, H.–(1) 30
Cai, Weizhong–(3) 1361
Caldwell, D. A.–(4) 2280
Campbell, J. C.–(3) 1426
Campbell, S.–(2) 483
Caneau, C.–(3) 1417
Cappy, Alain–(1) 255(E)
Cardinaud, Ch.–(4) 1823
Carlin, J. A.–(3) 1372
Carter, C. B.–(4) 2280
Carter, W. Brent–(3) 1326
Carter-Coman, C.–(3) 1300
Castaño, J. L.–(4) 1804
Castell, M. R.–(4) 2188
Cavanagh, Richard R.–(4) 1948
Cavus, Abdullah–(3) 1312
Caymax, M.–(1) 394
Chaikin, Paul M.–(2) 544
Chakraborty, Sahana–(4) 2159
Chan, C. H.–(2) 570
Chan, Yi-Jen–(1) 253
Chandrasekhar, D.–(4) 1937
Chang, C. P.–(1) 334
Chang, M. F.–(3) 1401
Chang-Liao, Kuei-Shu–(1) 250
Chao, I-Na–(3) 1459
Chao, Kuo-Jen–(1) 453
Charbonneau, S.–(2) 578; (4) 1790
Charbonnier, Francis–(2) 880
Chatterjee, T.–(3) 1463
Cheah, L. K.–(4) 2049
Chen, D. H.–(2) 697
Chen, H.–(3) 1377
Chen, J.–(2) 697; (3) 1305
Chen, L. C.–(4) 2280
Chen, L. J.–(4) 2013, 2026
Chen, M. C.–(4) 2026
Chen, P.–(3) 1330, 1334
Chen, W. M.–(4) 1928
Chen, Y. H.–(2) 570
Chen, Y. J.–(2) 511
Chen, Y. K.–(3) 1398, 1404
Chen, Z.–(2) 507
Cheng, K. Y.–(3) 1347, 1352, 1395
Cheng, T. S.–(4) 2237
Cheng, X.-C.–(4) 2291
Cheng, Zhaohui–(4) 2473
Cherrington, Blake E.–(2) 490
Chervela, N.–(3) 1343
Chevallier, J.–(4) 1777
Chichibu, Shigefusa–(4) 2204
Chikyow, Toyohiro–(4) 2538
Childress, J. R.–(4) 2275
Chizhov, I.–(3) 1339
Cho, A. Y.–(3) 1422
Cho, Guangsup–(4) 2079
Cho, Keum Jae–(1) 192
Cho, Kyoung Ik–(1) 238; (2) 773, 811, 815, 871
Cho, Kyoung-Ik–(1) 242; (2) 858
Cho, Sung Hee–(4) 2086
Cho, Sung-Pyo–(4) 2426
Choi, Eun-Ha–(4) 2079
Choi, H. J.–(2) 826
Choi, J. H.–(2) 736
Choi, Kyu Man–(2) 916
Choi, S. S.–(2) 826
Choi, Sang Soo–(4) 1992
Choi, W. B.–(2) 716
Chou, L. J.–(3) 1352, 1395
Chow, D. H.–(3) 1413
Chow, P.–(3) 1385
Chow, P. P.–(3) 1286
Christ, M.–(2) 693
Christensen, D. H.–(3) 1492
Christiansen, S.–(1) 19
Chu, D. P.–(1) 302, 377
Chu, Paul K.–(1) 197
Chu, S. N. G.–(3) 1422, 1467
Chua, S. J.–(2) 565
Chuang, M. C.–(4) 2026
Chung, Bokeon–(2) 700
Chung, C.-H.–(4) 2199
Chung, Chee Won–(4) 1894
Chung, Hong-Bay–(4) 1987
Chung, M. S.–(2) 906
Chung, Suk Jae–(2) 705
Claflin, B.–(4) 2154
Clark, J. G.–(1) 394
Clarysse, T.–(1) 260, 320, 355, 394, 401
Collart, E. J. H.–(1) 280, 298
Colligon, J. S.–(2) 589
Collins, D. C.–(1) 362
Condon, G. R.–(1) 23
Conley, J. F., Jr.–(4) 2134
Constancias, C.–(2) 841
Cook, J. M.–(4) 1998
Cook, J. W., Jr.–(3) 1282
Cooke, G. A.–(1) 302
Corcoran, S. F.–(1) 272
Correia, A.–(2) 654
Coulombe, Stephen A.–(1) 80
Coupeau, C.–(4) 1964
Craighead, H. G.–(4) 1943
Croke, E. T.–(4) 1937
Crosnier, Y.–(1) 223; (2) 561
Cuomo, J. J.–(2) 716
Current, M. I.–(1) 327
Currie, M.–(3) 1471
Cutler, P. H.–(2) 900, 906
Czotscher, K.–(4) 1864
Czuprynski, P.–(1) 147

Dagata, John–(2) 633
Dagenais, Mario–(2) 511
d'Agostino, R.–(4) 1867
Daleiden, J.–(4) 1864
Das Sarma, S.–(4) 2417
Davidson, J. L.–(2) 684, 732
Davis, P. R.–(4) 2057
Davis, W. A.–(4) 1943
Däweritz, L.–(4) 2404
Deen, M. J.–(2) 628; (4) 1812, 1881
Deenapanray, P. N. K.–(4) 1873
DeHerrera, M.–(3) 1484
Dekura, Takateru–(1) 38
De la Hidalga-W., F. J.–(4) 1812
De La Rue, R. M.–(4) 1818
DeLong, M. C.–(4) 2328
deLozanne, Alex–(4) 2215
de Lyon, T. J.–(3) 1321
DenBaars, S. P.–(2) 822
Deng, F.–(3) 1297
Deng, S. Z.–(2) 697
De Wolf, P.–(1) 320, 355, 367, 394
DeWolf, P.–(1) 401
Diatezua, D. M.–(2) 507
Dickson, M.–(2) 523
Diligenti, A.–(2) 665
Dimitrov, R.–(4) 2224
Dion, M.–(4) 1790
Dmitrienko, A. O.–(4) 2082
Doemling, M. F.–(4) 1998
Doezema, R. E.–(3) 1367
Doolittle, William A.–(3) 1300
dos Santos, M. C.–(4) 2105
Downes, James–(4) 2250
Downey, Daniel F.–(1) 286
Dowsett, M. G.–(1) 302, 377
Drawl, W.–(2) 681
Drigo, A.–(4) 2334
Drummond, T. J.–(3) 1498
Du, Q.–(3) 1401

2580

Duane, Michael–(1) 306
Dubray, J.-J.–(3) 1413
Ducroquet, F.–(2) 787
Duda, Laurent C.–(4) 2250
Dumpich, G.–(1) 77
Dunlap, H. L.–(3) 1413
Dupuis, R. D.–(4) 2215
Dutoit, M.–(2) 582

Eastman, L. F.–(4) 1943
Eckert, Patrick–(3) 1326
Economy, James–(1) 125
Edmond, J. A.–(3) 1282
Edwards, Hal–(1) 476; (2) 633
Eftekhari, G.–(4) 2115
Egami, N.–(2) 575
Eguchi, Toyoaki–(4) 2324, 2426
El-Masry, N. A.–(3) 1282
Elyukhin, V. A.–(3) 1289, 1456
Emiliani, V.–(4) 2350
England, J. G.–(1) 327
Eppell, Steven J.–(4) 2099
Erickson, John W.–(1) 197
Esser, N.–(4) 2350
Etoh, Masahiro–(2) 519
Eustis, T. J.–(4) 1943
Evans, D. A.–(4) 2350
Evoy, S.–(4) 1943
Eyckeler, M.–(4) 2224
Eyink, K. G.–(3) 1439, 1479

Faber, J. S.–(4) 2063
Fafard, S.–(2) 578; (4) 1790
Falcon, M.–(3) 1435
Faleev, N. N.–(3) 1289, 1456
Fang, X. M.–(3) 1459, 1463
Farrer, J. K.–(4) 2280
Feder, K. S.–(4) 2113
Feenstra, R. M.–(4) 2242
Felch, S. B.–(1) 435
Felch, Susan B.–(1) 286
Felker, B. S.–(1) 173
Feng, Y.–(2) 578
Ferry, D. K.–(4) 2165
Feudel, T.–(1) 440
Feurprier, Y.–(4) 1823
Filip, V.–(2) 888
Fitzgerald, E. A.–(3) 1471
Fleischman, Aaron J.–(2) 536
Foad, M. A.–(1) 316, 327
Fontaine, Chantal–(1) 204
Forchel, A.–(4) 2484, 2562
Forrest, R. L.–(3) 1451
Foster, J. E.–(2) 532
Fourré, Hervé–(1) 255(E)
Foxon, C. T.–(4) 2237
Fracassi, F.–(4) 1867
Franciosi, A.–(4) 2334
Fransen, M. J.–(4) 2063
Friedrichowski, S.–(1) 77
Frisch, A. M.–(4) 2350
Fritz, I. J.–(3) 1498
Fujimori, Y.–(2) 829
Fujita, Jun-Ichi–(4) 2439
Fujita, K.–(2) 575
Fukai, T.–(4) 2479
Fukui, Takashi–(4) 2387
Fursey, G. N.–(1) 232; (2) 910
Fyhn, M. F.–(4) 1777

Galligan, J. M.–(4) 2127
Gamo, K.–(4) 2313
Gamo, Kenji–(4) 2511
Ganière, J.-D.–(4) 2334
Gao, Yumin–(1) 197
García, B. J.–(4) 1804
García, N.–(2) 654
Garrido, B.–(4) 1851
Gaworzewski, P.–(1) 406
Gearhart, Steven S.–(1) 415
Gebretsadik, H.–(3) 1417
Gendry, M.–(4) 1786
Gertner, E. R.–(4) 2286
Gianardi, D. M., Jr.–(3) 1435
Gibis, R.–(1) 210
Gierak, J.–(4) 1919
Girard, J. C.–(4) 1964
Givargizov, E. I.–(2) 678
Glazanov, D. V.–(1) 232; (2) 910
Göhl, A.–(2) 693
Goldammer, K. J.–(3) 1367
Goldberg, R. D.–(4) 1790
Golding, T.–(3) 1385
Golding, T. D.–(3) 1451
Goodhue, W. D.–(3) 1430
Goorsky, M. S.–(3) 1381
Gopir, G.–(4) 1794
Goss, S. H.–(3) 1439
Gotoh, Y.–(2) 829, 833, 895
Gotthold, D. W.–(3) 1278
Gotza, M.–(2) 582
Gourbilleau, F.–(4) 1851
Gravesteijn, D. J.–(1) 137, 280, 298
Grazulis, L.–(3) 1439
Grenet, Genevieve–(4) 2421
Greve, D. W.–(4) 2242
Gribelyuk, M. A.–(1) 476
Griffis, D. P.–(4) 2494
Grilhé, J.–(4) 1964
Grudowski, P. A.–(4) 2215
Gu, C. Z.–(2) 710
Guerret-Piecourt, Christelle–(1) 204
Guidotti, Daniel–(2) 609
Guillermet, M.–(4) 1833
Guk, E. G.–(1) 426
Guo, Jinghua–(4) 2250
Guo, Y.–(1) 435
Gupta, I.–(2) 496
Guriev, A. I.–(3) 1289
Gutierrez-D., E. A.–(4) 1812

Haas, T. W.–(3) 1439
Habermann, T.–(2) 693
Hafizi, M.–(3) 1413
Hahm, Sung-Ho–(2) 762
Hale, J. S.–(3) 1484
Hamidi, A.–(4) 1983
Hamm, R. A.–(4) 2110
Hammel, P. C.–(4) 2275
Han, A. C.–(3) 1475
Han, I. T.–(4) 2052
Han, Jeong Yeul–(1) 192
Han, Seok-Yoon–(2) 724
Hanada, T.–(4) 2342
Hänni, W.–(1) 355
Hansing, C.–(3) 1426
Hansson, G. V.–(4) 1928
Hara, Akio–(1) 183
Harde, P.–(1) 210
Harnett, C. K.–(4) 1943
Harper, J.–(3) 1389
Harrington, William L.–(1) 286
Harrison, Christopher–(2) 544
Harrison, W. A.–(4) 2328
Hartford, C. L.–(1) 316, 411
Hartmann, N.–(4) 2254
Hasegawa, Hideki–(4) 2159, 2387
Hashizume, Tamotsu–(4) 2159
Hatakeyama, Tadao–(2) 553
Hausmann, S.–(4) 2574
Haworth, L.–(4) 2254
Hays, S. P.–(3) 1492
Hebert, F.–(4) 2118
Heinemann, B.–(1) 292
Heitz, R.–(3) 1330, 1334
Heitzmann, M.–(1) 147
Hellemans, L.–(1) 349, 355, 367, 401
Herrmann, T.–(4) 2350
Heun, S.–(4) 2334
Heyns, M.–(4) 2091
Hickernell, R. K.–(3) 1492
Hickman, R., II–(3) 1286
Hicks, S. E.–(4) 1818, 1860
Higa, Katsuya–(2) 651
Hill, P.–(4) 2254
Hillard, R. J.–(1) 316, 411
Hillenius, S. J.–(1) 334
Hillier, G.–(4) 1790
Hinds, B. J.–(4) 2171
Hingerl, K.–(4) 2342
Hinkle, C. L.–(4) 2171
Hinzer, K.–(2) 578
Hioki, Tatsumi–(1) 7
Hirayama, Y.–(4) 2543
Hirsch, L. S.–(4) 2261
Hiyamizu, S.–(2) 575
Ho, Francis–(1) 43
Hobert, H.–(2) 540
Hobler, Gerhard–(1) 386
Hobson, W. S.–(3) 1398
Hoekstra, Robert J.–(4) 2102
Hoffmann, C.–(4) 1864
Höfling, E.–(4) 2562
Hoflund, Gar B.–(3) 1446
Hogan, T.–(3) 1385
Hogg, S.–(4) 1901
Hohenecker, St.–(4) 2317
Hoke, W. E.–(3) 1408
Holland, C. E.–(2) 758
Holmes, A. L., Jr.–(3) 1278
Hong, J. P.–(2) 736, 741
Hong, J. W.–(4) 2125
Hong, M.–(3) 1395, 1398
Hong, S. S.–(2) 736
Hong, Young Kyu–(2) 729
Hopcus, A. B.–(4) 2399
Hopwood, J.–(2) 523
Horst, Scott C.–(2) 511
Horton, T. U.–(3) 1377
Hoshi, Yoshinobu–(1) 247
Hosono, A.–(2) 780, 790
Hosono, Akihiko–(2) 799
Houzé, Frédéric–(4) 2006
Howell, M.–(2) 732
Hren, J. J.–(2) 716
Hsieh, K. C.–(3) 1352, 1395
Hsieh, W. Y.–(4) 2013
Hsieh, Y. F.–(4) 2013
Htoon, H.–(4) 2215
Hu, E. L.–(4) 2199
Hu, S.–(4) 1983
Huang, M.–(2) 796
Hughes, O. H.–(4) 2237
Huh, Chul–(1) 192
Hultman, L.–(4) 1885
Hunter, A. T.–(4) 1937
Huq, S. E.–(2) 796
Hwang, W.-Y.–(3) 1422
Hwang, Wen-Yen–(3) 1361
Hwang, Yong-Sup–(4) 2019
Hwa Shin, Young–(2) 712
Hwu, R. J.–(3) 1467

Iacona, Fabio–(2) 619
Iani, F.–(2) 665
Ibbotson, Dale E.–(1) 131
Igarashi, Masaru–(1) 38
Ihm, Jeong Don–(2) 916
Ikeda, J.–(2) 803
Ikeda, Junji–(2) 818
Ikeda, Kei–(1) 159
Ikoma, Hideaki–(1) 183
Ilegems, M.–(2) 582
Iltgen, K.–(1) 292, 298
Im, H.-J.–(4) 2302
Inglefield, C. E.–(4) 2328
Inoue, K.–(2) 895
Iriguchi, T.–(2) 790
Isakozawa, S.–(4) 2532
Ishikawa, J.–(2) 829, 833, 895
Ishikawa, Junzo–(4) 2515
Ishikawa, Yasuhiko–(4) 2387
Ishitani, T.–(4) 1907, 2532
Ito, Fuminori–(2) 783
Iwakuro, Hiroaki–(4) 1846
Iwano, Hirotaka–(4) 2551
Iwasaki, Yukio–(2) 720
Iyer, S. S.–(3) 1489

Jackman, Rebecca J.–(1) 59
Jacob, K.–(1) 406
Jacobs, E. S.–(3) 1430
Jahnel, Franz–(1) 386
Jain, H.–(1) 362
Jan, G. J.–(2) 570
Jang, Jin–(2) 705
Jang, Soon-Ju–(2) 858
Janz, S.–(4) 1773
Jarausch, K. F.–(4) 2494
Jaros, M.–(4) 1794
Jeffs, N. J.–(4) 2237
Jensen, J. E.–(3) 1321
Jensen, K. L.–(2) 749; (4) 2038
Jeon, D.–(2) 700, 826; (4) 2052
Jeon, Young Jin–(4) 1992
Jeong, S.-Y.–(4) 2125
Jernigan, G. G.–(4) 1933
Jessen, G. H.–(4) 2177
Jessing, J. R.–(2) 777
Jessop, P. E.–(4) 1773
Ji, H.–(2) 710
Ji, H.-F.–(1) 116
Jiang, H.–(2) 710
Jiang, H. X.–(4) 2215
Jiang, Z. X.–(1) 373; (4) 1971
Jin, C. C.–(2) 710
Jin, Jeong-Gu–(2) 754
Jin, Y. X.–(2) 710
Jo, Sung Ho–(4) 2086
Jo, W.–(4) 1891
Johnson, F. G.–(4) 1808
Johnson, J. L.–(3) 1321
Johnson, M. A. L.–(3) 1282
Johnson, M. B.–(3) 1367, 1463
Johnson, S. M.–(3) 1321
Johnson, Shane–(3) 1502
Johnson, W. B.–(4) 1808
Johnson, Walt–(1) 447
Johs, B.–(3) 1484
Jokerst, Nan Marie–(3) 1300
Jones, C.–(1) 327
Jones, G. A. C.–(4) 2570
Jones, T. S.–(4) 2373
Joo, Jung Hoon–(2) 500
Jørgensen, Jan Friis–(2) 633
Joubert, O.–(1) 147; (4) 1833
Joyce, B. A.–(4) 2373
Joyce, R. J.–(4) 2449
Ju, Byeong Kwon–(2) 705
Jubber, M. G.–(4) 1860
Jung, J. E.–(2) 741
Jung, J. H.–(2) 826
Jung, Jae Hoon–(2) 705
Jung, Jaehoon–(2) 920

Jurkovic, M. J.–(3) 1401

Kabir, A. E.–(4) 2118
Kaczer, B.–(4) 2302
Kahn, A.–(4) 2218
Kahrizi, M.–(2) 605
Kajiyama, Tisato–(1) 121
Kalbitzer, S.–(4) 2455
Kalburge, A.–(3) 1330
Kamanin, A. V.–(1) 426
Kamath, K.–(3) 1343, 1417
Kamino, T.–(4) 1907, 2532
Kampen, T. U.–(4) 2224, 2317
Kaneko, Reizo–(1) 38
Kang, C. J.–(2) 826
Kang, Sangbeom–(3) 1300
Kang, Seung-Oun–(4) 2079
Kang, Seung-Youl–(2) 773, 811, 815, 871
Kang, Sung Weon–(1) 238, 242
Kang, W. P.–(2) 684, 732
Kao, Hung-Chung–(1) 253
Kapolnek, D.–(2) 822
Katsuragi, Ken–(2) 553
Kawabuchi, S.–(2) 780, 790
Kawabuchi, Shinji–(2) 799
Kawasaki, Seiji–(2) 720
Keenan, J. A.–(1) 435
Kelber, J. A.–(1) 30
Keller, B. P.–(2) 822
Keller, S.–(2) 822
Kerns, D. V.–(2) 684, 732
Khater, Marwan H.–(2) 490
Khodaparast, G. A.–(3) 1367
Khor, K. E.–(4) 2417
Khuwattanasil, Noppadol–(1) 121
Kiefer, R.–(4) 1864
Kim, C. K.–(2) 684, 732
Kim, Choonkyung–(1) 430
Kim, D. C.–(4) 1891
Kim, D.-Y.–(4) 2125
Kim, Dong-Won–(4) 2019
Kim, Esther–(3) 1270
Kim, H.–(4) 2547
Kim, H. R.–(2) 777
Kim, Hoon–(2) 705
Kim, Hyunchul–(1) 430
Kim, Il Yong–(4) 1992
Kim, J.–(1) 344, 394
Kim, J. M.–(2) 736, 741; (4) 2082
Kim, J. W.–(2) 741
Kim, Jae Ki–(2) 729
Kim, Jeong Ho–(2) 500
Kim, Jin Seung–(2) 729
Kim, Jong-Hee–(2) 558
Kim, Jongmin–(1) 430
Kim, Joong Kyun–(2) 500
Kim, Ju-Jin–(2) 729
Kim, Jung Hun–(2) 500
Kim, Junghwan–(4) 1808
Kim, K. Y.–(4) 1891
Kim, S. H.–(4) 2052
Kim, Sang Gi–(2) 811
Kim, Shin-Sung–(4) 2086
Kim, Yeo Hwan–(2) 916
Kim, Young-Guon–(4) 2079
Kim, Yuon Tae–(2) 871
Kimura, K.–(4) 2342
King, Oliver–(2) 511
Kinokuni, M.–(4) 2484
Kinoshita, T.–(2) 496
Kirschbaum, U.–(3) 1385
Kitamura, S.–(1) 1
Kitayama, Yoshihiko–(2) 553
Kittelson, D. B.–(2) 483
Klaassen, J. J.–(3) 1286
Klem, J. F.–(3) 1498
Klussmann, S.–(4) 1864
Knoblauch, A.–(4) 2455
Ko, Tae-Young–(2) 700
Kobayashi, M.–(3) 1316
Kodis, M. A.–(4) 2038
Koelle, Ulrich–(3) 1502
Koguchi, Nobuyuki–(4) 2538
Koh, Ken Ha–(2) 724
Kohl, Paul–(3) 1300
Kohmoto, S.–(1) 1
Koike, H.–(4) 1907, 2532
Kolodziejski, L. A.–(3) 1381
Komuro, Masanori–(4) 2506
Kondo, Naoto–(1) 216
Kondoh, E.–(4) 2091
Kong, H. S.–(3) 1282
König, H.–(4) 2562
Konkar, A.–(3) 1334
Konuma, Kazuo–(2) 783
Koontz, E. M.–(3) 1381
Koops, H. W. P.–(2) 862
Kopanski, J. J.–(1) 339, 463
Kopf, R. F.–(4) 2110
Korakakis, D.–(4) 2237
Korakakis, Dimitris–(4) 2250
Korn, M.–(4) 2350
Kosai, K.–(3) 1321
Koshida, Nobuyoshi–(2) 793
Kotcheryzhenkov, A. V.–(1) 232
Kovacic, S. J.–(4) 1773
Koyama, Hideki–(2) 793
Kozawa, T.–(2) 833
Kraft, R.–(2) 496
Kropewnicki, Tom–(3) 1300
Kropfeld, P.–(2) 787
Krüger, D.–(1) 292
Kruit, P.–(4) 2063
Kuan, T. S.–(3) 1489
Kubena, R. L.–(4) 2449
Kudriavtsev, Yu. A.–(3) 1289, 1456
Kudryavtsev, Yu. A.–(1) 426
Kuk, Young–(2) 826
Kumar, B. R.–(3) 1467
Kumura, Yoshinori–(4) 2489
Künzel, H.–(1) 210
Kuo, C.-H.–(3) 1484
Kuo, Chau-Hong–(3) 1502
Kuo, H. C.–(3) 1377
Kuo, J. M.–(3) 1398, 1404
Kuramoto, Yasuyuki–(4) 2473
Kurihara, Kenji–(1) 69
Kuroda, Shinji–(4) 1846
Kurps, R.–(1) 292
Kurz, H.–(4) 1983
Kushner, Mark J.–(4) 2102
Kwak, Byung Man–(4) 1992
Kwan, W. S.–(2) 628
Kwo, J.–(3) 1395, 1398
Kwon, Sang Jik–(2) 712
Kwong, Dim-Lee–(1) 453
Kwun, S.-I.–(4) 2125
Kyong, T.–(3) 1484

Labanda, Jose Gregorio C.–(4) 1885
Lafontaine, H.–(2) 599; (4) 1773
Lai, Han-Chao–(1) 250
Lai, Li-Shyue–(1) 253
Lai, T. C.–(3) 1467
Lakatha, H.–(1) 420
Lambrinos, M. F.–(2) 589
Lampert, W. V.–(3) 1439, 1479
Landheer, D.–(2) 605
Landsberger, L. M.–(2) 605
Langan, J. G.–(1) 173
Langdo, T.–(3) 1471
Lange-Gieseler, Rosa–(1) 386
Langouche, G.–(4) 1901
Lantier, R.–(4) 2334
Lareau, R. T.–(3) 1467
Larson, Lawrence–(1) 447
Lau, S. S.–(3) 1297
Laursen, T.–(4) 1937
La Via, Francesco–(2) 619
Lazzarini, L.–(4) 2334
Lazzarino, M.–(4) 2334
Lebedev, A. B.–(3) 1289
Lee, B. S.–(1) 435
Lee, Byoungho–(2) 920
Lee, Chi H.–(4) 1808
Lee, Chun Gyoo–(2) 807
Lee, H. M.–(4) 1891
Lee, Ho-Jun–(2) 500
Lee, Hyung-Ik–(4) 2528
Lee, Hyung-Il–(2) 762
Lee, Hyun-Young–(4) 1987
Lee, J. Y.–(2) 700
Lee, Jin Ho–(1) 238, 242; (2) 773, 811, 815, 871
Lee, Jong Duk–(2) 712, 807, 866, 916, 920; (4) 2086
Lee, Jong Ho–(2) 807
Lee, Jong-Hyun–(2) 762
Lee, Jung-Hee–(2) 762
Lee, K. L.–(1) 382
Lee, Kun-Jing–(4) 1808
Lee, Kyeong Kyun–(3) 1300
Lee, Mark–(1) 447
Lee, N.–(4) 2052
Lee, Nam-Yang–(2) 724
Lee, P. P.–(3) 1467
Lee, R. Y.–(2) 855
Lee, S.–(3) 1270
Lee, S. R.–(3) 1498
Lee, S. W.–(4) 2052
Lee, Sang Yun–(1) 238; (2) 773
Lee, Seung-Yun–(4) 2019
Lee, Soonil–(2) 724
Lee, V.–(3) 1321
Lee, Won-Jun–(4) 2019
Lee, Yong-Il–(2) 815
Lee, Youn-Seoung–(4) 2019
Leech, Patrick W.–(1) 227
Lemonias, P. J.–(3) 1408
Lenahan, P. M.–(4) 2134
Lenox, C.–(3) 1426
Leon, R.–(2) 578
Lera, Jared D.–(2) 826
Lerner, Peter–(2) 900
LeTarte, L.–(1) 420
Li, L. K.–(3) 1275
Li, Q.–(2) 684, 732
Li, Y.–(1) 457
Lim, Dae Ho–(2) 558
Lim, M. H.–(3) 1381
Lin, C. H.–(3) 1377
Lin, C.-H.–(3) 1389
Lin, Chih-Hsiang–(3) 1435
Lin, J.-S.–(1) 30
Lin, J. Y.–(4) 2215
Linder, K. K.–(3) 1417
Lindner, K.–(4) 2355
Lindstrom, S. C.–(3) 1367
Littler, C.–(3) 1385
Liu, C. T.–(1) 334
Liu, Henley L.–(1) 415
Liu, R.–(1) 334
Liu, W. K.–(3) 1367, 1459
Loboda, M. J.–(3) 1305
Lohau, J.–(1) 77
Look, D. C.–(3) 1275
Lopes, D.–(1) 327
Lothian, J. R.–(3) 1404
Lothianand, J. R.–(3) 1398
Lowney, J. R.–(1) 463
Lu, J.–(4) 2254
Lu, Y.–(2) 697
Lübbe, M.–(4) 2355
Lubyshev, D. I.–(3) 1339, 1361
Lucovsky, G.–(4) 2154, 2171, 2177, 2191
Ludeke, R.–(4) 2296
Luftman, H.–(1) 334
Luo, E. Z.–(4) 1953
Luo, Yuanyuan–(3) 1312
Lynch, William–(1) 306

Ma, J. X.–(4) 1953
Ma, Xianyun–(2) 745
Macdonald, J. E.–(4) 2254
Mackie, W. A.–(4) 2057
Madelon, R.–(4) 1851
Madhukar, A.–(3) 1330, 1334
Maeda, Kiyoshi–(2) 836
Maex, K.–(1) 164; (4) 2091
Magee, Charles W.–(1) 286
Maher, Dennis M.–(1) 471
Mailhot, S.–(4) 1773
Mair, R.–(4) 2215
Mais, N.–(4) 2562
Makita, K.–(1) 1
Mannaerts, J. P.–(3) 1395, 1398
Mantese, L.–(4) 2367
Mao, J. M.–(1) 14
Marchiando, J. F.–(1) 339, 463
Marcus, M. A.–(3) 1395
Marqués, M. I.–(2) 654
Marques, P. V. S.–(4) 1860
Marsh, J. H.–(4) 1818
Marton, D.–(3) 1451
Marzin, J. Y.–(4) 1919
Mashkova, E. S.–(2) 678
Mason, B. F.–(2) 599
Massoud, H. Z.–(4) 2191
Masui, Moto-hisa–(4) 2426
Matney, K. M.–(3) 1381
Matsubara, T.–(4) 2313
Matsui, Shinji–(4) 2439
Matsukawa, T.–(4) 2479
Matsukawa, Takashi–(4) 2489
Matsumoto, H.–(4) 2532
Matthai, C. C.–(4) 2358
Matthews, M. R.–(4) 2057
Matyi, R. J.–(1) 435
May, Gary–(3) 1300
Mayer, J. W.–(4) 1937
Mayo, W. E.–(3) 1404
Mazur, R. G.–(1) 316, 411
McCann, P. J.–(3) 1459, 1463
McCarson, B. L.–(2) 689
McClure, M. T.–(2) 689
McDaniel, D. L., Jr.–(3) 1435
McDermott, B. T.–(4) 2286
McDonald, Andrew J.–(1) 453
McGill, T. C.–(4) 2182, 2291
McGlothlin, R.–(1) 476
McLaughlin, A. J.–(4) 1860
McMurray, J. S.–(1) 344, 394
McMurry, P.–(2) 483
McNeil, John R.–(1) 80
McNulty, H. H.–(4) 2449
Mehregany, Mehran–(2) 536
Menzel, R.–(2) 540
Metzger, Robert A.–(3) 1300, 1326
Mhetar, Vijay R.–(4) 2121
Micovic, M.–(3) 1339, 1361
Midha, A.–(1) 54
Midzor, M. M.–(4) 2275
Miller, D. L.–(3) 1339, 1361

Mills, G.–(1) 54
Mimura, H.–(2) 803
Mimura, Hidenori–(2) 818
Mimura, R.–(4) 2484
Miner, C.–(4) 1790
Minhas, Babar K.–(1) 80
Mintairov, A. M.–(3) 1456
Mishima, Tetsuya–(4) 2324, 2426
Mishra, U. K.–(2) 822
Miskovsky, N. M.–(2) 900, 906
Mitchell, I. V.–(4) 1790
Mitchell, W. J.–(4) 2199
Miwa, S.–(4) 2342
Miyamoto, Iwao–(4) 2506
Miyamoto, Y.–(2) 851
Mizumura, Michinobu–(4) 2555
Mo, D.–(2) 697
Mogren, S.–(4) 2469
Mokina, L. A.–(1) 426
Molchanov, A. V.–(2) 678
Möller, H.–(3) 1507
Mollot, F.–(1) 223
Mönch, W.–(4) 2224
Monemar, B.–(4) 1928
Monget, C.–(4) 1833
Morante, J. R.–(4) 1851
Morimoto, H.–(2) 780, 790
Morimoto, Hiroji–(2) 799
Morkoc, H.–(2) 507
Moser, B. G.–(3) 1377
Motohiro, Tomoyoshi–(1) 7
Moustakas, Theodore D.–(4) 2250
Moy, A. M.–(3) 1347, 1352
Mukhametzhanov, I.–(3) 1330
Müller, B.–(1) 406; (4) 2334
Müller, G.–(2) 693
Müller, S.–(4) 1864
Murad, S. K.–(1) 54
Murase, K.–(4) 2313
Murphy, M. J.–(4) 1943
Murry, S. J.–(3) 1435
Murthy, C. S.–(1) 440
Müssig, H.–(4) 2562
Myburg, G.–(4) 1873
Myers, T. H.–(4) 2261

Na, Seong Jun–(2) 724
Na, Zhiyong–(1) 471
Nagamachi, Shinji–(4) 2515
Nagao, M.–(2) 829, 895
Nagase, Masao–(1) 69
Nahm, Sahn–(2) 858
Nakada, Yoshinobu–(2) 645
Nakamoto, Masayuki–(2) 770
Nakamura, Jun–(4) 2324, 2426
Nakamura, Ryuichi–(1) 183
Nakamura, S.–(3) 1316
Nakamura, Shuji–(4) 2204
Nakane, Hideaki–(2) 875
Nakano, Shizuka–(4) 1914
Nakayama, M.–(4) 2313
Nakayama, Masayoshi–(4) 2511
Nam, Jung Hyun–(2) 916
Namatsu, Hideo–(1) 69
Nambu, Y.–(1) 1
Nannini, A.–(2) 665
Napolitani, E.–(4) 2334
Narayanan, S.–(2) 613
Nassar, R.–(1) 109
Nassar, Raja–(4) 2499
Nau, D.–(2) 693
Neo, Y.–(2) 803
Neogi, Suneeta S.–(1) 471
Neubauer, G.–(1) 394
Neugebauer, J.–(4) 2242
Ni, W. X.–(4) 1928
Nicolaescu, D.–(2) 888
Nie, H.–(3) 1426
Niedermann, Ph.–(1) 355
Niehuis, E.–(1) 298
Nihei, Yoshimasa–(4) 2473
Niimi, H.–(4) 2177
Nijhawan, S.–(2) 483
Nikishin, N. V.–(4) 2082
Nikishin, S. A.–(3) 1289, 1456
Nimmagadda, S. V.–(2) 613
Ning, Y. Q.–(2) 710
Nishii, Kiyoaki–(2) 651
Nishikawa, Osamu–(2) 836
Nishizawa, Masayasu–(4) 2324, 2426
Noda, Shoji–(1) 7
Noda, T.–(4) 2547
Noël, Sophie–(4) 2006
Noguchi, Kazumi–(1) 38
Nomura, Setsuo–(1) 104
Nordgren, Joseph–(4) 2250
Northrup, J. E.–(4) 2242
Nosho, B. Z.–(4) 2381, 2399
Nozawa, Hiroshi–(2) 515
Numada, Yoshihiro–(2) 836
Nutsch, A.–(4) 1864
Nxumalo, J. N.–(1) 457
Nylandsted Larsen, A.–(4) 1777
Nys, J. P.–(1) 137

Oehrlein, G. S.–(4) 1998
Ogiso, Hisato–(4) 1914
Oh, Chang Woo–(2) 807
Oh, Myung Hwan–(2) 705
Ohdomari, I.–(4) 2479
Ohdomari, Iwao–(4) 2489
Ohno, Naoto–(4) 2506
Ohta, K.–(2) 575
Ohtake, A.–(4) 2342
Ohtani, Yoshikatsu–(2) 836
Ojha, J. J.–(4) 1773
Okabe, Jun–(4) 2489
Okamoto, Akihiko–(2) 783
Okamoto, Kazuya–(2) 818
Okamoto, Y.–(2) 575
Oki, A. K.–(3) 1475
Okumura, Hajime–(2) 645
Okuyama, F.–(2) 888
Olson, S. R.–(3) 1475
Ono, Tomio–(2) 770
Ooi, B. S.–(4) 1818
Orchard-Webb, J.–(4) 1881
Ormsby, T. J.–(1) 302
Osaka, Toshiaki–(4) 2324, 2426
Osburn, Carlton M.–(1) 286
Oshima, K.–(2) 851
Oshita, Yoshio–(1) 159
Otter, F. A.–(4) 2127
Overzet, Lawrence J.–(2) 490
Owari, Masanori–(4) 2473
Owen, J. H. G.–(4) 2399

Paggel, J. J.–(4) 2334
Palmström, C. J.–(4) 2280
Pang, S. W.–(2) 765
Panitz, J. A.–(1) 23
Paraskevopoulos, A.–(1) 210
Park, Byung Gook–(2) 807, 866
Park, Chan–(2) 729
Park, Chong-Ook–(4) 2019
Park, D.–(2) 507
Park, Dong-Il–(2) 762
Park, Gun-Sik–(2) 754
Park, J. M.–(2) 906
Park, Jeong-Young–(2) 826
Park, Jung Il–(2) 724
Park, Kwang Ja–(2) 724
Park, Kyung Ho–(2) 724
Park, Miri–(2) 544
Park, N. S.–(2) 741
Park, Seong-Ju–(1) 192
Park, Soon-Soo–(2) 762
Parker, D. L.–(2) 777
Parker, R. K.–(2) 749
Parkinson, B.–(3) 1456
Pastor, C. J.–(4) 1804
Patten, E. A.–(3) 1321
Pattyn, H.–(4) 1901
Paul, Kateri E.–(1) 59, 88
Pawlik, Marek–(1) 286
Pearson, Peter–(1) 386
Pei, S. S.–(3) 1389, 1435
Pelz, J. P.–(4) 2302
Pelzel, R. I.–(4) 2399
Penczek, J.–(2) 855
Peng, S. Q.–(2) 697
Peng, Y. C.–(4) 2013
Pérez-Rodríguez, A.–(4) 1851
Perovic, D. D.–(2) 599; (4) 2188
Perry, S. S.–(3) 1270
Pesant, Jean Claude–(1) 255(E)
Petrich, G. S.–(3) 1381
Phillips, J.–(3) 1343
Phillips, J. R.–(4) 2494
Pickrell, G. W.–(3) 1347
Pierson, R. L.–(4) 2286
Pilione, L.–(2) 681
Piotrowicz, S.–(2) 561
Piotto, M.–(2) 665
Piqueras, J.–(4) 1804
Pittman, R.–(4) 2286
Piva, P. G.–(4) 1790
Plavitu, C. N.–(2) 888
Pletschen, W.–(4) 1864
Ploog, K. H.–(4) 2404
Ploog, Klaus H.–(4) 2229
Polezhaev, S. A.–(1) 232
Pollak, Fred H.–(3) 1312
Polley, C.–(3) 1286
Porte, L.–(4) 1786
Posselt, M.–(1) 440
Pozina, G.–(4) 1928
Prewett, P. D.–(2) 796
Priester, Catherine–(4) 2421
Proost, J.–(4) 2091
Puttock, M.–(1) 147

Qin, Dong–(1) 98
Qiu, B. C.–(4) 1818

Rack, M. J.–(4) 2165
Radelaar, S.–(1) 373
Raineri, Vito–(2) 619
Rakhshandehroo, M. R.–(2) 765
Ramachandran, T. R.–(3) 1330, 1334
Ramam, A.–(2) 565
Rao, N. P.–(2) 483
Raymond, Christopher J.–(1) 80
Reetz, M. T.–(1) 77
Reeves, Geoffrey K.–(1) 227
Register, Richard A.–(2) 544
Reifsnider, J. M.–(3) 1278
Reithmaier, J. P.–(4) 2562
Ren, F.–(3) 1398
Ressel, Peter–(1) 227
Reynolds, D. C.–(3) 1275
Rha, Sa-Kyun–(4) 2019
Richards-Babb, M. R.–(4) 2261
Richter, Lee J.–(4) 1948
Richter, W.–(4) 2350, 2358
Rietman, Edward A.–(1) 131
Rimai, L.–(1) 142
Rimini, Emanuele–(2) 619
Ringel, S. A.–(3) 1372, 1471
Rizk, R.–(4) 1851
Robach, Y.–(4) 1786
Robey, S. W.–(4) 2413
Robillard, M.–(4) 1773
Rockett, A.–(2) 507
Rogers, John A.–(1) 59, 88
Rolfe, S. J.–(2) 599
Romanato, F.–(4) 2334
Romano, L. T.–(4) 2261
Romig, Terry–(1) 447
Ronsheim, P. A.–(1) 382
Rose, P. D.–(4) 2570
Rossow, U.–(4) 2350, 2355, 2367
Roth, J. A.–(3) 1413
Roth, John–(3) 1502
Roukes, M. L.–(4) 2275
Routh, B. P., Jr.–(4) 2057
Rubini, S.–(4) 2334
Rueger, N. R.–(4) 1998
Rumyantsev, S.–(4) 1881
Russell, P. E.–(4) 2494
Ryan, R. W.–(4) 2110
Ryu, Y. S.–(2) 736

Sabin, E. W.–(4) 1841
Sacks, R. N.–(3) 1372
Sadakane, S.–(2) 895
Sadra, K.–(1) 116
Sadwick, L. P.–(3) 1467
Saito, Yasushi–(2) 875
Saitoh, Keiya–(4) 2555
Saitoh, W.–(2) 851
Saka, H.–(4) 2522
Sakaguchi, Kiyoshi–(4) 2462
Sakai, Tadashi–(2) 770
Sakai, Takamasa–(4) 2159
Sakaki, H.–(4) 2547
Sakamoto, Tetsuo–(4) 2473
Saku, T.–(4) 2543
Sakuma, Naoshi–(2) 770
Salviati, G.–(4) 2334
Samavedam, S. B.–(3) 1471
Santos, M. B.–(3) 1367, 1459
Sawaragi, H.–(4) 2484
Sawayanagi, Mineo–(4) 2426
Sayedi, S. M.–(2) 605
Schäfer, J.–(4) 2177
Schaff, W. J.–(4) 1943
Schelhase, S.–(1) 210
Schenk, Andreas–(4) 2296
Schetzina, J. F.–(3) 1282
Schiltz, A.–(4) 1833
Schlaf, R.–(3) 1456
Schlesser, R.–(2) 689, 716
Schneider, A.–(4) 2254
Schneider, M.–(4) 1919
Schneir, Jason–(2) 633
Schoenborn, Phillipe–(4) 2102
Schoessler, C.–(2) 862
Schreck, M.–(2) 693
Schuler, O.–(1) 223
Schuler, Olivier–(1) 255(E)
Schulte, D. W.–(3) 1356
Schultz, Ch.–(4) 2350
Schultz, J. A.–(3) 1270
Schwall, D.–(4) 2127
Schwartz, Daniel K.–(1) 51
Schwartzman, A. F.–(3) 1293
Schwoebel, P. R.–(2) 758
Sciortino, P. F.–(3) 1422
Seaford, M. L.–(3) 1479
Sekine, Takahiro–(2) 836
Sekine, Tetsu–(4) 2462
Semendy, F.–(3) 1489
Seo, Jae Myung–(1) 192

Sergent, A. M.–(3) 1395
Seryogin, G. A.–(3) 1456
Shanabrook, B. V.–(4) 2381
Sharp, L. I.–(3) 1456
Shaw, M. J.–(4) 1794
Shen, Jeng-Jung–(3) 1326
Sheng, Xia–(2) 793
Sherbondy, J. C.–(1) 411
Shi, Frank F.–(1) 125
Shi, X.–(4) 2049
Shibata, Tomohiro–(2) 515
Shih, A.–(4) 2038
Shih, C. K.–(4) 2215
Shih, Chih-Kang–(1) 453
Shimawaki, H.–(2) 803
Shimizu, Ryuichi–(4) 2528
Shimomura, S.–(2) 575
Shin, M. S.–(4) 2242
Shinada, T.–(4) 2479
Shinada, Takahiro–(4) 2489
Shirafuji, Junji–(2) 720
Shmakov, S. L.–(4) 2082
Shmidt, N. M.–(1) 426
Shubina, T. V.–(3) 1289
Shuman, V. B.–(1) 426
Sidorov, M.–(4) 2165
Sieg, R. M.–(3) 1471
Sievers, Barry–(3) 1326
Silvestre, C.–(4) 1933
Simionescu, Alexander–(1) 386
Singh, R.–(2) 613
Singh, Raj–(4) 2250
Sitar, Z.–(2) 689, 716
Sivananthan, S.–(3) 1309
Skowronski, M.–(4) 2242
Slinkman, J.–(1) 344
Smith, A. R.–(4) 2242
Smith, Arthur R.–(1) 453
Smith, David J.–(4) 1937
Smith, H. E.–(1) 339, 420
Smith, H. I.–(3) 1381
Smith, Kevin E.–(4) 2250
Smith, R. P.–(4) 2286
Sobiesierski, Z.–(4) 2358
Sobolewski, M. A.–(1) 173
Sohail H. Naqvi, S.–(1) 80
Solére, A.–(4) 1786
Song, Kug-Hyun–(2) 724
Song, Yoon-Ho–(1) 238; (2) 773, 811, 815, 871
Sorba, L.–(4) 2334
Sota, Takayuki–(4) 2204
Sou, I. K.–(1) 14
Spangenberg, B.–(4) 1983
Spindt, C. A.–(2) 758
Stagarescu, Cristian B.–(4) 2250
Stanchina, W. E.–(3) 1413
Stark, T. J.–(4) 2494
Steckl, A. J.–(3) 1305
Steimetz, E.–(4) 2358
Stepanova, A. N.–(2) 678
Stevens, K. S.–(3) 1293
Stiévenard, D.–(1) 137
Stillman, G. E.–(3) 1377
Stock, S.–(3) 1300
Stone, Dennis–(2) 511
Storm, W.–(1) 164
Stranick, Stephan J.–(4) 1948
Strathman, M. R.–(1) 435
Stratton, F. P.–(4) 2449
Strausser, Yale–(2) 633
Strecker, B. N.–(3) 1459
Streetman, B. G.–(3) 1278, 1426
Streetman, Ben G.–(1) 453
Streit, D. C.–(3) 1475
Stritzker, B.–(2) 693
Strunk, H. P.–(1) 19
Stutzmann, M.–(4) 2224
Su, D.–(1) 327
Subramanian, S.–(3) 1356
Sudarshan, T. S.–(2) 745
Sugimoto, Y.–(1) 1
Sugino, Takashi–(2) 720
Suh, B. J.–(4) 2275
Suh, Kyung-Soo–(2) 858
Sukharev, Valeriy–(4) 2102
Sullivan, G. J.–(4) 2286
Summers, C. J.–(2) 855
Sun, Z.–(4) 2049
Suzuki, M.–(2) 833
Suzuki, N.–(2) 780

Tabara, Suguru–(2) 553
Taferner, W. F.–(3) 1479
Taga, Y.–(2) 833
Tahara, K.–(2) 803
Tahara, Kaoru–(2) 818
Takahara, Atsushi–(1) 121
Takahashi, K.–(3) 1316
Takahashi, Masanori–(4) 2473
Takahashi, Yasuo–(1) 69
Takai, M.–(2) 780, 790
Takai, Mikio–(2) 799
Takanashi, Yoshifumi–(1) 216
Takaoka, S.–(4) 2313
Takeda, Shuuichi–(4) 2506
Takeda, Yasuhiko–(1) 7
Takekawa, K.–(2) 575
Takihara, Hirotaka–(2) 670
Tamamura, Toshiaki–(2) 515
Tamargo, Maria C.–(3) 1312
Tanabe, Masafumi–(2) 553
Tanaka, Masaaki–(4) 2267
Taniguchi, Jun–(4) 2506
Taniguchi, Y.–(4) 2532
Tate, A.–(4) 2110
Tay, B. K.–(4) 2049
Taylor, P. C.–(4) 2328
Teichert, J.–(4) 2574
Temkin, H.–(3) 1289, 1456
Tennant, D. M.–(4) 2113
Teschke, O.–(4) 2105
Tételin, C.–(1) 137
Théron, D.–(1) 223
Thomas, D. R.–(3) 1439
Thomas, S.–(3) 1377
Thompson, Daniel W.–(1) 312
Thompson, P. E.–(4) 1933
Thoms, S.–(1) 54
Thomson, D. J.–(1) 457
Tiemeijer, P. C.–(4) 2063
Ting, D. Z.-Y.–(4) 2182
Ting, S. M.–(3) 1471
Tittelbach-Helmrich, K.–(1) 406
Tiwald, Thomas E.–(1) 312
Toda, Akitoshi–(1) 38
Tokumoto, Hiroshi–(4) 1914
Tomich, D. H.–(3) 1439, 1479
Tomita, N.–(2) 575
Torabi, A.–(3) 1408
Tormen, B.–(4) 1833
Trampert, Achim–(4) 2229
Tränkle, G.–(4) 1864
Trassaert, S.–(2) 561
Trenkler, T.–(1) 349, 367, 394
Tsai, H. S.–(3) 1404
Tseng-Chung Lee, J.–(1) 131
Tsuji, H.–(2) 829, 895
Tsukada, Kenji–(4) 2324
Tsurumi, Naohiro–(4) 2387
Tu, C. W.–(3) 1297, 1479; (4) 2395
Turban, G.–(4) 1823
Twigg, M. E.–(4) 1933

Ueda, Masahiro–(4) 2515
Uh, Hyung Soo–(2) 866, 916
Ukraintsev, V. A.–(1) 476
Umemura, Shigeru–(1) 38
Underwood, Robert D.–(2) 822
Ungier, L.–(3) 1356
U'Ren, G. D.–(3) 1381

Vajo, J.–(4) 2449
Vajo, J. J.–(4) 1937
Valizadeh, R.–(2) 589
Vallier, L.–(1) 147; (4) 1833
van Berkum, J. G. M.–(1) 280, 298
Vandervorst, W.–(1) 164, 260, 320, 349, 355, 367, 394, 401
Vanhaelemeersch, S.–(1) 164
Van Hove, J. M.–(3) 1286
Vanoverschelde, A.–(2) 787
van Rooy, Th. L.–(4) 2063
Vantomme, A.–(4) 1901
Vasile, M.–(1) 109
Vasile, Michael J.–(4) 2499
Vasileska, D.–(4) 2165
Vedagarbha, V.–(2) 613
Venables, D.–(1) 362
Venables, David–(1) 471
Venkateswaran, U.–(3) 1356
Venzor, G.–(3) 1321
Veprek, Stan–(1) 19
Vereecke, G.–(4) 2091
Versen, M.–(4) 2567
Vieu, C.–(4) 1919
Vigil, J. A.–(3) 1321
Visser, C. C. G.–(1) 373
von Criegern, Rolf–(1) 386

Wada, Kazumi–(4) 2204
Wagner, B. K.–(2) 855
Wagner, Dennis–(1) 447
Wakao, K.–(3) 1316
Wakaya, F.–(4) 2313
Wakaya, Fujio–(4) 2511
Walck, S. D.–(3) 1439
Waldman, J.–(3) 1430
Wallace, W. E.–(4) 1958
Wallart, X.–(1) 137
Wan, Junzuo–(3) 1312
Wang, F.–(4) 2171
Wang, L.–(1) 435
Wang, M. T.–(4) 2026
Wang, P. C.–(4) 2026
Wang, Qiming–(4) 1901
Wang, W.–(2) 532
Wang, W. B.–(2) 710
Wang, W. I.–(3) 1275, 1401, 1489
Wang, Wei–(1) 415
Wang, Y. C.–(3) 1404
Wang, Y. Z.–(2) 710
Wang, Z.–(2) 507
Washington, D.–(3) 1385
Wassermann, E. F.–(1) 77
Watanabe, Masafumi–(2) 836
Waters, K.–(3) 1270
Weaver, J. M. R.–(1) 54
Wedel, M.–(2) 693
Weemers, K.–(1) 280, 298
Wei, A. X.–(2) 697
Weichold, M. H.–(2) 777
Weimann, G.–(4) 1864
Weimer, M.–(3) 1389
Weinberg, W. H.–(4) 2199, 2381, 2399
Weinzierl, S. R.–(1) 411
Weiss, B. L.–(2) 681
Weisser, S.–(4) 1864
Wen, H. J.–(4) 2296
Wen, Y. C.–(3) 1456
Wendt, A. E.–(2) 532
Werner, W. S. M.–(1) 420
Wesch, W.–(2) 540
Westwood, D. I.–(4) 2254, 2358
Whang, Chung-Nam–(4) 2019
Whang, Ki-Woong–(2) 500
Whitesides, George M.–(1) 59, 88, 98
Whitman, L. J.–(4) 2381
Wieck, A. D.–(4) 2567
Wiemann, C.–(4) 2567
Wijewarnasuriya, P. S.–(3) 1309
Wildt, D.–(4) 1804
Wilkinson, C. D. W.–(4) 1818, 1860
Williams, C. C.–(1) 344, 394
Williams, R. L.–(4) 1773
Willis, R. F.–(3) 1339
Wilshaw, P. R.–(2) 796
Wilson, I. H.–(1) 14; (4) 1953
Wilson, Rick–(2) 511
Winter, M.–(1) 77
Wisitsora-at, A.–(2) 684, 732
Wittmaack, K.–(1) 272
Wohlert, D. E.–(3) 1352
Wojak, G.–(2) 716
Wojtowicz, M.–(3) 1475
Wolan, John T.–(3) 1446
Wolfe, D. M.–(4) 2171
Wolfframm, D.–(4) 2350
Wong, V. V.–(3) 1381
Wood, M. C.–(3) 1467
Woodward, John T.–(1) 51
Woollam, John A.–(1) 312
Wormington, M.–(3) 1381
Wowchak, A. M.–(3) 1286
Wu, C. I.–(4) 2218
Wu, M. F.–(4) 1901
Wu, O. K.–(3) 1321
Wu, W. L.–(4) 1958
Wu, Z.–(2) 483

Xia, Younan–(1) 98
Xie, Jushan–(4) 2499
Xie, Q.–(3) 1330
Xie, Tianbao–(4) 2057
Xu, J. B.–(1) 14; (4) 1953
Xu, J. F.–(2) 684, 732
Xu, N. S.–(2) 697

Yaguchi, T.–(4) 1907, 2532
Yamada, Atsushi–(2) 818
Yamaguchi, A.–(2) 851
Yamaguchi, Hiroshi–(4) 2555
Yamaguchi, Toru–(1) 69
Yamamoto, Yoshihisa–(1) 43
Yamanaka, Kazushi–(4) 1914
Yamazaki, Kenji–(1) 69
Yan, Chi–(3) 1435
Yan, X.–(4) 1953
Yanagisawa, J.–(4) 2313
Yanagisawa, Junichi–(4) 2511
Yang, B. H.–(3) 1435
Yang, Bin–(4) 2229
Yang, F.–(2) 578
Yang, Gye Mo–(2) 558
Yang, H.–(4) 2191
Yang, Hui–(4) 2229
Yang, Kwangsun–(1) 430
Yang, Qinqing–(4) 1901
Yang, Rui Q.–(3) 1435
Yang, Y. R.–(4) 2013
Yang, Z.–(3) 1489
Yano, Akihiro–(2) 783
Yao, Shude–(4) 1901
Yao, T.–(4) 2342
Yap, K. S.–(3) 1372
Yaradou, O.–(2) 787

2584

Yasuda, T.–(4) 2342
Yasuda, Yukio–(4) 2551
Yater, J. E.–(4) 2038
Yevgen'ev, A. O.–(1) 232
Yi, S. I.–(4) 2199
Yocom, P. N.–(2) 855
Yokoo, K.–(2) 803
Yokoo, Kuniyoshi–(2) 818
Yokota, Yuuko–(2) 720
Yoo, H. M.–(3) 1356
Yoo, Hyung Joun–(1) 238, 242; (2) 773, 811, 815, 871
Yoo, Jae Soo–(4) 2086
Yoon, B.-G.–(2) 906
Yoshida, Hiro–(1) 247
Yoshida, Toshiyuki–(4) 2159
Yoshikawa, A.–(3) 1316
Yoshikawa, T.–(1) 1
Young, A. P.–(4) 2177
Yu, Byoung Gon–(1) 242
Yu, E. T.–(4) 2395
Yu, Hongbin–(4) 2215
Yuan, G.–(2) 710
Yuan, S.–(3) 1459
Yuba, Y.–(4) 2313
Yurre, T. A.–(1) 426

Zahn, D. R. T.–(4) 2254, 2317, 2355
Zaidman, E. G.–(4) 2038
Zaima, Shigeaki–(4) 2551
Zaknoune, M.–(1) 223
Zborowski, J. T.–(3) 1451
Zeng, K.–(4) 2215
Zettler, T.–(4) 2358
Zhang, B. L.–(2) 710
Zhang, D.–(3) 1389
Zhang, F. L.–(2) 855
Zhang, Kai–(3) 1361
Zhang, W.–(1) 109
Zhang, Y.-H.–(3) 1484, 1507
Zhang, Yong-Hang–(3) 1502
Zhang, Z.–(3) 1451; (4) 2275
Zhao, Y.–(3) 1297
Zheng, X. G.–(2) 697
Zhirnov, V. V.–(2) 678
Zhong, G.–(2) 523
Zhou, H.–(1) 54
Zhou, Li–(1) 447
Zhou, T. M.–(2) 710
Zhou, Wei–(1) 227
Zhou, X.–(3) 1343
Ziemann, P.–(2) 483
Zinck, J. J.–(3) 1413; (4) 2381
Zinke-Allmang, M.–(4) 2188
Zorman, Christian A.–(2) 536
Zoulkarneev, A.–(2) 741
Zubrilov, A. S.–(3) 1289
Zuo, S. L.–(4) 2395
Zypman, Fredy R.–(4) 2099

AMERICAN VACUUM SOCIETY

Under United States copyright law, a transfer of copyright from the author(s) must be explicitly stated in writing to enable the publisher to disseminate the work to the fullest extent. The following transfer agreement must be completed, signed and returned to the Editor's office before the manuscript can be accepted for publication. Further information is available from the Office of Rights and Permissions, American Institute of Physics (AIP), 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Tel: (516) 576-2268; Fax: (516) 576-2499; Internet: rights@aip.org.

TRANSFER OF COPYRIGHT AGREEMENT

Copyright in the unpublished and original article, including the abstract forming part thereof, entitled _____

_____

_____

Title of Article

(the "Article"), submitted by the following author(s) [the "Author(s)"] _____

_____

_____

Names of all Authors

is hereby assigned and transferred to the American Vacuum Society ("AVS") for the full term thereof throughout the world, subject to the terms of this Agreement and to acceptance of the Article for publication in _____

_____

Name of Journal

AVS shall have the right to publish the Article in any medium or form, or by any means, now known or later developed. AVS shall have the right to register copyright to the Article in its name as claimant whether separately or as part of the journal issue or other medium in which the Article is included. The Author(s) reserve all proprietary rights other than copyright, such as patent rights. If the Article was prepared under a United States Government contract, the Government shall have rights in the copyright to the extent required by the contract, and the Author(s) shall notify AVS of any such rights. The Author(s) represent and warrant: (1) that the Article is original with them; (2) that the Article does not infringe any copyright or other rights in any other work, or violate any other rights; (3) that the Author(s) own the copyright in the Article or are authorized to transfer it; and (4) that all copies of the Article the Author(s) make or authorize will include a proper notice of copyright in AVS's name. If each Author(s) signature does not appear below, the signing Author(s) represent that they sign this Agreement as authorized agents for and on behalf of all the Authors, and that this Agreement and authorization is made on behalf of all the Authors.

_____

Name (print) Date

_____

Signature

If the manuscript has been prepared as a work for hire, this Agreement should be signed by the employee (or commissioned party) (**above**) and by the employer (or commissioning party) (**below**). In such event, the employer (or commissioning party) agrees to be bound by all of the obligations of the Author(s) under this Agreement.

_____

Name of Employer (or Commissioning Party) (print)

_____

Signature Title Date

A work prepared by a U.S. Government officer or employee as part of his or her official duties is not eligible for U.S. copyright; however foreign copyright laws may differ. Thus this form should be signed even by U.S. Government officers or employees. Signing of this form will not affect U.S. Copyright law provisions in the case of works of the U.S. Government. **If all the authors are in this category, check the box here and return the signed form.** ☐

See page 2 (over) for rights granted by AVS to authors.

TRANSFER OF COPYRIGHT (page 2)

AVS hereby grants to the Author(s) a nonexclusive license to do the following:

(a) After publication of the article by AVS, to publish all or part of the Article in any compilation or publication limited to the Author(s)' own works, or in any compilation or publication approved by AVS, and to make copies of the Article for the Author(s)' own use for lecture or classroom purposes;

(b) To make an oral presentation of the Article;

(c) If the Article was prepared as a work made for hire, to allow the employer or commissioning party to make copies of the Article for its own internal use.

AVS and the Author(s) further agree that, after publication of the Article by AVS, third parties wishing to republish the Article must obtain written permission from both the Author(s) and AVS; provided, however, that only AVS's permission is necessary with respect to republication of entire journal issues which contain the Article.

Please sign the applicable portion of the form and return it promptly to the Editor's Office, Journal of Vacuum Science and Technology, 10 Park Plaza, Ste. 4A, Caller Box 13994, Research Triangle Park, NC 27709-3994. We cannot begin the production process until this signed form has been received. Please help us to expedite the publication of your paper with your immediate response.

JOURNAL OF VACUUM SCIENCE AND TECHNOLOGY
INFORMATION FOR CONTRIBUTORS

1. Articles. This journal publishes two types of articles: original research reports comprising a complete, detailed, self-contained description of a research effort and Review papers that provide an updating of an area of science or technology pertinent to the scope of this journal. All articles must contain an abstract (see item **5**).

2. Brief Reports and Comments. These are short submissions that either report a brief item of research or technology that does not require extensive descriptive writing or that comment on an item previously published in the journal. Brief Reports and Comments do not require the format structure of regular journal articles and do not require an abstract.

3. Rapid Communications. These are short submissions which contain important new results and are intended for accelerated publication. Rapid Communications are given priority in handling and do not require an abstract.

4. Shop Notes. These are "how to do it" papers. They should be written and illustrated so that the reader may easily follow whatever instruction or advice is being given. An abstract is not required for a Shop Note.

5. Abstracts. An *abstract must accompany each article.* The abstract should be typed (double-spaced) *on a separate page.* It should be adequate as an index and as a summary. It should give all subjects, major and minor, concerning which new information is presented. It should give the conclusions of the article and all numerical results of general interest. An abstract is usually reprinted verbatim in abstract journals. Therefore, great care should be used in writing it.

6. Manuscripts. Manuscripts are accepted with the understanding that they have not been copyrighted, published, or accepted for publication elsewhere. Manuscripts must be in English, typewritten, double-spaced on one side of the page, on $8\frac{1}{2}\times 11$ in. durable, opaque, white paper. *An original and two* (2) *copies must be submitted.* Wide margins should be provided to permit editorial instructions to the printer. *Authors should adopt a particular way of writing their names in the "by-line" and use the same name for all publications.* This practice makes indexes more useful and less confusing. For submission of electronic files for publication see item **15** below.

7. Style. Authors are expected to follow the conventional writing, notation, and illustration style prescribed in the American Institute of Physics *Style Manual* which is available from the AIP. For ordering information, visit http://www.aip.org/epub/styman.html. Authors should also study the form and style of printed material in this journal. The SI units should be used.

8. Illustrations: Illustrations published in the journal are either scanned by AIP using a digital scanner or received electronically from the author, and integrated with the text of the article, creating completely electronic pages. To receive optimal quality, we strongly encourage you to send electronic graphics files to AIP, rather than laser output. (Note: If you are submitting electronic graphics files, you are still required to send hardcopies of the figures to the Editorial Office. Adherence to electronic submission instructions is crucial. If your electronic files are received after AIP has already processed the hardcopy illustrations, the electronic files will not be used.) Please adhere to the following guidelines when preparing your illustrations for submission:

Sizing Illustrations (Electronic Graphics Files and Hardcopy)

- **Prepare illustrations in the final published size, not oversized or undersized.** Size your illustrations according to your journal's specifications. Submit each illustration at the final size in which it will appear in the journal. The standard is 8.5-cm maximum width (3-3/8″ or 20.5 picas) for one column. This is especially important for screened or shaded illustrations; reduction of screened/shaded originals during the digitizing process introduces an unacceptable moiré pattern.
- **Ensure a minimum of 8-point type size (2.8 mm high; 1/8″ high) and 1-point line width within illustrations**. Ensure that line weights will be **0.5 points** or greater in the final published size. **Line weights below 0.5 points will reproduce poorly.** Avoid inconsistencies in lettering within individual figures, and from one figure to the next. Lettering and symbols cannot be handwritten. Avoid small open symbols that tend to fill in if any reduction is necessary.

Preparation of Hardcopy Illustrations for Scanning

- **Number figures in the order in which they appear in text.** Label illustrations with their number, the name of the first author, and the journal, on the **front** of the figure well outside the image area.
- **Place only one figure per page (including all parts).** Place all parts of the same figure on one sheet of white bond paper, spaced 1/4 in. apart, using a glue stick or wax on the back of the illustration and leaving a 2-in. bottom margin. Label all figure parts with (a), (b), etc. Make sure each figure is straight on the page. *Photocopies of artwork are not acceptable.*
- **Do not use correction fluid or tape on illustrations.** The scanner is extremely sensitive and reproduces all flaws (e.g., correction fluid, tape, smudges, dust). *Do not write on the back of the figure* because it will be picked up by the scanner.
- **Authors' laser-generated graphics are acceptable only if the lettering and lines are dark enough, and thick enough, to reproduce clearly,** especially if reduction is required. Maximum black–white contrast is necessary. Choose a laser printer with the highest dot-per-inch (dpi) available (i.e., the highest resolution possible). Remember that fine lines in laser-generated graphics tend to disappear upon reduction, even if the oversized original looks acceptable.
- **Submit continuous-tone photographs in final published size on white glossy or matte paper.** Avoid glossy paper stock that is off-white, ivory, or colored because contrast within the illustration will be lost in reproduction. Print the photograph with more contrast than is desired in the final printed journal page. Avoid dull, textured paper stock, which will cause illustrations to lose contrast and detail when reproduced.

Preparation of Electronic Graphics Files

- We recommend that all halftone art (screened art), shaded figures, and combinations (line art + halftone) be submitted electronically. Computer-generated illustrations output to desktop printers produce a screen. These figures are most problematic in the scanning process, because scanning screened output produces an unacceptable moiré pattern.
- **Acceptable formats:** Graphics must be submitted as PostScript, EPS (using either Arial or Times Roman fonts), or TIFF (lzw compressed). Do not send application files, e.g., Corel Draw, etc.
- **Settings:** Set the graphic for **600 dpi** resolution for line art, **264 dpi** for halftones (noncompressed), and **600 dpi** for combinations (line art + halftone). Save the files to grayscale (B/W), not color.
- Make sure there is only **ONE** figure per file. Each figure file should include all parts of the figure. For example, if Figure 1 contains three parts (a, b, c), then all of the parts should be combined in a single file for Figure 1.
- You are still required to send hardcopies of all figures to the Editorial Office, along with a hardcopy of the manuscript.
- Do not FTP the graphics files to the Editorial Office or AIP unless otherwise instructed to do so. Full instructions will be sent to you twice: Once on a hardcopy form after initial submission of your article, and once via e-mail after your article has been accepted for publication.

Detailed instructions for submitting electronic graphics to AIP and a glossary of terms may be found on the AIP Physics Information Netsite at www.aip.org/epub/submitgraph.html

9. Mathematics. Mathematical expressions should be typewritten as completely as possible, with unavailable symbols being carefully inserted in ink. Special care should be given to make equations and formulas clear to the typesetter, and all handwritten letters and symbols that cannot be typed should be identified *in the margin and only the first time that such occur.* Clearly distinguish capital and lowercase letters where there could be confusion. Use fractional exponents to *avoid root signs.* Extra symbols should be introduced to avoid complicated exponents or where it is necessary to repeat a complicated expression a number of times. Use the solidus (/) wherever possible for fractions. Do not repeat mathematical derivations that are easily found elsewhere in the literature; merely cite the references.

10. References. *References should be collected and typed (double-spaced) on a separate page.* A list of the abbreviations for the names of journals appears in the *Style Manual.* They should be numbered consecutively and arranged thus:

[1]A. B. Smith, Phys. Rev. **41**, 852 (1932).
[2]H. Lamb, *Hydrodynamics,* 6th ed. (Cambridge University Press, Cambridge, England, 1940), pp. 573, 645.

11. Tables. All but the simplest tabular material should be organized into separate tables. Tables should be numbered with Roman numerals and typed on sheets at the end of the running text. Each table must have a caption typed at the top of the table, which makes the data in the table intelligible without reference to the text. Avoid complicated column headings. If necessary, use symbols which are explained in the caption. Type a double horizontal line below the caption, a single line below the headings, and another double line at the end of the table. Long tables should, if possible, be submitted in a form ready for direct photoreproduction. Detailed instructions for preparing camera-ready tables are available from the Editorial Office.

12. Publication Charges comprise the major source of income for this journal. Consequently, author's institutions or companies are asked *and urged* to pay a publication charge—which entitles them to 100 reprints without covers—of $95 per printed page.

13. Correspondence and Alterations. *Send manuscripts to the Editorial Office, Journal of Vacuum Science and Technology, 10 Park Plaza, Suite 4A, Caller Box 13994, Research Triangle Park, North Carolina 27709-3994;* telephone: (919) 361-2787; FAX: 919-361-1378; E-mail: jvst@jvst.org. Manuscripts to be published in special issues (papers from conferences) should be sent to the Guest Editor appointed for the particular Conference. Please do not send the Editor any correspondence about proofs, reprints, or publication charges.

If your paper is accepted for publication, you will be notified by the Editor. You will also be notified when your paper is sent to the American Institute of Physics. *All* subsequent correspondence about your paper should be addressed to the Editorial Supervisor, JVSTA *or* JVSTB, American Institute of Physics, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Reference must be made to the author, journal, and scheduled date of issue.

14. Proofs of Articles will be sent to you and should *be returned promptly to the Editorial Supervisor, JVSTA or JVSTB, AIP.* A few alterations in proof are unavoidable, but the cost of making extensive alterations or of correcting mistakes caused by careless preparation of the manuscript will be charged to the author. *Unless corrected proofs are returned punctually, your paper will be published in uncorrected form and will be so labeled.*

15. Compuscripts. AIP is accepting author-prepared computer files for use in production. If you have used REVTeX, LaTeX, Microsoft Word, or WordPerfect to compose your manuscript, AIP may be able to use your file to produce author proofs. If you are interested in participating in the program, please indicate so in the cover letter that accompanies your original submission. Please also include an electronic mail address if available. Do not send a disk with your original submission. Wait for further instructions from the editorial office on when to send the disk (or, if the editorial office accepts electronic submissions, when to e-mail the file). Do not send the file directly to AIP Production unless otherwise requested. AIP uses translation software to convert REVTeX, LaTeX, MS Word, or WordPerfect files into Xyvision composition files for production. For this project to be of benefit to both the author and the production staff, if is imperative that the guidelines as documented in either the REVTeX toolbox or the MS Word/WordPerfect toolbox be followed precisely. The REVTeX toolbox is available via anonymous FTP on the Internet from ftp.aip.org in the directory /pub/revtex. The MS Word/WordPerfect toolkit is also available via anonymous FTP from ftp.aip.org. Move to the directory /ems, then follow the instructions given on the screen. If electronic retrieval is not possible, you may receive the toolbox or toolkit on disk by contacting toolkits@aip.org. The REVTeX toolbox and the Word/WP toolkit, as well as general information regarding the compuscript program, are accessible via the AIP Physics Information Netsite (www.aip.org/epub/compuscripts.info.html). The Word/WP author toolkit FAQ ("Frequently Asked Questions") may be found at www.aip.org/pubservs/toolkit.html."

16. Electronic Mail. Authors may include their internet e-mail addresses along with all other footnotes in the following format: Electronic mail: smith@foo.bar.com

Over 7,000 Reasons to *Join* the American Vacuum Society

- *Receipt of 7,000+ pages of technical material*
- *Membership rates on conferences and symposia*
- *Discounts on short courses and educational materials*

All for under a penny a reason! Our programs allow you to:

Interact with Technology Leaders at Symposia and Conferences

Each fall, scientists gather from around the world to attend the American Vacuum Society (AVS) National Symposium —a week-long forum for science and technology exchange in the areas of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum. The symposium features papers from eight technical divisions and two technology groups, short courses, topical conferences on emerging technologies, and an equipment exhibition.

Local chapters host symposia on topics of interest to their local technical communities and equipment exhibitions during the year. Joining the AVS and a local chapter or division is an excellent way to interact with technology leaders and stay informed.

Increase Your Job Skills—More Than 70 Short Courses by Industry Experts

AVS short courses have provided hard-to-find job training for more than 20 years in applied vacuum technology, surface analysis and materials characterization, and materials, thin films, and coatings: processing and properties. The AVS offers over 70 short courses that have been developed by industry experts, and regional programs are held all year long to meet the training needs of local technical communities across the U.S.

Stay Informed—Receive Over 7,000 Pages of AVS Publications Each Year

AVS members receive the *Journal of Vacuum Science and Technology*—the first major physics journal available on CD-ROM. Refereed papers cover surface science, interfaces, nanometer-scale science, STM, plasmas, thin films, and vacuum technology. The *AVS Newsletter* announces meetings and events and includes interesting articles and conference reviews.

Surface Science Spectra is a unique laboratory reference tool published as a journal and on disk. It's an evolving database (including related sample and instrument information) of peer-reviewed XPS, Auger, and SIMS spectra of a wide range of materials.

Maintain Your Technical Edge with a Host of Educational Resources

The AVS has an ongoing program to develop and provide convenient, cost-effective educational and training resources. Formats include monographs (single-subject information); a series of book reprints entitled *AVS Classics;* various instructional materials; recommended practices information; and video tapes that provide industrial training, offer tutorial lectures, or present topical conferences.

Join Us Today—Under a Penny a Reason!

Membership dues are $75 annually ($20, students) and include receipt of *JVST, Physics Today,* and the *AVS Newsletter;* discounts on events and educational materials; and free membership in any of the AVS' chapters or divisions. (There is a separate subscription fee for *Surface Science Spectra.)*

The AVS is a volunteer-based, nonprofit technical society dedicated to advancing the science and technology of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum.

For more information, contact the AVS, 120 Wall St., 32nd Floor, New York, NY 10005, 212-248-0200, fax 212-248-0245, e-mail avsnyc@vacuum.org, http://www.vacuum.org.

...N'T ...SE YOUR TECHNICAL EDGE ...

Attend the premier annual symposium on materials, surfaces, interfaces, and processes!

American Vacuum Society 45th International Symposium November 2–6, 1998 Baltimore, Maryland

Eight technical division programs will be held on issues related to films, microelectronics, nanostructures, processing, surfaces, and vacuum, including:

- Applied Surface Science.
- Electronic Materials and Processing.
- Nanometer-Scale Science and Technology.
- Plasma Science and Technology.
- Surface Science.
- Thin Films.
- Vacuum Metallurgy.
- Vacuum Technology.

In addition, the week-long annual symposium will feature:

Four Topical Conferences

- Science of Micro-electromechanical Systems.
- Organic Electronic Materials.
- Partial Pressure Measurements for Process Control.
- Selected Energy Epitaxy.

Three Technical Group Programs

- Biomaterial Interfaces.
- Magnetic Interfaces and Nanostructures.
- Manufacturing Science and Technology.

40+ Short Courses

- Applied Vacuum Technology.
- Surface Analysis and Materials Characterization.
- Materials, Thin Films, and Coatings: Processing and Properties.

150+ Exhibitors

- Vacuum and Deposition Equipment.
- Analysis Systems.
- Vacuum-Related Services.

For more information on the symposium, visit the AVS Web site at http://www.vacuum.org or contact Angela Mulligan, (212) 248-0200, fax (212) 248-0245, e-mail avsnyc@vacuum.org.

The American Vacuum Society is a volunteer-based, nonprofit organization dedicated to advancing the science and technology of films, microelectronics, nanostructures, processing, surfaces, and vacuum and to providing a variety of educational opportunities. It has eight technology-based divisions, three technical groups, 21 local-area chapters, and about 6,000 members.

AVS BALTIMORE

Advertisers Direct Response

Journal of Vacuum Science & Technology B

Microelectronics and Nanometer Structures, Processing, Measurement, and Phenomena

Listed below are the companies that have advertised in this issue. If you would like to know more about a product or service please complete the information below and fax this form to the appropriate company.

◊ JULY/AUGUST 1998 ◊

ADVERTISERS:

A3 Ametek-Process & Analytical Instruments Div
Fax: (412) 826-0399

A9 Cerac Inc.
Fax: (414) 289-9805

A4 CHA Industries
Fax: (510) 683-3848

A15 Hiden Analytical Inc.
Fax: 44 1925 416 518

A17 International Radiation Detectors Inc.
Fax: (310) 534-3665

Cov 3 Kurt J. Lesker Co.
Fax: (412) 233-4275

Cov 2 Leybold Inficon Inc
Fax: (315) 437-3803

A11 LK Technologies
Fax : (812) 332-4493

A11 Omicron Associates
Fax: (412) 831-9828

A15 Omley Industries
Fax: (1-800) 717-3355

Cov 4 Quad Group, Inc.
Fax: (509) 458-4555

A7 Stanford Research
Fax: (408) 744-9049

A1 Varian Vacuum Division
Fax: (617) 860-5437

A13 VLSI Standards Inc.
Fax: (408) 428-9555

FROM:

_____
Your Name

_____
Title

_____
Institution/Dept.

_____
Address

_____
City/State/Zip

_____
Phone/Fax

1. **Specialty:**
   a. ☐ Applied Surface Science
   b. ☐ Electronic Mat'l. & Processing
   c. ☐ Nanotechnology
   d. ☐ Plasma Science & Technology
   e. ☐ Surface Science
   f. ☐ Thin Film
   g. ☐ Vacuum Metallurgy
   h. ☐ Vacuum Technology
   i. ☐ Other: _____

2. **Need for information:**
   j. ☐ Immediate k. ☐ Future

3. **Purchasing Authority**
   l. ☐ Specialty m. ☐ Recommend

INDEX TO ADVERTISERS

Ametek-Process & Analytical Instruments Div A3
Cerac Inc. A9
CHA Industries A4
Hiden Analytical, Inc. A15
International Radiation Detectors A17
Kurt J. Lesker Co. ..... Cover 3
Leybold Inficon, Inc. ..... Cover 2
LK Technologies A11
Omicron A11
Omley Industries A15
Quad Group, Inc. ..... Cover 4
Stanford Research Systems A7
Varian Vacuum Division A1
VLSI Standards, Inc. A13

Advertising Sales Office

American Institute of Physics
500 Sunnyside Boulevard
Woodbury, NY 11797-2999
Telephone (516) 576-2440
Fax (516) 576-2481

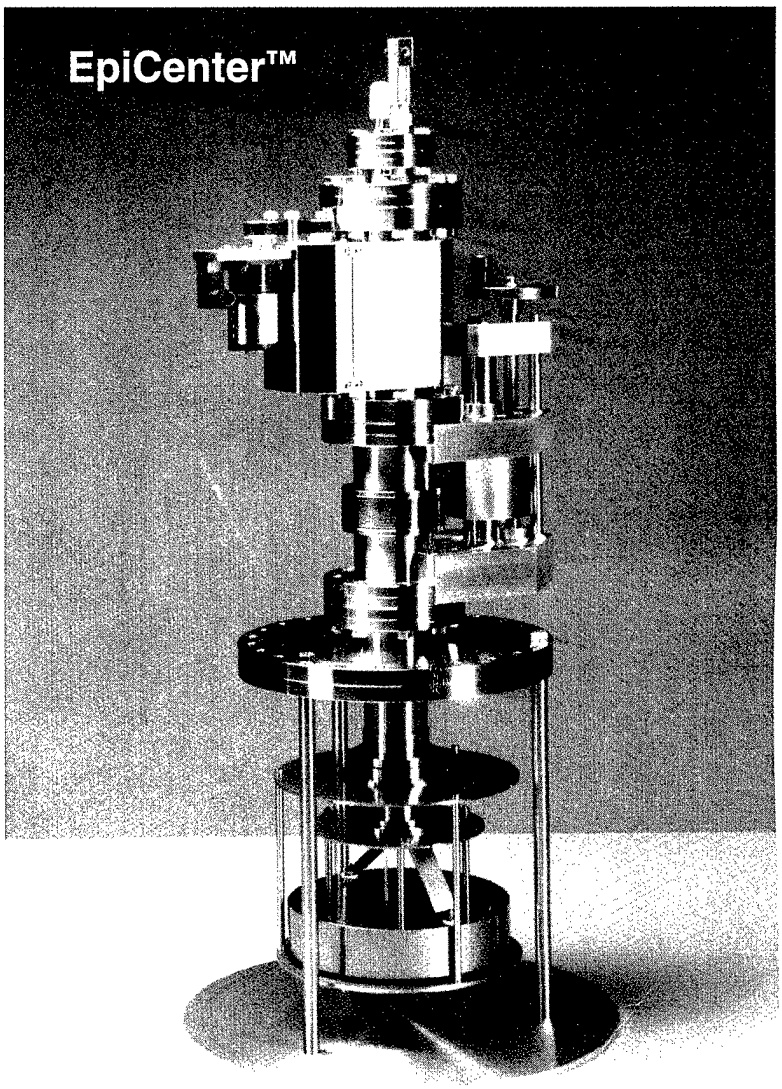
Advertising Manager: Richard T. Kobel

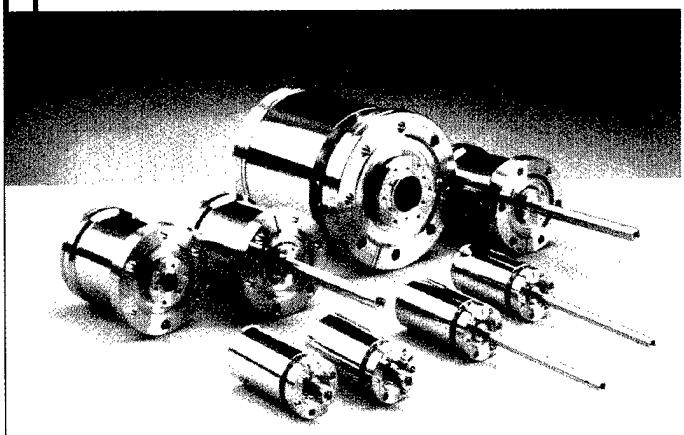
Advertising Sales Representative: Robert G. Finnegan

Production Manager: Betty Aroesty

MBE Wafer Heaters
The new standard . . .

- **Wafer sizes from 50 to 300mm**
- **Uniform heating to 1200°C**
- **Clean, high speed rotation**
- **Low particulate, UHV operation**
- **Manual or motorized wafer exchange**
- **In-line and right angle models**





The innovative **EpiCenter™** substrate heating and rotation stage uses the latest magnet technology for high torque, bellows-free rotation...the same technology used in **MagiDrive™** rotary drives (see left).

Visit us on the Web!
www.lesker.com

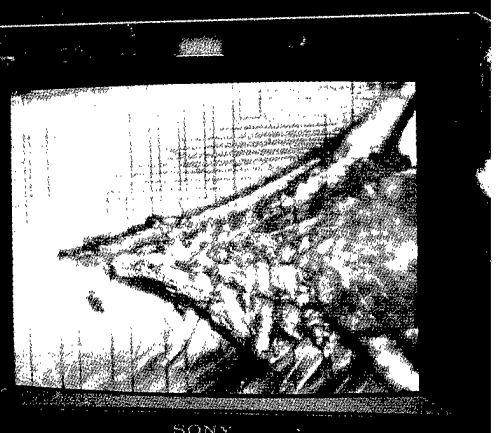
Kurt J. Lesker Company
1515 Worthington Avenue, Clairton, PA 15025
Sales: 800-245-1656 or 412-233-4200
Fax: 412-233-4275 sales@lesker.com
Int'l. Sales: international@lesker.com

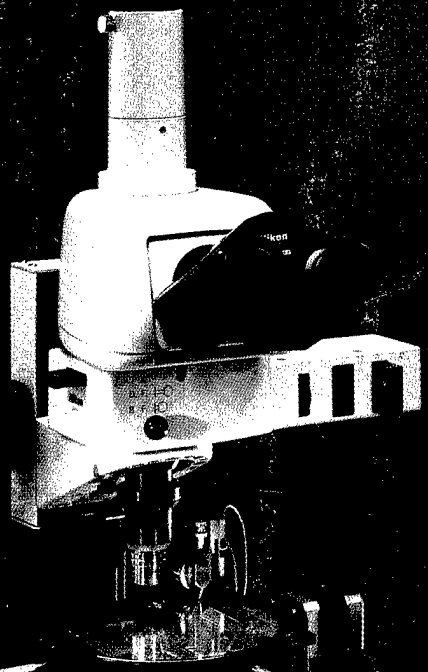
The Most Versatile Diamond Scratch Tester on Earth!

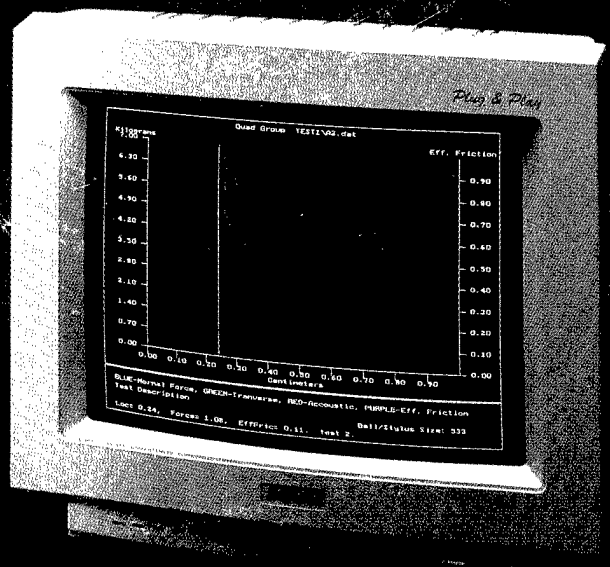
Tribo + Thin Film Adherence and Friction

Wear and Durability (Optional)

Price only $34k (not including viewing options)

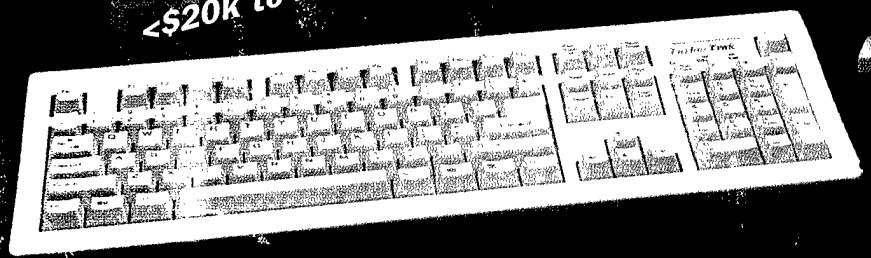






Small Footprint

Price <$20k to Romulus III owners



Universal Mechanical Strength Tester

Quad Group

1815 South Lewis Street • Spokane, WA 99224 • (509) 458-4558 • (800) 342-2430 • Fax: (509) 458-4555
E-Mail: quadgroup@spokane.net • www: spokane.net/quadgroup