DTIC
SELECTED
APR 2 7 1995

A

Physical Metallurgy and
Materials Science

METALLURGICAL AND MATERIALS TRANSACTIONS

Volume 25A–No. 9 September 1994 Pages 1781-2068

N00014-93-1-0026

19950426 048

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

ASM INTERNATIONAL
The Materials Information Society

TMS
Minerals • Metals • Materials

INSTRUCTIONS FOR AUTHORS

METALLURGICAL AND MATERIALS TRANSACTIONS publishes contributions on all aspects of research and significant engineering advances in materials science and metallurgy. Approved manuscripts will be published according to subject matter in Section A or B. Assignment will be made in the Editor's Office, but author guidance will be appreciated concerning appearance in Section A or B and the classification of the manuscript according to current nomenclature used in the Table of Contents.

Metallurgical and Materials Transactions A publishes contributions on all aspects of physical metallurgy and materials science, with a special emphasis on relationships among the processing, structure, and properties of materials.

Metallurgical and Materials Transactions B publishes contributions on the theoretical and engineering aspects of the processing of metals and other materials, including studies of electro- and physical chemistry, mass transport, modeling, and related computer applications.

All manuscripts, including symposia, will be judged by qualified reviewers according to established criteria for technical merit. The review procedure begins in the editorial office as a Key Reader is assigned by the Editor or Associate Editor. The Key Reader chooses one or two reviewers for the manuscript and submits his recommendation, based on his own and the reviewers' judgments, to the editorial office. The Editor or Associate Editor then makes a final decision on the paper.

Submission of a manuscript is representation that it has neither been copyrighted, published, nor submitted for publication elsewhere. Prior publication is a basis for rejection. Appearance in a conference proceedings or similar special presentation with limited distribution is not necessarily prior publication. In such cases the article should be so referenced. There are two classes of papers:

Technical Publication: Manuscripts should represent completed original work embodying the results of extensive field, plant, laboratory, or theoretical investigation, or new interpretations of existing problems. Material must be considered to have significant permanent value. In addition to technical acceptability, material should be presented clearly and concisely.

Communication: This class provides more rapid publication of short items. Abstracts and divisional headings are not used. The length for a Communication should be about 1000 words of text, accompanied by whatever figures and tables may be required to support this text. These fall in the following categories:

(a) theoretical or experimental work of immediate current interest;
(b) discussion of Technical Papers and Communications. These should contribute to the original article by providing confirmation or additional interpretation. They will be referred to authors for reply. Discussion and authors' reply will be published concurrently;
(c) announcements and summaries of computations and computer programs or other compilations which are available on request from the author or a specified agency.

MANUSCRIPT

1. Send to: *Metallurgical and Materials Transactions*
   Carnegie Mellon University
   Schenley Park
   Pittsburgh, PA 15213
2. **Submit three copies, each complete with abstract, tables, and figures.** All copy (abstract, text, footnotes, references, figure captions) should be typed double-spaced on one side of 22 × 28 cm (8½ × 11 in.) paper, with a margin of at least 2.5 cm (1 in.) all around. **Three sets of high quality micrographs** are required for review purposes; Xerox-type copies are not acceptable. *Printer's copy of all figures (drawings and halftones) should be retained by the author until requested by the editor (see items 3a and 3b).*
   a. Copyright transfer must accompany the manuscript. See section entitled **COPYRIGHT POLICY** which follows.
   b. A separate list of figure captions should be included (double-spaced) in addition to the caption appearing with each figure.
   c. All tabulated data identified as tables should be given a table number and a descriptive caption. Tables should be numbered consecutively, using Roman numerals.
3. **Figures (halftones and line drawings) should be designed for final printing in single column 8.5 cm (3¼ in.) width.** Double column 18 cm (7 in.) treatment will be used only when required by the complexity of the material.
   a. **Line drawings.** These should be no larger than 22 × 28 cm (8½ × 11 in.). Lettering should be large enough to be 2 mm (1/16 in.) after reduction. Glossy photocopies of larger drawings are preferred.
   b. **Photographic material.** One mounted set of glossy prints with captions should be furnished for the printer's use. These should be protected by cardboard to avoid creases and markings. Staples, tape or clips should not be used. If necessary, indicate suitable framing, position, and proportion on a working copy.
   c. **Color printing** of photographic material can be arranged. This will involve an additional charge. Authors should correspond with the editor regarding specific papers.
   d. Original drawings and photographic material will be returned to the author by the printer after publication.
4. **The abstract is the author's summary of a scientific paper and is included in the review procedure.** It should indicate newly observed facts, conclusions, and the essential parts of any new theory, treatment, apparatus, technique, *etc.* It should be concise and informative and only in exceptional cases exceed 200 words.
5. **References** should be double-spaced and listed on a separate sheet. The required format is shown in the section entitled **STYLING OF REFERENCES.**
6. If the paper depends on unpublished work, three copies of the unpublished material should be included to assist the referees in their evaluation.
7. On the printer's copy, symbols should be clearly identified, *i.e.*, "Greek omicron," "One and ell" and "oh and zero" should be differentiated by a marginal note to avoid ambiguity.
8. Extensive revision at the author's option of text or figures in proof is costly and will be charged to the authors. Typographical errors may be corrected and references updated without charge.
9. There are no facilities for translating or making editorial revisions of foreign contributions. All publication is in English, and papers must be submitted in proper form. Canadian and British spelling should be changed to Americanized version.
10. The policy of *Metallurgical and Materials Transactions* is to use the International System of Units (SI). For guidelines, see National Bureau of Standards Special Publication 003-003-02380-9 (for sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402). If other units are required for special situations, a conversion to SI must be provided in parentheses or in a table.
11. Avoid the use of trade names and proprietary information whenever possible. Such use can occasionally be justified if this is the best way to specify a particular material or process.
12. A series of papers dealing with separate aspects of a subject should be cross referenced. Divisions, such as Part I, Part II, *etc.*, are not recommended as they complicate the review procedure and abstract listings. If such divisions are considered essential, they should be justified by the author at the time of submission and are subject to approval in review.
13. A **Table of Symbols** should be included when symbols are used extensively throughout a paper.

(continued on cover 3)

METALLURGICAL AND MATERIALS TRANSACTIONS A

Physical Metallurgy and Materials Science

VOLUME 25A, NUMBER 9 **SEPTEMBER 1994**

Metallurgical and Materials Transactions A publishes contributions on all aspects of physical metallurgy and materials science, with a special emphasis on relationships among the processing, structure, and properties of materials. All papers, including symposia contributions, are reviewed by a committee consisting of one member of the Board of Review and one or two other reviewers. See inside covers for detailed instructions for authors.

EDITORIAL STAFF
Editor: David E. Laughlin
Editor Emeritus: Gerhard Derge†
Associate Editors: Alan W. Cramb
R.J. Fruehan
S. Mahajan
Thaddeus B. Massalski
Production Editor: Dora Moscatello
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
FAX: (412) 268-7169
E-mail: mettrans@andrew.cmu.edu

JOINT COMMISSION FOR METALLURGICAL AND MATERIALS TRANSACTIONS
C. Parker, *Chair*
B. Attwood
D. Granger
M. Ozgu
A. Romig
J. Salsgiver

Alexander R. Scott
Edward L. Langer

JOINT MANAGERS
Edward L. Langer
ASM INTERNATIONAL

Alexander R. Scott
The Minerals, Metals & Materials Society

SOCIETY PRESIDENTS
ASM INTERNATIONAL
John G. Simon

The Minerals, Metals & Materials Society
Lionel C. Kimerling

Iron & Steel Society
Harry E. Follwell

BUSINESS STAFF
TMS Subscriptions and Advertising:
Linda L. Gibb, *TMS*

ASM Subscriptions:
Faye D. Balser, *ASM INTERNATIONAL*

Published jointly by

The Minerals, Metals & Materials Society
420 Commonwealth Drive
Warrendale, PA 15086

ASM INTERNATIONAL
Materials Park, OH 44073

Metallurgical and Materials Transactions "A" is published monthly by THE MINERALS, METALS & MATERIALS SOCIETY (TMS), 420 Commonwealth Drive, Warrendale, PA 15086, telephone (412) 776-9000, and ASM INTERNATIONAL, Materials Park, OH 44073, telephone (216) 338-5151. TMS and ASM are not responsible for any statements made or opinions expressed in their publications.

Subscription rates per year—Metallurgical and Materials Transactions "A": Individual non-members and multi-user facilities (libraries, corporate information centers, documentation services, military installations, government and private laboratories, etc.) $696.00 U.S., $696.00 other countries; individual members of TMS or ASM $50.00 U.S., $50.00 other countries. Metallurgical and Materials Transactions "B": Individual non-members and multi-user facilities $498.00 U.S., $498.00 other countries; individual members of TMS or ASM $36.00 U.S., $36.00 other countries. Combined Metallurgical and Materials Transactions "A" and "B": Individual non-members and multi-user facilities $1134.00 U.S., $1134.00 other countries; individual members of TMS or ASM $82.00 U.S., $82.00 other countries. Back issues of Metallurgical and Materials Transactions "A" and "B" are available by contacting TMS or ASM. Bound volumes of Metallurgical and Materials Transactions are available through TMS. Claims to either ASM or TMS for issues not received must be made within 120 days of expected delivery date.

Copyright 1994 by THE MINERALS, METALS & MATERIALS SOCIETY and ASM INTERNATIONAL . . . indexed by major abstracting services . . . second class postage paid at Warrendale, PA 15086, and at additional entry offices. Send all 3579 forms to Metallurgical and Materials Transactions, TMS, 420 Commonwealth Drive, Warrendale, PA 15086.

Authorization to photocopy items for internal and/or personal use of specific clients is granted upon request to TMS or ASM. Libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, may photocopy items provided that the base fee of $00.75 per copy is paid directly to CCC, 21 Congress St., Salem, MA 01970.

Metallurgical and Materials Transactions "A" prior to 1972 is available from University Microfilms International, 300 North Zeeb Rd., Ann Arbor, MI 48106.

ISSN: 1073-5623/83

CODEN: MTTABN 25A(9) 1781-2064 (1994)

BOARD OF REVIEW

H.I. Aaronson
J. Ainsworth
S. Ankem
S.D. Antolovich
S.A. Argyropoulos
M.J. Aziz
A.S. Ballantyne
R.G. Bautista
C. Beckermann
W.D. Bennon
J.T. Berry
R. Berryman
C.P. Blankenship, Jr.
C.L. Briant
C.R. Brinkman
A. Bronson
W.E. Brower, Jr.
V. Brusic
A. Bustos
M. Byrne
R.W. Carpenter
K.S. Chan
K.-M. Chang
P. Chaubal
P.K. Chaudhury
H. Chen
N.S. Cheruvu
H. Cialone
M.J. Cieslak
J.A. Clum
T.H. Courtney
J.A. Dantzig
D.L. Davidson
G.P. Demopoulos
P. Desclaux
M. Dollar
F.M. Doyle
S. Dregia
D. Dreisinger
T.W. Eagar
G.R. Edwards
N. El-Kaddah
J.W. Elmer
W.H. Emling
M. Enomoto
B. Farouk
S.G. Fishman
R.A. Fournelle
D. Frear
P.D. Funkenbusch
A. Garg
W.M. Garrison
D.R. Gaskell
W.W. Gerberich
P.C. Glaws
J. Goldak
T. Gross
J. Grubb
D. Gupta
J.E. Hack
K.N. Han
R.L. Harris
M.J. Haun
J.B. Hiskey
J.M. Howe
Hsun Hu
S.-C. Huang
G. Irons
A.F. Jankowski
K.V. Jata
W.C. Johnson
R.H. Jones
J.S. Kallend
S.L. Kampe
M.E. Kassner
T.Z. Kattamis
A.H. King
C.C. Koch
B. Lally
C. Landefeld
J.D. Landes
T.G. Langdon
J. Lankford
E.J. Lavernia
H.M. Ledbetter
C.G. Levi
J.J. Lewandowski
P.K. Liaw
B. Lichter
K. Liddell
J.C. Liu
H. Margolin
S. Marsh
M. Mataya
R. Matway
J. Mazumder
A. McLean
T.R. McNelley
S.K. Menon
P.J. Meschter
M. Meshii
M.A. Meyers
G.M. Michal
B. Mishra
F. Mohamed
N.R. Moody
M.B. Mooiman
J.J. Moore
J.E. Morral
A.E. Morris
D.R. Morris
W.C. Moshier
K. Mukherjee
K.L. Murty
R.H. Nafziger
S. Nourbakhsh
D.L. Olson
M. Ozgu
B. Ozturk
R.A. Page
U. Pal
C. Pande
J.M. Papazian
H.G. Paris
C.A. Parker
R.D. Pehlke
T. Piwonka
M.R. Plichta
J.H. Poveromo
M. Przystupa
V.A. Ravi
D. Rego
W.T. Reynolds, Jr.
C.G. Rhodes
G. Richards
R.E. Ricker
J.M. Robertson
P. Rohatgi
A.K. Sachdev
I.V. Samarasekera
M. Sarikaya
A. Saxena
H. Schade
J.H. Schneibel
S.K. Sharma
G.J. Shiflet
G.K. Sigworth
J.E. Smugeresky
G. Spanos
S.K. Srivastava
G.R. Stafford
K.P. Staudhammer
D.M. Stefanescu
R. Stevenson
C. Suryanarayana
J.M. Tartaglia
B.G. Thomas
J.A. Todd
J.M. Toguri
A.K. Vasudevan
J.M. Vitek
V. Voller
P.W. Voorhees
G.W. Warren
G.S. Was
C.M. Wayman
C.A. Wert
J.A. Wert
S.H. Whang
J.E. Wittig
M. Wuttig

DTIC QUALITY INSPECTED 5

Contents

Pacific Rim Symposium—Part I

C.M. Wayman	1787	The Phenomenological Theory of Martensite Crystallography: Interrelationships
H.I. Aaronson and M.G. Hall	1797	A History of the Controversy over the Roles of Shear and Diffusion in Plate Formation above M_d and a Comparison of the Atomic Mechanisms of These Processes
J.W. Christian	1821	Crystallographic Theories, Interface Structures, and Transformation Mechanisms
B.C. Muddle, J.F. Nie, and G.R. Hugo	1841	Application of the Theory of Martensite Crystallography to Displacive Phase Transformations in Substitutional Nonferrous Alloys
U. Dahmen	1857	A Comparison between Three Simple Crystallographic Principles of Precipitate Morphology
G.C. Weatherly and W.-Z. Zhang	1865	The Invariant Line and Precipitate Morphology in Fcc-Bcc Systems
Gary R. Purdy and Wenzheng Zhang	1875	On the Role of Interphase-Boundary Structure in Plate Growth by Diffusional Mechanisms
J.P. Hirth	1885	Ledges and Dislocations in Phase Transformations
G.J. Shiflet and J.H. van der Merwe	1895	The Role of Structural Ledges as Misfit-Compensating Defects: fcc-bcc Interphase Boundaries
Yiwen Mou	1905	Intrinsic Ledges at Interphase Boundaries and the Crystallography of Precipitate Plates
James M. Howe	1917	Atomic Site Correspondence and Surface Relief in the Formation of Plate-Shaped Transformation Products
M.G. Hall and H.I. Aaronson	1923	Formation of Invariant Plane-Strain and Tent-Shaped Surface Reliefs by the Diffusional Ledge Mechanism
Kenneth C. Russell	1933	Effects of Precursor Matrix Events on Subsequent Nucleation
M.K. Kang, Y.Q. Yang, Q.M. Wei, Q.M. Yang, and X.K. Meng	1941	On the Prebainitic Phenomenon in Some Alloys
Masato Enomoto	1947	Thermodynamics and Kinetics of the Formation of Widmanstätten Ferrite Plates in Ferrous Alloys
Mats Hillert	1957	Diffusion in Growth of Bainite
G. Spanos	1967	The Fine Structure and Formation Mechanism of Lower Bainite
Yasuya Ohmori, Hiroyuki Ohtsubo, Yun Chul Jung, Shuji Okaguchi, and Hiroo Ohtani	1981	Morphology of Bainite and Widmanstätten Ferrite
S.K. Liu, L. Yang, D.G. Zhu, and J. Zhang	1991	The Influence of the Alloying Elements upon the Transformation Kinetics and Morphologies of Ferrite Plates in Alloy Steels
Hong-Sheng Fang, Jia-Jun Wang, and Yan-Kang Zheng	2001	Formation Mechanism of Bainitic Ferrite and Carbide
Kaneaki Tsuzaki, Aki Kodai, and Tadashi Maki	2009	Formation Mechanism of Bainitic Ferrite in an Fe-2 Pct Si-0.6 Pct C Alloy

Mechanical Behavior

In-Gyu Lee, Amit K. Ghosh, Ranjan Ray, and Sunil Jha	2017	High-Temperature Deformation of *B*2 NiAl-Base Alloys
D.-G.C. Syu and A.K. Ghosh	2027	Forging Limits for an Aluminum Matrix Composite: Part I. Experimental Results
D.-G.C. Syu and A.K. Ghosh	2039	Forging Limits for an Aluminum Matrix Composite: Part II. Analysis
D.-G.C. Syu and A.K. Ghosh	2049	Stress-State Dependence of Strain-Hardening Behavior in 2014 Al/15 Vol Pct Al_2O_3 Composite

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification .....
By .....
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

Part I

Proceedings of the Pacific Rim Conference on The Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products

Foreword

The question as to whether the atomic movements involved in the formation of plates during phase transformations in crystalline solids taking place at temperatures above the martensite range occur by closely coordinated glide-type motions, *i.e.*, shear, or by biased random walk diffusion has been a controversial issue since it was raised by J.M. Robertson (1929), E.S. Davenport, and E.C. Bain (1930) and R.F. Mehl and C.S. Barrett (1931). Contrary to reasonable expectation, the increasing availability of sharper experimental tools (particularly visual and analytical transmission electron microscopy) and improved theoretical analyses has made this controversy more intense, albeit also more precisely defined. Since it has become clear that at least part of the difficulty being encountered in resolving this basic issue is one of communication between the two schools of thought, the Phase Transformations Committee of ASM International (Chaired by Professor Hayden Chen during the organization of this conference and then by Professor James M. Howe) decided to hold a conference focused as tightly as practicable solely upon this question. All speakers at this conference were invited, as were those additional conferees who have maintained an active interest in the subject. Approximately equal numbers of papers were commissioned from the two schools, as well as several from leading authorities in related phase transformation research areas who were not identified with either school. In addition to time for discussion scheduled after the presentation of each paper, most of the last day of the conference was devoted to two General Discussions sessions in which the main issues which arose during the Conference were publically discussed at length. Recordings were made of these Discussions. After some editing by the organizers, the individual discussions were returned to those who had given them for modification into final form. The compilation of these discussions, together with the large set of papers obtained on the formation of α_1 plates in Cu-Zn, Cu-Zn-Al, and Ag-Zn alloys, will appear soon in Part II of these proceedings, also to be published in this journal.

Because interest in this controversy has increasingly shifted in recent years to Japan and China, it was decided the conference would be held in the Hawaiian Islands—in particular on the "big island" of Hawaii—because these islands have won wide acceptance and the vote of preference as a meeting place for persons from the Pacific Rim countries. The Kona Hilton Hotel provided an outstanding milieu for the conference in the form of a relaxed and informal but very comfortable environment in which private discussions flourished as vigorously as did those held in public. (Indeed, considerable correspondence and research has already resulted from this conference.)

The organizers would like to express their gratitude to the Office of Naval Research (Dr. Robert C. Pohanka, Dr. George Yoder, and Dr. A.K. Vasudevan) and the National Science Foundation (Dr. Bruce A. MacDonald) for financial support which made possible the participation of a number of the conferees. Ms. Althena Franklin of the Kona Hilton Hotel staff and Mrs. Melody Guthrie, whose services as conference manager were provided by the Naval Research Laboratory, handled the operation of the conference so smoothly and efficiently that the organizers were able to devote their full attention to the technical business of the conference itself. We are grateful to the members of the Phase Transformations Committee of ASM International for their support and constructive

suggestions during the formulation of the conference program. Thanks are also due to Professors Tsun Ko (University of Science and Technology Beijing), Makato Kikuchi (Tokyo Institute of Technology), Gregory Olson (Northwestern University), Ken'ichi Shimizu (Kanazawa University), George Weatherly (McMaster University), Hayden Chen (University of Illinois at Urbana–Champaign), Barry Muddle (Monash University), Yasuya Ohmori (Ehime University), George Krauss (Colorado School of Mines), Mats Hillert (Royal Institute of Technology), T.Y. Hsu (Xu Zuyao) (Shanghai Jiao Tong University), James Howe (University of Virginia), and William Reynolds, Jr. (Virginia Polytechnic Institute and State University) for their service as session chairmen. Finally we thank Professor David E. Laughlin, Editor, and Ms. Dora Moscatello, Production Editor, Metallurgical and Materials Transactions, for supervising the review procedures for the conference papers.

C.M. Wayman, General Chairman
University of Illinois at Urbana–Champaign

H.I. Aaronson, Administrative Chairman
Geo-Centers, Inc., stationed at the Naval Research Laboratory

J.P. Hirth, Advisory and General Discussions Chairman
Washington State University

B.B. Rath, Government Relations Chairman
Naval Research Laboratory

Symposium Organizers

The Phenomenological Theory of Martensite Crystallography: Interrelationships

C.M. WAYMAN

The phenomenological theory of martensite crystallography is based on the observation that the habit plane is an invariant plane in which all lines in it are neither distorted nor rotated by the displacement vector of the invariant plane strain. The structural change is effected by a Bain deformation which follows intuitively from the orientation relationship and through which atomic interchanges do not occur. A lattice invariant deformation is also incorporated to insure that the martensite-parent interface is semicoherent and glissile. These factors and the orientation relationship as well are uniquely interrelated within the theoretical framework and must be so for any transformation alleged to follow the theory.

I. INTRODUCTION

IN one of the most classical metallurgical articles of the century,[1] Bain published "The Nature of Martensite" in 1924. Some 70 years later, this article still provides fascinating reading about its crystallography, morphology, kinetics, and mechanical properties. Indeed, on the subject at hand here, it may be said that Bain laid some crucial cornerstones of the presently accepted phenomenological theory of martensite crystallography (PTMC). This theory became fully developed, more or less, in the 1950s, following much intervening work of increasing degrees of sophistication and definitiveness. Today, "the Bain strain" is common knowledge, even outside metallurgical circles. It is the purpose of the present article to review the development and assess the validity of the PTMC. As other articles from this conference will suggest, the theory may transcend its original intent, and an important aspect of the theory, the invariant line strain, is coming into closer scrutiny.

II. THE LATTICE CORRESPONDENCE

All martensitic transformations involve a correspondence by means of which lattice points in the parent phase are uniquely related on a one-by-one basis to those in the product (martensite). The original proposal by Bain[1] is drawn to scale and copied from the original paper. Bain's correspondence relates the body-centered tetragonal (bct) cell (heavy lines) within two unit face-centered cubic (fcc) cells to the body-centered cubic (bcc) unit cell of α-iron. A simple upsetting of the former produces the bcc lattice (it was then thought that martensite was cubic α-iron). This contraction along the c-axis and expansion along the a axis is almost obvious from inspection, but there are many possible fcc-bcc correspondences, another case in point given in Figure 2. Jaswon and Wheeler[2] considered many such correspondences between the two lattices and showed mathematically that of all considered, Bain's correspondence involved the smallest principal strains. For this, and other reasons as well, the Bain strain has gained universal acceptance as a model for the fcc-bcc transformations; and even in other alloy systems involving different crystal structures and correspondences, the pure strain effecting the structural change is most frequently referred to as the Bain strain.

Bain argued in favor of his distortion by further noting that "there are also extensions and compressions locally developed as shown by the roughening of a polished surface of austenite (fcc parent) after martensitization is brought about by cooling to liquid-air temperatures". But he went further to say that "it is not to be supposed that the long needles (actually plates of martensite) one sees in an austenite-martensite mixture are wholly of one orientation", reasoning that one orientation would result in too much strain accumulation. Thus, Bain noted a correspondence and distortion, a shape change of a transformed region, and provided the notion that a crystal of martensite is not really a single crystal (which had undergone solely the Bain distortion). These three features in one way or another have been retained in present thinking and the crystallography. Bain said nothing quantitative about the shape change but did specify numerically the principal strains (distortions) involved in his correspondence.

III. THE ORIENTATION RELATIONSHIP BETWEEN THE MARTENSITE AND PARENT LATTICES

Referring to Figure 1, it is easy to see that certain planes and directions should remain parallel after the Bain

C.M. WAYMAN, Professor, is with the Department of Materials Science and Engineering, University of Illinois, Urbana, IL 61801.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

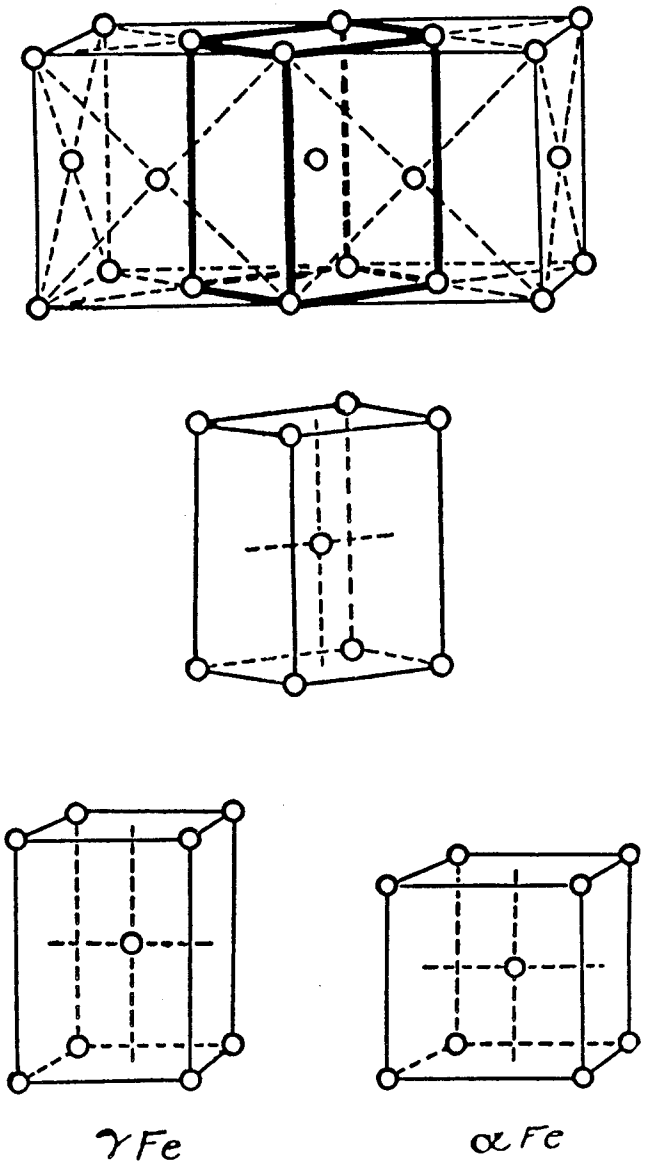


Fig. 1—Model for the fcc to bcc austenite (γ)-martensite (α) transformation in steels as proposed by Bain.[1] At the time (1924) it was thought that martensite was the same as cubic α-iron.

distortion has occurred ("martensitization"). For example, using subscripts P and M for the fcc parent and bcc martensite, respectively, we have

$$[001]_P \parallel [001]_M$$

$$(001)_P \parallel (001)_M$$

$$[110]_P \parallel [010]_M$$

$$[1\bar{1}0]_P \parallel [100]_M$$

$$(110)_P \parallel (010)_M$$

etc.

In other words, the orientation relationship, as expressed by the parallelism of important planes and directions, is rational. Not much later it was shown that this simplicity is not true. Both Kurdjumov and Sachs[3] and Nishiyama[4] using X-ray diffraction found that these important planes and directions (as well as some others) were not really parallel, in some cases detecting deviations of more than a few degrees. Thus, the corresponding cells in the martensite and the parent are also rotated with respect to each other. Important planes and directions were no longer precisely parallel and it became clear that the martensite-parent orientation relation was irrational. Kurdjumov and Sachs[3] and Nishiyama[4] did find, however, that one set of close-packed planes and directions (in those planes) were parallel, *i.e.*,

$$(111)_P \parallel (011)_M$$

$$[\bar{1}01]_P \parallel [\bar{1}\bar{1}1]_M$$

At this juncture, both Kurdjumov and Sachs[3] and Nishiyama[4] focused on the orientation relationship and attempted to account for it. They did not employ the Bain strain, but rather suggested austenitic slip shears combined with some rather *ad hoc* atomic adjustments to yield the observed result.

At the next juncture, people began to note that martensite had a definite habit plane and that the surface relief (macroscopic shape strain) attendant to the formation of a plate resembled a shear, and various "shear" mechanisms were put forth. However, Machlin and Cohen changed the general tone of thinking when they showed conclusively that the "macroscopic shear" accompanying the formation of a plate was not a simple shear,[5] and Bowles[6] described the strain as a simple tilting of the surface about its intersection with the habit plane.

IV. THE MARTENSITE SHAPE STRAIN (SURFACE RELIEF)

In more detail, Bowles and Mackenzie[7] termed the shape strain an invariant plane strain (IPS) according to which the displacement of any point is in a common direction and proportional to its distance from the invariant (habit) plane.

A schematic representation of the IPS shape change is shown in Figure 3. The displaced fiducial scratch **DF** was originally the straight line **DE**, and the habit plane **ABC** is the invariant plane. It is very important to comment that the direction of displacement in an IPS does not in general lie in the habit plane. In other words, the macroscopic displacement by an IPS has both a shear and a volume component, the latter directing the overall displacement out of the plane. In the case of pure shear (like in mechanical twinning), the displacement vector lies in the plane because there is no volume change. But martensitic transformations, in general, have both a shear and a volume component, the latter arising from the transformation volume (structural) change. There is no volume (structural) change in twinning. Figure 4 depicts three-dimensionally an IPS in which a cube is distorted into a parallelepiped. Plane **ABCD** is the invariant plane and the displacement direction is given by the vector **d** which has both normal and parallel components. Clearly, a volume change has occurred, given by the normal component of the displacement vector.

V. THE SUBSTRUCTURE IN MARTENSITE

Although not known by Bain,[1] Kurdjumov and Sachs,[3] Nishiyama,[4] and contemporary investigators,

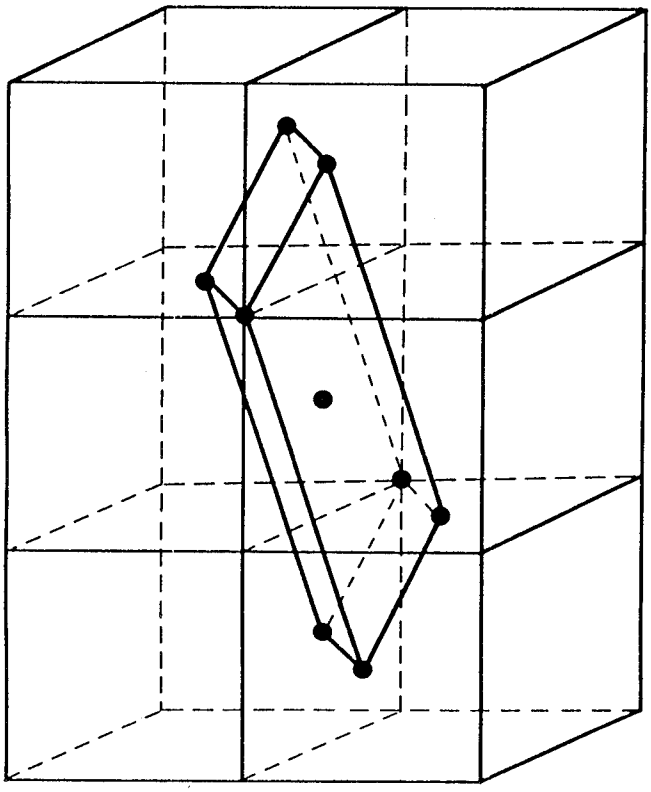


Fig. 2—A different lattice correspondence for the fcc-bcc martensitic transformation involving larger distortions than the Bain strain shown in Fig. 1.

it has been known for some time since that most martensites have a fine structure consisting of internal twins or stacking faults. Somehow this must fit into the theory, and so did the two-shear models around the time of Greninger and Troiano[8] which did not provide an accounting of *all* the known geometrical aspects of the transformation, principally because the shape change was taken to be a simple shear (not IPS) on the habit plane. However, the important point was made by Greninger and Troiano that a shear on the habit plane failed to produce the correct martensite crystal structure. As will be seen, the martensite substructure acts to counteract certain dimensional changes brought about by the Bain strain. It should be mentioned that the substructure in martensite is typically on the scale of 100 Å, and this inhomogeneity (twinning, faulting, and slip in some instances) is frequently referred to as the "inhomogeneous shear" in current theories, *i.e.*, in contrast to the "shape shear" which is actually an IPS (with both a shear and normal component).

Figure 5 is a transmission electron micrograph showing a single martensite plate in an Fe-Ni-C alloy which is finely internally twinned. These twins are referred to as transformation twins. Figure 6 shows schematically the appearance of internally twinned (a) and internally slipped martensite (b). Note that the jogs in the interface act in "push-pull" to average out the long-range interface distortion. Figure 7 shows that the same macroscopic effect is given by the effective shear angle, γ. Thus, mathematically there is no difference between Figures 7(a) and (b) and use is made of this "convenience" in theories

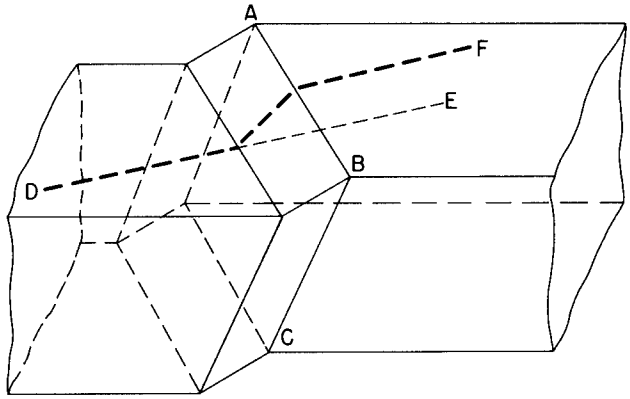


Fig. 3—Idealized representation of IPS caused by the formation of martensite plate with habit (invariant) plane **ABC**. The straight scratch **DE** before transformation is displaced to the position **DF** by the martensite plate.

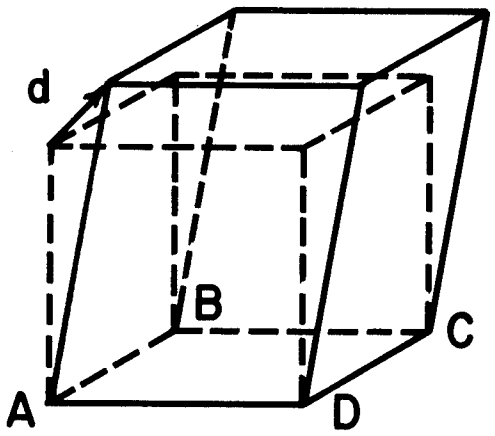


Fig. 4—Idealized representation of a general IPS. The plane **ABCD** is invariant and the direction of displacement **d** has both volume (normal) and shear components (parallel). The cube becomes a parallelepiped of larger volume.

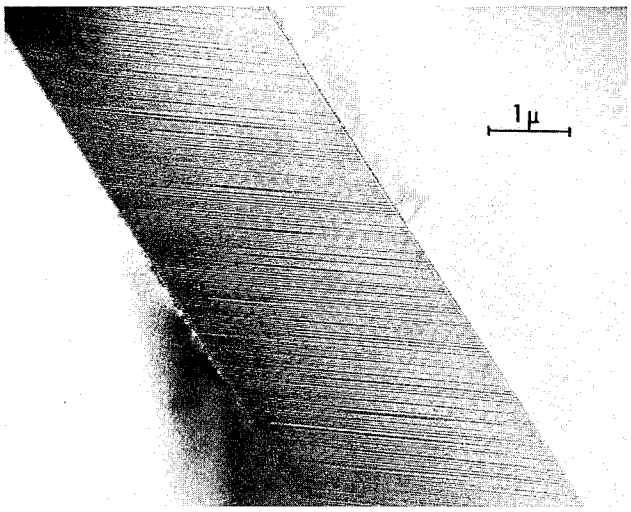


Fig. 5—Transmission electron micrograph showing internally twinned martensite plate in a matrix of retained austenite in an Fe-Ni-C alloy. The fine twins in the martensite are transformation twins.

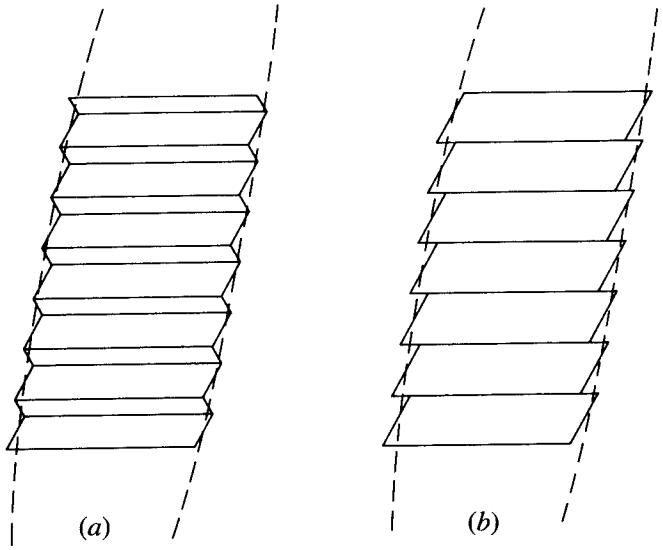


Fig. 6—Schematic appearance of internally twinned (*a*) and internally slipped (*b*) martensite plates. The serrated interfaces in both cases prevent the localized Bain strain from accumulating, thus leaving the interface (habit plane) one of zero average distortion.

which follow. As Christian has pointed out,[9] when slipped martensite results, the interior of the plates is theoretically featureless, but arrays of dislocations exist at the martensite-parent boundary. On the other hand, twinned plates contain numerous twin boundaries, but there are no interface dislocations.

Clearly, the Bain distortion leaves no plane undistorted because everything goes along with the upsetting process. However, the inhomogeneous shear acts to lengthen vectors shortened by the Bain strain and *vice versa*, and thus undistorted planes (not invariant planes) can be found. One problem still remains at this point. Even though the required structural change has been effected (by the Bain distortion) and an undistorted contact plane has been provided for (by the inhomogeneous shear), the habit plane is still not unrotated as required from observation. Thus, a rigid body rotation is additionally incorporated along with the Bain distortion and inhomogeneous shear, and these are the three phenomenological steps describing the total transformation. There is no time sequence implied as to which step occurs when. The combined effect of these three operations must, of course, be equivalent to the shape deformation. Within the framework just described, different crystallographic features such as the habit plane and orientation relationship can be predicted by supposing that the inhomogeneous shear occurs on different crystallographic plane and directions.

VI. FORMULATION OF THE THEORY

The 1950s were richly rewarding years for martensite. A number of fundamentally identical models evolved, making extensive use of matrix algebra. The terminologies were different and the inner workings must have appeared disparate, but in retrospect all paths converged.

The basic equation of the crystallographic theory is

$$\mathbf{P}_1 = \mathbf{R}\bar{\mathbf{P}}\mathbf{B} \qquad [1]$$

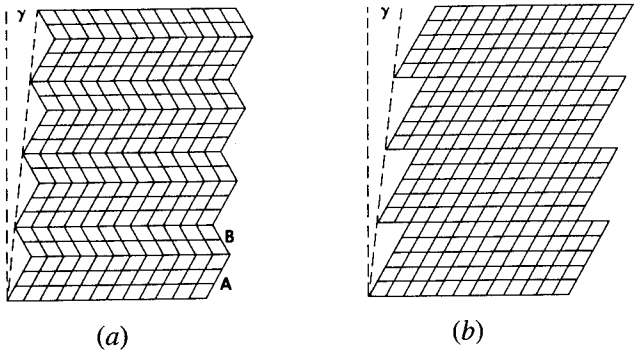


Fig. 7—Idealized diagram showing the mathematical equivalence of twinning (*a*) and slip (*b*). Regions **A** and **B** are twin related, and the overall macroscopic shear angle γ is identical in each case.

where **B** is the Bain distortion, $\bar{\mathbf{P}}$ is a simple (inhomogeneous) shear (following, mathematically, the Bain distortion), **R** is the rigid body rotation previously mentioned, and $\mathbf{P}_1$ is the IPS shape deformation. **B**, $\bar{\mathbf{P}}$, **R**, and $\mathbf{P}_1$ are all (3 × 3) matrices. The matrix product $\mathbf{R}\bar{\mathbf{P}}\mathbf{B}$ is equivalent to the shape deformation $\mathbf{P}_1$ and the rotation **R** rotates the plane left undistorted by $\bar{\mathbf{P}}\mathbf{B}$ to its original position. That is, $\mathbf{P}_1 = \mathbf{R}\bar{\mathbf{P}}\mathbf{B}$ is an IPS.

Although Eq. [1] shows the inhomogeneous shear $\bar{\mathbf{P}}$ following the Bain distortion, it is to be noted that the same mathematical result is obtained by "allowing" the shear to occur in the parent phase prior to the Bain distortion. By following this latter procedure, there are certain computational simplifications to be gained, and in this case the basic equation becomes

$$\mathbf{P}_1 = \mathbf{RBP} \qquad [2]$$

where **P**, as before, represents a simple shear.

With regard to the previous discussion, it will be noted that the IPS shape deformation can be expressed as

$$\mathbf{P}_1 = \mathbf{I} + m\mathbf{dp}'$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + m[d_1 d_2 d_3]\,(p_1 p_2 p_3)$$

$$= \begin{pmatrix} 1 + md_1p_1 & md_1p_2 & md_1p_3 \\ md_2p_1 & 1 + md_2p_2 & md_2p_3 \\ md_3p_1 & md_3p_2 & 1 + md_3p_3 \end{pmatrix} \qquad [3]$$

where $\mathbf{p}'$ (prime meaning transpose) being a plane normal is written as a (1 × 3) row matrix in contrast to **d**, a lattice vector, which is a (3 × 1) column matrix.

Going back to Eq. [2], it is noted that **P** is a simple shear and therefore of the invariant plane strain form ($\mathbf{I} + m\mathbf{dp}'$). Further, the inverse of **P**, $\mathbf{P}^{-1} = \mathbf{I} - m\mathbf{dp}'$, corresponds to a simple shear of the same magnitude on the same plane, but in the opposite direction. Thus, both **P** and $\mathbf{P}^{-1}$ are IPSs. It is then convenient to rewrite Eq. [2] as

$$\mathbf{P}_1\mathbf{P}_2 = \mathbf{RB} \qquad [4]$$

where $\mathbf{P}_2 = \mathbf{P}^{-1}$. Since $\mathbf{P}_1$ and $\mathbf{P}_2$ are IPSs, their product **RB** is an invariant *line* strain, **S**, defined by the planes (*i.e.*, their intersection) which are invariant to $\mathbf{P}_1$ and

$\mathbf{P}_2$.[7] Once the invariant line strain $\mathbf{S}$ is known, all the crystallographic features of a given martensitic transformation can be predicted. It is beyond the scope of the present account to go into the details of the invariant line strain analysis, but some highlights can be mentioned. The Bain correspondence and distortion are known from the lattice parameters of the parent and martensitic phases. $\mathbf{R}$ can be determined once the plane $\mathbf{p}_2'$ and direction $\mathbf{d}_2$ of $\mathbf{P}_2$ are assumed.

The important results of the invariant line strain analysis, noting that the shape strain is $\mathbf{P}_1 = \mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1'$ and that the simple shear (preceding the Bain distortion) is $\mathbf{P}_2 = \mathbf{I} + m_2\mathbf{d}_2\mathbf{p}_2'$, are as follows (where the magnitudes, directions, and planes of the component invariant plane strains are given respectively by m, $\mathbf{d}$, and $\mathbf{p}'$).

$$\mathbf{d}_1 = [\mathbf{S}\mathbf{y}_2 - \mathbf{y}_2]/\mathbf{p}_1'\mathbf{y}_2 \qquad [5]$$

$$\mathbf{p}_1' = (\mathbf{q}_2' - \mathbf{q}_2'\mathbf{S}^{-1})\mathbf{q}_2'\mathbf{S}^{-1}\mathbf{d}_1 \qquad [6]$$

where $\mathbf{y}_2$ is any vector lying in $\mathbf{p}_2'$ (except the invariant line $\mathbf{x}$) and $\mathbf{q}_2'$ is any normal (other than $\mathbf{n}'$, the row unit eigenvector of $\mathbf{S}^{-1}$, *i.e.*, $\mathbf{n}'\mathbf{S}^{-1} = \mathbf{n}'$) to a plane containing $\mathbf{d}_2$. The normalization factor of $\mathbf{d}_1$, in Eq. [5] is $1/m_1$ and therefore, $\mathbf{P}_1$, m_1, $\mathbf{d}_1$, and $\mathbf{p}_1'$ are all determinable. The matrix $\mathbf{R}$ is determined from the requirement that $\mathbf{x}$ and $\mathbf{n}'$, which are displaced by the Bain distortion, must be totally invariant. $\mathbf{R}$ defines the orientation relationship within any small region of the martensite plane not involving $\mathbf{P}_2$. Thus, the assumed correspondence and lattice parameters determine $\mathbf{B}$, the assumption of $\mathbf{p}_2'$ and $\mathbf{d}_2$ allows $\mathbf{R}$ to be determined, and $\mathbf{RB} = \mathbf{S}$ defines the elements of $\mathbf{P}_1$.

The previous description parallels the theoretical development given by Bowles and Mackenzie,[7] but the treatments of Wechsler *et al.*[10] and Bullough and Bilby[11] are equivalent. The Bowles-Mackenzie theory has been chosen as exemplary not only because of its comparative mathematical simplicity, but also because this theory introduced the concept of an invariant line strain, which is coming into focus in interpreting certain transformations, as other articles from this conference will show.

Some variations in the basic theory just presented include the introduction of a dilatation parameter,[7] which in effect slightly relaxes the requirements that the habit plane be undistorted, and the incorporation of two inhomogeneous shear systems[12,13] such that

$$\mathbf{P}_1 = \mathbf{RBS}_2\mathbf{S}_1 \qquad [7]$$

where $\mathbf{S}_2$ and $\mathbf{S}_1$ are the two inhomogeneous shear systems involved. Neither of these modified approaches is without criticism, and the double shear approach relative to the original single shear approach loses generality.[14]

VII. DETERMINATION OF AN INVARIANT PLANE STRAIN

One cannot simply look at transformational surface upheavals (or downheavals) and infer that an invariant plane strain has occurred. A good deal of analysis is required. Another question concerns whether surface "disturbances" are actually indicative of what has happened within the material, *e.g.*, a plate of martensite. These problems are addressed in this section.

Dunne and Wayman[15] studied the formation of mechanical twins in single crystal cubes of Sn, where the twinning system is known to be $(301)[\bar{1}03]$. Twinning is clearly the simplest known IPS, and in all cases the twinning (habit) plane was determined by measuring its trace in two surfaces and then following standard stereographic techniques. The surface tilt (displacement of the invariant plane strain) was measured by scratch displacements or interference fringe counting. Given the habit plane, the displacement of three or more scratches on one surface determines the IPS displacement vector. The same can also be determined by measuring plate tilts on two surfaces (two tilt method) or, in an overdetermined way, the displacement can be obtained from tilts on three surfaces (the twin goes all the way around the corner of a cube). Their results are given in Table I. It can be seen that the angular difference between the experimental and theoretical twin plane normal ($\mathbf{p}' - (301)$) is less than one degree. Most stereographic projections are off this much. Similarly, the angle between the twinning plane and direction ($\mathbf{p}' - \mathbf{d}$) is orthogonal within experimental error, and there is very little scatter in the (correct) twinning magnitude. Of significance also is that the Bowles-Mackenzie dilatation parameter δ is unity to within four decimal places, which means that the interface is truly undistorted, as demanded. The conclusions are, unquestionably, that all three methods employed produce unusual accuracy, and that what is measured at the surface(s) is truly indicative of the behavior of the subsurface twins. Thus, the surface measurements are not superficial and may be relied upon to study the strains associated with displacive phase transformations, and moreover, distortion (if it exists) in the interface can be detected.

The next point concerns determining the magnitude and displacement of the displacement vector of the IPS. First, the habit plane must be determined by standard two-surface stereographic analysis using, for example, a geologist's optical goniometer. Then, at least three nonparallel scratches must be made (*e.g.*, diamond paste) on one surface. After transformation, these scratches are displaced by the martensite, and the displacement needs to be measured accurately (using a comparator, for example). The procedure is carried out stereographically as shown in Figure 8 where the initial and final positions of a scratch define a great circle. Three such scratches intersect at $\mathbf{d}$, the direction of displacement of the invariant plane ($\mathbf{p}'$) strain. Normalization of the vector $\mathbf{d}$ yields the magnitude of the shape strain. This procedure is laborious, but must be used for definitive results. The austenite-martensite orientation relationship is usually obtained by the microbeam Laue back-reflection method, where the distance between spots is measured accurately with a comparator.

VIII. VINDICATIONS OF THE PTMC

There are many success stories which can be told here, but time and space permit a limited, but representative, selection. One of the first involves a study of a Au-47.5 at. pct Cd alloy by the late Professor Read and associates. This alloy undergoes a cubic-to-orthorhombic

Table I. Measurements of the Twinning Shear Elements Referred to the Crystals Basis[15]

Number	p′	Angular Difference p′ − (301) (Degree)	d	Angular Difference d − [$\bar{1}$03] (Degree)	ρ′ Angular Difference p′ − d (Degree)	m	δ
Scratch displacement method							
1	$\begin{pmatrix} 3.0171 \\ -0.0635 \\ 1.0000 \end{pmatrix}$	0.99	$\begin{bmatrix} -1.0000 \\ -0.0230 \\ 3.0795 \end{bmatrix}$	0.97	89.49	0.09625 ± 0.00376	1.00030 ± 0.00203
2	$\begin{pmatrix} 3.0917 \\ -0.0323 \\ 1.0000 \end{pmatrix}$	0.92	$\begin{bmatrix} -1.0000 \\ -0.0081 \\ 2.8902 \end{bmatrix}$	1.02	91.71	0.10156 ± 0.00609	0.99898 ± 0.00342
3	$\begin{pmatrix} 3.0577 \\ 0.0424 \\ 1.0000 \end{pmatrix}$	0.84	$\begin{bmatrix} -1.0000 \\ -0.0457 \\ 2.8200 \end{bmatrix}$	2.15	92.10	0.09726 ± 0.00456	0.99881 ± 0.00157
4	$\begin{pmatrix} 3.0410 \\ -0.0353 \\ 1.0000 \end{pmatrix}$	0.67	$\begin{bmatrix} 1.0000 \\ 0.0316 \\ 3.0603 \end{bmatrix}$	1.05	89.85	0.09905 ± 0.00099	1.00008 ± 0.00045
5	$\begin{pmatrix} 3.0940 \\ -0.0022 \\ 1.0000 \end{pmatrix}$	0.78	$\begin{bmatrix} -1.0000 \\ -0.0265 \\ 3.0339 \end{bmatrix}$	0.84	90.50	0.10274 ± 0.01734	0.99971 ± 0.00361
Two tilt method							
5(a)	$\begin{pmatrix} 3.0940 \\ -0.0022 \\ 1.0000 \end{pmatrix}$	0.78	$\begin{bmatrix} 1.0000 \\ -0.0077 \\ 3.0938 \end{bmatrix}$	0.82	—	0.10305	—
Three tilt method							
6	$\begin{pmatrix} 3.0793 \\ 0.0185 \\ 1.0000 \end{pmatrix}$	0.65	$\begin{bmatrix} -1.0000 \\ 0.0096 \\ 3.0547 \end{bmatrix}$	0.47	90.18	0.09871	0.99990

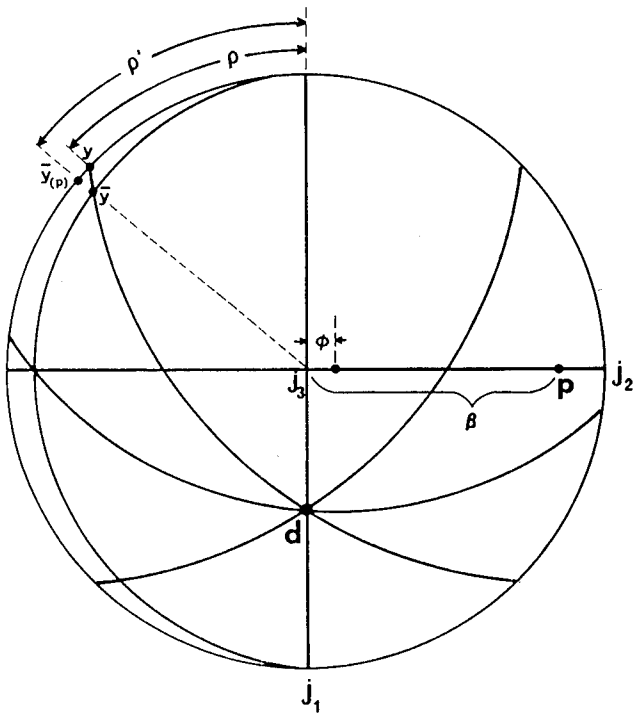


Fig. 8—Stereographic analysis of the shape-strain direction **d** in a martensitic transformation. The scratch **y** before transformation becomes transformed into the scratch $\bar{\mathbf{y}}$. **y**, and $\bar{\mathbf{y}}$ define a great circle, and three such great circles intersect at the unique point **d**, the direction of displacement of the IPS shape deformation. The habit plane is **p**.

martensitic transformation in the vicinity of 60 °C and observations on this alloy were the seeds of the WLR[10] theory. By transforming a rod-like specimen of this alloy in a temperature gradient, it was possible to freeze the parent-martensite interface in place and make experimental observations on both sides of the interface. By choosing the appropriate thermal gradient, it was possible to nucleate a single domain of martensite which propagated from one end of the original crystal by means of the passage of a single-interface habit plane. The habit plane was readily measured, and the kink of the crystal about the single interface permitted the shape deformation to be deduced. It was also noted that the martensite was twinned, seen even with a low-power optical microscope, the twinning plane being {111} of the orthorhombic martensite, which according to the lattice correspondence used was generated from the {101} plane in the cubic phase. The predicted and observed experiments are in excellent agreement.[16] Table II is taken from their work.

It could be argued that the outstanding theory-experiment agreement in the Au-Cd case is of limited validity because the habit plane, traversing the lateral dimensions of a rodlike single crystal, is totally unconstrained, *i.e.*, sees no resistance from its matrix. But the theory works equally well (in fact, better, using more refined experimental techniques) on plates of martensite formed within grains of the matrix as in steels and many other commercial alloys. A case in point here is Fe-Pt alloys near the composition Fe_3Pt as studied by Efsic and

Table II. Comparison of Experimental and Theoretical Results for Au-47.5 At. Pct Cd[16]

Quantity	Experimental	Theoretical	Difference
Habit plane, $\mathbf{p}_1$	$\begin{bmatrix} 0.696 \\ -0.686 \\ 0.213 \end{bmatrix}$	$\begin{bmatrix} 0.6968 \\ -0.6810 \\ 0.2250 \end{bmatrix}$	<1.5 deg
Direction of shape deformation, $\mathbf{d}_1$	$\begin{bmatrix} 0.660 \\ 0.729 \\ 0.183 \end{bmatrix}$	$\begin{bmatrix} 0.6510 \\ 0.7322 \\ 0.2001 \end{bmatrix}$	<1.5 deg
Angle of shape deformation shear	2.94 deg	3.28 deg	0.36 deg
Orientation relationship	$(001)_c \parallel (001)_o$	$(001)_c$ 2 deg 40 min from $(001)_o$ [twin orientation 1] $(010)_c$ 2 deg 27 min from $(010)_o$ [twin orientation 2]	2.5 deg
	$[111]_c \parallel [011]_o$	$[111]_c$ 18 min from $[011]_o$ [twin orientation 1] $[111]_c$ 18 min from $[011]_o$ [twin orientation 2]	0.3 deg

the author.[17] Habit planes were determined by two-surface analysis, the orientation relationship determined by the microbeam Laue technique, and the shape strain determined by the three-scratch displacement method as shown in Figure 8.

The experimental results from the Fe-Pt study are shown in the stereographic projection (Figure 9). Two determinations of the orientation relationship were made and found to be too close to be distinguished stereographically. The agreement with the theory was excellent. Likewise, the theoretical positions of the habit plane normal (**p**) and shape-strain direction (**d**) lay within the small scatter displayed by the experimental quantities. The predicted magnitude of the shape strain $m_1 = 0.229$ compares very favorably with the average of five experimental results, $m_1 = 0.216$, and it was concluded that the overall agreement between experimental measurements and theoretical predictions was "essentially perfect."

Another vindication of the crystallographic theory lies in its ability to explain certain martensite morphologies of the self-accommodating type. In many ferrous and nonferrous alloys as well, the plates of martensite form with a spear-like or diamondlike morphology, as shown in Figure 10 for a Cu-40 pct Zn alloy almost fully transformed to martensite, where the various spears (wedges) and diamonds (back-to-back spears) are evident in surface relief. Saburi and Wayman[18] studied the diamond morphology in six different alloy systems and came to the same conclusion in each case. Figure 11, for a Cu-Zn-Ga alloy depicts the idealized diamond morphology (a plate group). Figure 11(a) and (b) plots the four habit

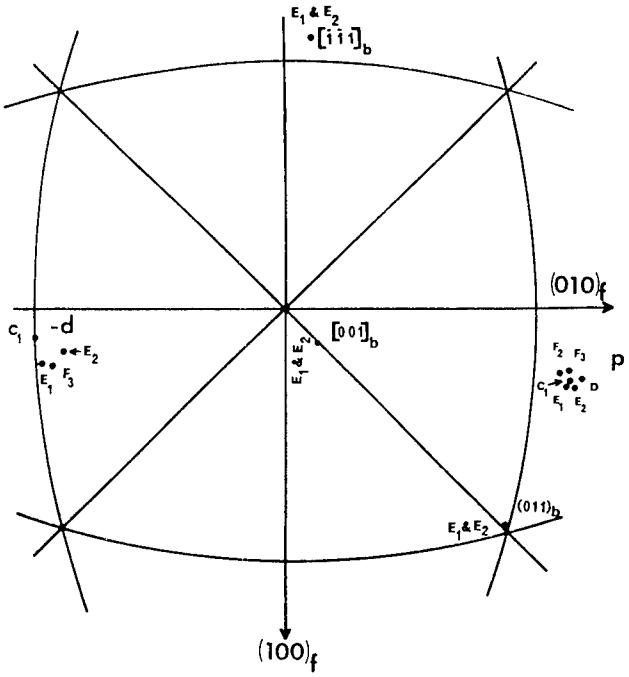


Fig. 9—Portion of a stereographic projection showing experimental values of **p** and **d** for a martensitic transformation in an Fe-24.5 at pct Pt alloy. Two measurements of the orientation relationship (E_1 and E_2) were too close to be resolved in the scale of the drawing. $-\mathbf{d}$ is plotted so that both the habit plane (**p**) and shape-strain direction appear on the upper side of the stereographic projection.[17]

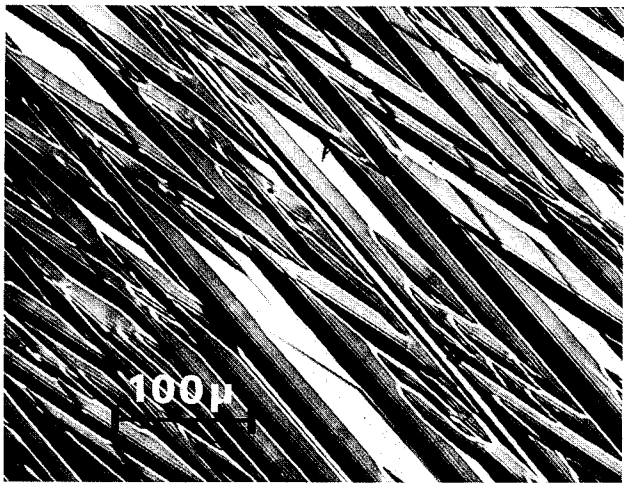


Fig. 10—Surface relief optical micrograph of martensite plates formed in a Cu-40 pct Zn alloy. The white areas are untransformed parent. Note that the plates form as pairs with a spear morphology or as back-to-back spears giving rise to a diamond morphology. These plates are self-accommodating. See text.

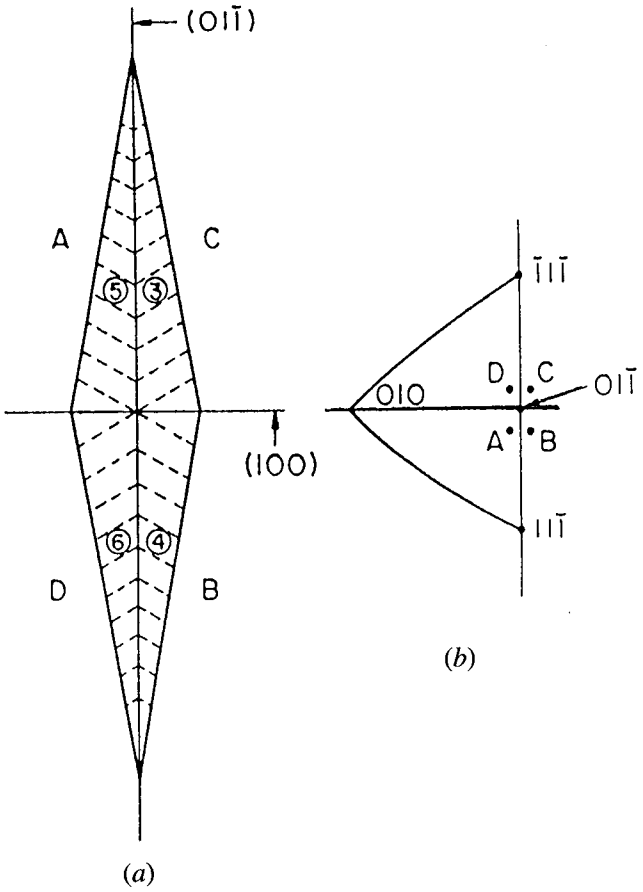


Fig. 11—Diamond-shaped plate morphology (*a*) frequently seen in martensitic transformations. The habit planes of plates **A**, **B**, $\mathbf{C}_2$ and **D** are plotted stereographically in (*b*). The planes (100) and (011) are junction planes—not habit planes.[18]

planes (experimental) clustered around a (011) pole of the parent. The (011) and (100) planes of the parent are bisectors. The observed habit planes successfully calculated using the theory and the corresponding shape-strain matrices were determined as well and are listed in Table III for variants **A**, **B**, **C**, and **D**. The last row grouping lists the numerical average of results **A**, **B**, **C**, and **D** and proves to be essentially the identity matrix

Table III. Calculated Shape Deformation Matrices[18]

	Shape Deformation, CuZnGa		
A	1.003	0.016	−0.016
	0.016	1.080	−0.079
	0.019	0.095	0.906
B	1.003	0.016	−0.016
	−0.019	0.906	0.095
	−0.016	−0.079	1.080
C	1.003	−0.016	0.016
	0.019	0.906	0.095
	0.016	−0.079	1.080
D	1.003	−0.016	0.016
	−0.016	1.080	−0.079
	−0.019	0.095	0.906
Average	1.003	0	0
	0	0.993	0.008
	0	0.008	0.993

(0.008 is very close to zero). This means that the four individual shape strains taken as a group cancel each other and that the diamond-shaped volume **ABCD** is essentially distortionless on a macroscopic scale, *i.e.*, the total transformation strain energy is essentially nil. This "cooperation" among martensite plates is common.

A detailed study of the martensitic transformations in Ti-Ta alloys has been reported by Bywater and Christian,[19] who studied compositions up to 53 wt pct Ta. They noted that martensitic transformations in Ti and its alloys are unusual in that the principal distortions of the Bain strain are rather small. One of the three is nearly unity in pure Ti and becomes exactly unity in high Ta content alloys. In such a case, the Bain strain *per se* results in an undistorted plane of contact. It was indeed observed that such plates contained no internal twins because there is no need for an inhomogeneous shear in the transformation, *i.e.*, $\mathbf{P}_1 = \mathbf{RB}$.

Watanabe and Wayman[20] studied the austenite-martensite transformation in high-carbon aluminum steels and focused on crystallographic aspects. Of particular interest was the fact that for a given carbon content, additions of aluminum produce an "excessively" high tetragonality in the martensite. This means that, theoretically, the amount of Bain "upsetting" will be less. In studying an Fe-7Al-2C (wt pct) alloy it was found that the martensite tetragonality was $c/a = 1.14$ compared to 1.08 in very high carbon steels. Because of the less severe Bain strain, the experimental shape strain magnitude was 0.12 in contrast to a value of 0.22 for plain carbon steels, Fe-Ni-C steels, and carbon-free Fe-Ni alloys. This lower value of m_1 for the Al-steels is consistent with the smaller Bain strain according to Eq. [1], *i.e.*, a smaller Bain strain gives rise to a smaller shape strain.

The above successes are not meant to imply that all martensitic transformations can be crystallographically explained using the PTMC. Indeed, for some time the well-known transformations in Cu-Al-Ni[20,21] and Ni-Ti[22] shape memory alloys remained enigmas until workers observed that these martensites were internally twinned according to a type II mechanism. Heretofore, the more familiar type I twinning was employed in typical martensite calculations. These transformations are now explained by the PTMC. There are, however, two familiar cases which have yet to be explained theoretically by the PTMC: the {111}-lath and the {225}-plate martensites found in steels.

IX. SUMMARY

Most martensitic transformations studied to date are very much in compliance with the crystallographic theory introduced in the 1950s, and accordingly a martensitic transformation may be described in the following way, as introduced by Clark and the author in 1970:[24]

"A martensitic transformation is in general described by an invariant-plane strain relief effect at a free surface, an irrational habit plane and orientation relationship, a lattice correspondence between substitutional atoms which implies the absence of long-range diffusion, and the presence of an internal inhomogeneity

such as twinning, all of these features being uniquely related in accordance with the phenomenological crystallographic theory, according to which this body of information is completely internally consistent."

In effect, according to this description, one uses all aspects of the crystallography to define martensite.

The previous description is not to say that transformations not commonly recognized as martensitic do not fit the picture. Indeed, the formation of CuAu II by ordering[25] and the formation of "suboxides" in Ta[26] conform very well to all aspects of the theory. However, CuAu II and TaO may have *lattice* correspondences and lattice correspondences are what the theory is all about. But greater rigor both in the laboratory and at the computer is required before it can be said that the PTMC does indeed apply to a given nonmartensitic transformation. Or in short, *caveat emptor*.

ACKNOWLEDGMENTS

The author is indebted to various colleagues, many of whom have attended this conference, for discussing the subject at hand with him. In addition, the author is indebted to the Air Force Office of Scientific Research, the Army Research Office, and the Electric Power Research Institute for sponsoring research on martensitic transformations.

REFERENCES

1. E.C. Bain: *Trans. AIME*, 1924, vol. 70, p. 25.
2. M.A. Jaswon and J.A. Wheeler: *Acta Crystallogr.*, 1948, vol. 1, p. 216.
3. G.V. Kurdjumov and G. Sachs: *Z. Phys.*, 1930, vol. 64, p. 325.
4. Z. Nishiyama: *Sci. Rep. Tohoku Imp. Univ.*, 1934, vol. 23, p. 638.
5. E.S. Machlin and M. Cohen: *Trans. AIME*, 1951, vol. 191, p. 746.
6. J.S. Bowles: *Acta Crystallogr.*, 1951, vol. 4, p. 162.
7. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 129, 138, and 224.
8. A.B. Greninger and A.R. Troiano: *Trans. AIME*, 1949, vol. 185, p. 590.
9. J.W. Christian: *The Theory of Transformations in Metals and Alloys*, Pergamon Press, Oxford, 1965.
10. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. AIME*, 1953, vol. 197, p. 1503.
11. R. Bullough and B.A. Bilby: *Proc. Phys. Soc. B*, 1956, vol. LXIX, p. 1276.
12. A.F. Acton and M. Bevis: *Mater. Sci. Eng.*, 1969–1970, vol. 5, p. 19.
13. N.H.D. Ross and A.G. Crocker: *Acta Metall.*, 1970, vol. 18, p. 405.
14. D.P. Dunne and C.M. Wayman: *Acta Metall.*, 1971, vol. 19, p. 2327.
15. D.P. Dunne and C.M. Wayman: *Acta Metall.*, 1970, vol. 18, p. 981.
16. D.S. Lieberman, M.S. Wechsler, and T.A. Read: *J. Appl. Phys.*, 1955, vol. 26, p. 473.
17. E.J. Efsic and C.M. Wayman: *Trans. TMS-AIME*, 1967, vol. 239, p. 873.
18. T. Saburi and C.M. Wayman: *Acta Metall.*, 1979, vol. 27, p. 979.
19. K.A. Bywater and J.W. Christian: *Phil. Mag.*, 1972, vol. 25, p. 1249.
20. M. Watanabe and C.M. Wayman: *Metall. Trans.*, 1971, vol. 2, pp. 2221 and 2229.
21. S. Ichinose, Y. Funatsu, and K. Otsuka: *Acta Metall.*, 1985, vol. 33, p. 1613.
22. K. Otsuka: *Proc. ICOMAT-86, Int. Conf. on Martensitic Transformations*, The Japan Institute of Metals, Sendai, Japan, 1987, p. 35.
23. K.M. Knowles and D.A. Smith: *Acta Metall.*, 1981, vol. 29, p. 101.
24. H.M. Clark and C.M. Wayman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 59.
25. R. Smith and J.S. Bowles: *Acta Metall.*, 1960, vol. 8, p. 405.
26. J. Van Landuyt and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, pp. 803 and 815.

A History of the Controversy over the Roles of Shear and Diffusion in Plate Formation above M_d and a Comparison of the Atomic Mechanisms of These Processes

H.I. AARONSON and M.G. HALL

The evolution of the controversy over the roles played by shear and by diffusion in the formation of plate-shaped transformation products above M_d is traced along its principal pathways. Growth by shear is taken to occur through the glide of dislocations, whereas diffusional growth is considered to occur by means of the biased random walk of individual atoms. Both mechanisms are based upon growth ledges, but during shear, risers, and in some cases terraces, are mobile in non-fcc/hcp transformations; however, during diffusional growth, only the risers (and more likely, only kinks on the risers) are mobile when there is a stacking sequence difference across the terrace of a growth ledge. The aura of the powerful and elegant phenomenological theory of martensite crystallography (PTMC) appears to have persuaded many researchers that even above M_d, plates usually form by shear because one or more of the multiple requirements for the applicability of the PTMC appears to be fulfilled. As the preceding article by Wayman in these conference proceedings emphasizes, however, *all* of these requirements must be fulfilled in a self-consistent manner if a plate is to qualify as the product of a shear transformation mechanism. Successful application of the invariant line (IL) component of the PTMC to precipitate plates and needles or rods by Dahmen and co-workers provides a phenomenological reason why the crystallographies and surface reliefs of ledgewise shear and ledgewise diffusional growth sometimes appear to be identical. While the IL-containing broad faces of a martensite plate minimize the transformation strain energy, the same situation in plates formed by ledgewise diffusional growth tentatively appears to engender a minimum in the kinetics of growth ledge generation and/or the lateral growth kinetics of such ledges.

I. INTRODUCTION

AS part of "setting the stage" for the proceedings of this conference on the roles of shear and diffusion in the atomic processes involved in the formation of plate-shaped transformation products, it seems appropriate to include a fairly detailed history of the origins of this controversy and the various paths along which it has developed. This historical account takes up most of the present article. Because a number of opportunities has recently become available for discussing in detail the current status of the shear and the ledgewise diffusional views on plate formation,[1–6] the historical account is followed by only a brief comparison of these two viewpoints.

II. A HISTORY OF THE SHEAR *vs* DIFFUSION CONTROVERSY

A. *Early History (1929 to 1934)*

Perhaps the first clear intimation that the carbide-containing ferrite plates developed at intermediate reaction temperature in steel may form by high-velocity shear is due to Robertson,[7] who proposed in 1929 that the progressive darkening of the optical microstructure with decreasing reaction temperature results from increasing amounts of carbide precipitation consequent upon the taking of larger proportions of carbon into solid solution by growing ferrite plates. However, the controversy itself may be fairly said to have begun the following year, when Davenport and Bain[8] essayed a similar hypothesis. Nonetheless, Hoyt,[9] in a discussion to their classic article, concluded that the configuration of what is now known as the time-temperature-transformation (TTT) diagram demonstrated that the transformation products that were not clearly martensite must have formed by a nucleation and growth process, controlled by the kinetics of nucleation, and thus, implicitly, by a diffusional mechanism. And in 1931, when Mehl and Barrett[10] examined the crystallography of hexagonal close packed (hcp) γ $AlAg_2$ formation from face-centered cubic (fcc) α Al-Ag, they concluded that while the transformation mechanism could be represented as a shear process, uncertainty remained as to whether the actual atomic transport during the crystal structure change takes place by shear or diffusion. Several investigators [10A, 10B, 10C], recognizing the contradiction inherent in the concept of a transformation requiring a composition change taking place by shear, proposed that the necessary diffusion takes place in the matrix phase prior to the occurrence of the phase transformation.

H.I. AARONSON is with Geo-Centers, Inc., Ft. Washington, MD, and is stationed in the Physical Metallurgy Branch, Code 6320, Naval Research Laboratory, Washington, D.C. 20375-5343. M.G. HALL, Lecturer, is with the School of Metallurgy and Materials, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

B. *Surface Relief-Based History (1952 to Date)*

1. *Introduction*

An invariant plane strain (IPS) surface relief effect is produced when the interception of a growing plate by a free surface that was polished flat prior to transformation causes the intercepted area to rotate about a line of intersection between the plate and the surface, as illustrated in Figure 1.[11] Note that the rotated surface remains flat. Thus, a linear scratch present on the surface before transformation has occurred will be abruptly deflected—while remaining straight—as it passes over the transformed area of the surface. The scratch resumes its original direction upon returning to the matrix phase but is now laterally displaced to position STT′S′ in Figure 1. (See the article by Wayman[12] in these conference proceedings for a more detailed discussion of the IPS concept and of the measurements that must be made in order to confirm the presence of such a surface relief.)

During a shear or martensitic transformation, it follows automatically—at least in substitutional alloys—that neither a change in composition nor a change from long-range order to a different level of long-range order, short-range order, or disorder can accompany the transformation, because the transformation mechanism is based upon dislocation glide rather than any diffusional process.[13,14,15] Hence, in the following historical account, the occurrence of either of these changes in conjunction with an IPS surface relief is seen to run directly counter to the soon widely held belief that such a relief, when accompanied by an appreciable shear component, provides a unique identification of a martensitic transformation mechanism at the atomic scale. These contradictions, repeatedly observed and still more often disputed, have provided one of the principal "foundation stones" of the present conference.

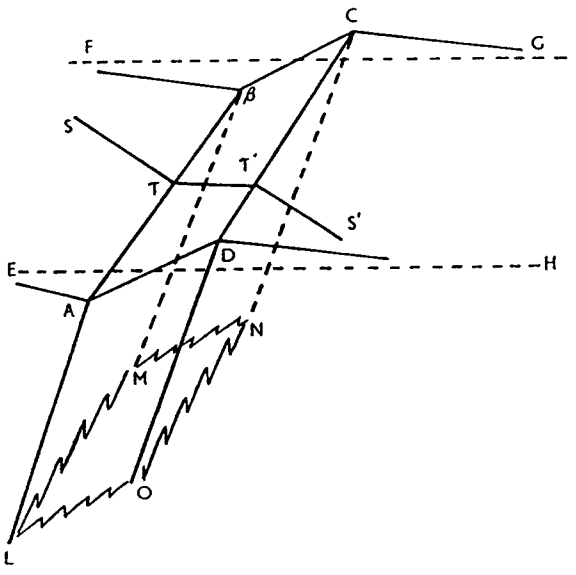


Fig. 1—The shape deformation (and invariant plane strain surface relief effect—ABCD) produced by a martensite plate (ABCDOLMN) intercepting a free surface (defined by the parallel dashed lines EH and FG). Note that the tilted surface ABCD remains planar and causes an initially straight scratch to be deflected to TT′ where it crosses the martensite plate and then laterally displaced to final position T′S′.[11]

2. *Formation of Invariant Plane Strain (IPS) surface reliefs during growth of plates with a change in composition or in long-range order*

In 1952, Ko and Cottrell[16] observed the growth of lower bainite at temperature with hot-stage optical microscopy. While contrary to expectation, *e.g.* References 7 and 8, the bainite structures were observed to grow slowly, their formation was seen to be accompanied by a pronouncedly geometric surface upheaval qualitatively similar to that associated with martensite plate formation at a free surface. The following year Ko[17] reported an equivalent surface relief observation on Widmanstatten ferrite plates formed at a much higher temperature. These results led Ko to propose that in interstitial alloys, displacement of the substitutional atoms forming the crystallographic frameworks of the matrix and product phases could take place by shear whereas that of the interstitial atoms could occur simultaneously by means of diffusion.

Also in 1953, Wechsler *et al.*[18] and Bowles and Mackenzie[19] presented in different, but essentially equivalent,[11,13] forms what is now known as the phenomenological theory of martensite crystallography (PTMC). This theory defined the surface relief attending formation of a martensite plate as an IPS type of shape strain. Hull[20] then proposed that the appearance of an IPS relief be used to identify the operation of the shear mode of transformation. In a particularly influential early review, which appeared in 1956, Bilby and Christian[13] strongly re-enforced this view and emphasized that "The change of shape, with its extra energy, is apparently tolerated only because there exists an easy growth mechanism, not requiring atomic diffusion, which leads to the rapid production of a new phase and a net lowering of the free energy." They immediately noted that the absence of a change in composition or in long-range order follow directly. Growth of martensite plates by glissile dislocation arrays in their interphase boundaries was taken almost for granted in this review, because this seemed to be the only means available for producing the near-sonic growth kinetics sometimes observed.[21] The first attempt to derive such a mobile interfacial structure was made 3 years earlier by Frank.[22]

The problem of applying the PTMC to transformations producing plate-shaped products, taking place above the temperature range normally associated with martensite, was first addressed in detail by Christian.[14] He expressed skepticism about the existence of sessile dislocation boundaries at the interfaces of such plates. Christian concluded that if such interfacial structures do exist and if they can be displaced normal to themselves, say by dislocation climb, the IPS surface reliefs associated with them would disappear as quickly as they form under the driving force of strain energy reduction, because the diffusion distances involved in growth and IPS relief elimination are comparable. However, when the diffusion distance is small relative to that required to eliminate the IPS relief, as in the formation of the long-range ordered AuCu II plates from disordered Au-50 at. pct Cu,[23] the relief was expected to survive.

Christian[24] further developed his views on the transformation mechanism of plates formed above the usual

martensite range in another important review. The presence of a shear component in the IPS relief was taken to demonstrate the presence of a correspondence between lattice sites but not necessarily between individual atoms in the matrix and product phases,[29] as earlier noted by Bowles and Mackenzie.[19] Further, the IPS relief was proposed to appear whether the interphase boundary has a fully coherent, partially coherent, or an incoherent structure with the shear component being absent only in the latter case.[24] However, Christian has privately remarked to us that when the shear component is not present, the free surface will not be tilted, and thus, the characteristic IPS surface relief will not appear. These two statements represent a further development of the views of Hull[20] and of Bilby and Christian[11,13] on the significance of the IPS surface relief effect. Comment was also included on an article by Laird and Aaronson,[25] in which a possible IPS surface relief effect was identified in a transformation, the precipitation of hcp γ $AlAg_2$ from fcc α Al-Ag, wherein a composition change supposedly occurs throughout the growth process. (The composition change was subsequently supported by analysis of the kinetics of hot-stage transmission electron microscopy (TEM) measurements of the growth kinetics of individual ledges showing that they are controlled by the volume interdiffusion of Al and Ag atoms.[26]) Christian noted, though, that the details of the surface relief had not been clearly defined. Liu and Aaronson[27] subsequently used interferometry to show that the shape change is initially the half-twinning shear expected to accompany an fcc/hcp transformation. They also measured the tilt angle as a function of isothermal reaction time. As shown in Figure 2,[27] the surface relief did diminish, but so did orders of magnitude more slowly than originally predicted.[14] They remarked that the prolonged survival of the shape change is simply the result of the driving force for relaxation being considerably less than that for formation. Christian[28] now accepts that a lattice correspondence "cannot easily be destroyed in the growth process during transformations which clearly involve diffusion."

Observations on surface relief effects, which included

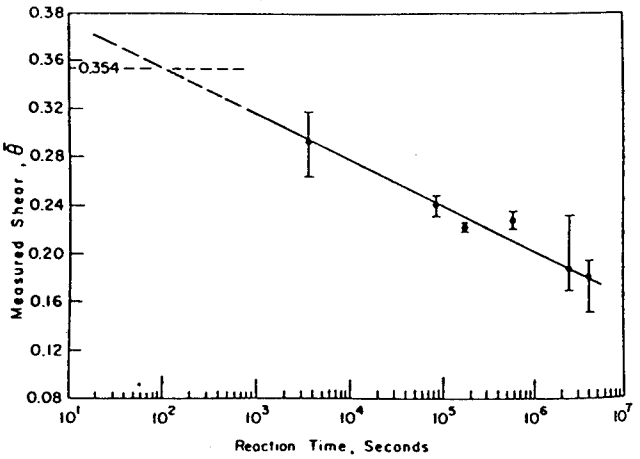


Fig. 2—Tilt angle of the surface relief generated by the interception of a polished surface by γ $AlAg_2$ plates as a function of the logarithm of the isothermal reaction time.[27]

a shear component attending plate formation in binary substitutional alloys, were made some years earlier. Probably because the transformation crystallography tended to agree with the PTMC; however, the implication was usually drawn that, at least during the stage of transformation studied, the composition of the precipitates was the same as that of the matrix. Thus, the report by Bowles and Tegart[29] that B2-ordered CuBe plates precipitated from fcc α Cu-4.4 wt pct Be yielded a relief effect similar to that produced by martensite plates during the early stages in the growth of the CuBe plates was interpreted by Christian[14] in terms of an implied absence of a composition change.

Similarly, the formation of surface reliefs quantitatively similar to those associated with martensite plates formed at subzero temperatures in β Cu-Zn alloys (where α_1' Cu-Zn plates meet a prepolished surface[30]) led initially to hints that the composition of these plates might not differ from that of their matrix during the initial stages of growth.[14,29,30] Warlimont[29A] noted that the crystal structure of α Cu-Zn plates is actually orthorhombic, the same as that of martensite in these alloys. Also, as in this martensite, the α Cu-Zn plates are densely arrayed with stacking faults.[31,32] Cornelis and Wayman[33] redesignated these plates as α_1 and confirmed that their surface reliefs are of the IPS type. Flewitt and Towner[34] reported the first microanalysis of α_1 plate composition. Their data appeared to support early suggestions that the composition of these plates is initially the same as that of their ordered β matrix. However, it was soon pointed out that the effective spot diameter of the electron probe used by Flewitt and Towner was comparable to the sum of the thickness of a plate and of the associated solute concentration field.[35] Cornelis and Wayman[36] repeated the composition analysis study with a higher resolution instrument, using thin foils instead of the bulk specimens employed by Flewitt and Towner (thereby permitting achievement of a smaller effective "spot" diameter), but obtained the same result as their predecessors: the zinc concentration in α_1 is the same as that in the ordered β matrix until a later stage in plate thickening. Lorimer et al.[37] used an analytical electron microscope with a still higher spatial resolution to re-examine specimens used by Cornelis and Wayman. They secured an entirely different result: the α_1 plates had a lower zinc concentration than that of their matrix phase even at the smallest thickness at which their composition could be reliably determined, *e.g.*, about 300 nm. Recently, Hamada *et al.*[38] have succeeded in measuring the composition of α_1 plates as thin as about 3 nm and demonstrated that they, too, contained a significantly smaller zinc concentration than the immediately adjacent ordered β matrix. Thus, both γ $AlAg_2$ and α_1 Cu-Zn plates have now been shown to differ in composition from their matrix phase at early stages in their thickening and yet to yield the IPS surface relief effect-cum-shear component characteristic of a martensitic crystallography.

In the same context of a violation of a direct consequence of a martensitic transformation mechanism, Smith and Bowles[23] have shown that the transformation crystallography and the IPS surface relief effect of AuCu II plates are predicted by the PTMC to a remarkably high level of accuracy despite the circumstance that the Au-50

at. pct Cu matrix is disordered and AuCu II is a face-centered orthorhombic superlattice incorporating 10 orthorhombic unit cells. Aaronson and Kinsman[39] called AuCu II plates the product of a massive transformation. Bowles and Wayman,[40] on the other hand, continued to accept these plates as the product of a shear transformation occurring "by the gliding of transformation dislocations at a rate determined by the rate of ordering at the dislocations."[41]

3. *Non-IPS surface relief effects associated with growth of monocrystalline plates above M_d*

Christian[14] characterized an IPS shape change "as the tilting of an originally flat, polished surface.. . ." Now suppose that the surface relief consists of a tent, each of whose sides can be described as resulting from the "tilting of an originally flat, polished surface." An obvious explanation for a tent-shaped relief is so that it arises from two plates formed face-to-face.[42]* Plates or laths of proeutectoid ferrite in Fe-C alloys,[43–47] chromium-rich precipitates formed in copper-rich Cu-Cr alloys,[48] and proeutectoid α in a hypoeutectoid Ti-Cr alloy[49] are among those precipitates producing tent-shaped reliefs. Direct experimental evidence, however, has shown that both ferrite[44,46,47] and proeutectoid α'[49] plates producing such reliefs are single crystals. The TEM studies of chromium-rich plates and laths in Cu-Cr alloys have provided no evidence that they are actually bicrystals.[50,51] Inspection of very large numbers of reliefs produced by ferrite plates in Fe-C alloys has further shown that even tent-shaped reliefs are rare; the usual relief morphology tends instead to be considerably more complex.[44,52] Hence, these observations of surface reliefs consisting of two or more tilted planar surfaces associated with monocrystalline plates are in direct conflict with the requirements and predictions of the PTMC—but also continue to be suggestive of at least a formalistic linkage between plate formation by shear and diffusional processes.

*Christian[28] notes that a monocrystalline tent-shaped relief can develop only when there is "a degeneracy in the crystallographic solution." A near-degeneracy will tend to produce "banded structures (*e.g.*, In-Tl)," *i.e.*, tents of "indefinite length."

C. *Transformation Crystallography-Based History (1954 to Date)*

The PTMC requires as "input data" only the crystal structures of the parent and product phases, the lattice parameter of each, the choice of a lattice correspondence based upon strain energy minimization, and the plane and direction of the lattice invariant deformation (LID) if this occurs by slip or, Christian privately notes, the plane *or* direction of the twinning shear if the LID takes place in this manner. The theory yields, as "output," the lattice orientation relationship, habit plane, shape strain (of which the surface relief is the manifestation thereof at the free surface), and clues as to the lattice invariant deformation (slip or twinning) and the interphase boundary structure required to make the transformation mechanistically feasible. The surface relief effect has been separately discussed in Section B because it is easy to observe and report, *e.g.*, Reference 53, particularly when the detailed analysis, noted by Wayman in the preceeding article as necessary to identify an IPS, is not performed. In the present Section, orientation relationship and habit plane "output data" obtained from application of the PTMC to transformations occurring above the M_d temperature will be considered.

Publication of the PTMC by Wechsler *et al.*[18] and Bowles and Mackenzie[19] in 1953 was quickly followed by attempts to apply the theory to plates formed in transformations above the martensite range where diffusion accompanying the transformation is a realistic possibility and thus where the PTMC might be considered inapplicable. In the context of quantitative measurement of one or more "output parameters" and comparison with calculations based upon the theory, the first to essay such a contribution appears to have been Garwood,[30] who in 1954 measured the habit plane of what are now known as α_1 Cu-Zn plates. He found the habit plane to be essentially the same as that of martensite formed at subzero temperatures in Cu-Zn alloys.[54] On this basis, Garwood suggested that the Ko–Cottrell mechanism of shear and diffusion is operative during the formation of σ_1 Cu-Zn plates. (The diffusional aspects of this proposal were later criticized in detail.[35]) Twenty years later, Cornelis and Wayman[33] confirmed Garwood's habit plane result and added that the orientation relationship is also the same as that of subzero martensite, an IPS surface relief is present, and irregularly spaced stacking faults in the plates imply the motion of the relevant partial dislocations that provide the lattice invariant deformation, just as they do in the martensite plates formed at much lower temperatures. In 1979, Kostic and Hawbolt[55] obtained equivalent results from the "parallel" formation of α_1 Ag-Cd plates.

Returning to chronological sequencing, Bowles and Tegart[29] determined the orientation relationship and habit plane for CuBe precipitated from an α Cu-Be alloy 1955 and found them in agreement with the predictions of the PTMC. Otte and Massalski[56] analyzed published habit plane data on linear $\alpha:\beta$ Cu-Zn boundaries[57] and on the broad faces of Fe_4N and Fe_3P plates.[58] They obtained reasonable or better accountings for these data through application of the PTMC, even though the experiment on two-phase brass was conducted on an annealed rather than on a transformed microstructure. Smith and Bowles[23] reported in 1960 that AuCu II plates formed from a disordered Au-50 at. pct Cu matrix fulfilled the habit plane, orientation relationship, surface relief, and internal inhomogeneity specifications of the PTMC with exceptional accuracy. In one of the articles in these proceedings, Muddle *et al.*[61] critically review the various crystallographic aspects of α_1 Cu-Zn, γ $AlAg_2$, and AuCu II plate formation. They appear to accept that the first two transformations exhibit a composition change and the third a change in long-range order throughout the growth process, but he concludes that all three transformations are nonetheless consistent with the PTMC in the sense of their transformation crystallography.

Also in 1960, Bowles and Kennon[62] investigated the bainite reaction from the standpoint of the PTMC. They were able to account, on this basis, for the transformation crystallography of lower bainite but not of upper bainite. Evidently on the assumption that the proportion

of carbon in the matrix inherited by the ferritic component of bainite in steel increases with decreasing reaction temperature, Christian[24] proposed in 1962 that the habit plane of upper bainite should be similar to that of low-carbon martensite, while that of lower bainite ought to be that of martensite in a steel of the same composition. The habit plane determinations of Smith and Mehl[63] were noted to agree with this prediction. In 1968, Srinivasan and Wayman[64,65] determined the orientation relationship, habit plane, shape strain, and three-dimensional shape of lower bainite plates in a high-chromium steel. In order to account for the habit plane thus determined, they found that it was necessary to assume either a counter-physical isotropic contraction of this plane or irrational LID modes. Despite the absence of appreciable amounts of retained austenite in the Fe-0.4 wt pct C alloy used, Watson and McDougall[45] succeeded in determining the orientation relationship and habit plane as well as the surface tilt of proeutectoid ferrite plates in 1973. They interpreted these data in terms of a martensitic mechanism for the formation of these plates. However, their habit plane data showed unusual scatter. Some of the ferrite plates examined exhibited tent-shaped reliefs, as previously noted (and no attempt was made to ascertain experimentally whether or not such plates were single crystals or bicrystals[66]). And Wayman[67] has pointed out that their orientation relationship and shape strain data are inconsistent. Later, the shear measured by Watson and McDougall was shown to correspond to a strain energy too high to permit plate formation by shear throughout most of the proeutectoid ferrite range in their alloy.[68] Hoekstra[69] and Sandvik,[70] in 1980 and 1982, respectively, made careful measurements of the habit plane of bainite in two quite different steels but were unable to explain them on the basis of the PTMC when using rational LIDs.

Whereas attempts to explain the transformation crystallography of proeutectoid ferrite and the ferritic component of bainite with the PTMC have proved, in the main, to be disappointing, Bowles and co-workers[71,72] reported in 1977 on the examination of another interstitial alloy transformation, the formation of β $VH_{0.45}$ plates from the V-H solid solution, with quite different results. The surface relief, orientation relationship, habit plane, and even the interphase boundary structure were accurately consistent with the PTMC. These results are in good accord with the Ko–Cottrell[16,17] mechanism of the shear of the substitutional atoms "framework" and the diffusion of the interstitial species.

D. *Growth Kinetics and Interphase Boundary Structure-Based History (1962 to Date)*

1. *Introduction*

Attempting to compile this subsection as a single chronological account yielded too complex a result. Accordingly, this material has been broken down into logical subdivisions, whose components are then usually presented in order of publication.

2. *Interphase boundary structure*

a. *Introduction*

Except for identifying the original discoverers of the linear misfit-compensating defects, comprising (with the coherent areas between them) the structure of partially coherent interphase boundaries, and describing each linear defect, these historical summaries will focus upon the question as to whether the defect structure is glissile (thus capable of supporting transformation by shear) or sessile (thereby restricting the transformation mechanism to diffusional growth).

b. *At broad faces of plates*

Misfit dislocations

The misfit dislocation was originally predicted theoretically by Frank and van der Merwe.[73,74] The first direct observations of this linear defect were made in 1961 with TEM on oriented overgrowths. Matthews[75] observed misfit dislocations between PbSe and PbS, successively deposited from the vapor phase, and at nearly the same time, Delavignette *et al.*[76] made similar observations upon interphase boundaries formed by an unknown compound deposited upon $CrBr_2$. As mentioned in Section B–2 in 1962 Christian[14] expressed doubt about the existence of interphase boundaries with a structure whose movement must be accompanied by dislocation climb. In the same symposium proceedings, Aaronson[77] proposed a dislocation structure for fcc:hcp interfaces formed by $\{0001\}_{hcp}//\{111\}_{fcc}$, $\langle 1120\rangle_{hcp}//\langle 110\rangle_{fcc}$ that must be displaced by climb (or an equivalent process) in the direction normal to the interphase boundary.

In 1964, Whitton[78] made the first observation of misfit dislocations on the broad faces of precipitate plates. He found these dislocations on the broad faces of UC_2 plates that had precipitated from UC. The dislocations were arrayed in two orthogonal sets and were described by $\mathbf{b} = a\langle 100\rangle$, with $\mathbf{b}$ parallel to the $\{100\}$ interface plane. Hence, these boundaries had a sessile misfit dislocation structure. Three years later, Laird and Aaronson[25] determined that the misfit dislocations on the (0001) broad faces of γ $AlAg_2$ plates are $a/6\langle 112\rangle$ Shockley partials in a fcc α Al-Ag alloy. These dislocations appear in arrays of one, two, or three sets of equivalent but nonparallel Burgers vectors. In order to maintain the correct stacking sequence, successive parallel partials must be stepped down by two $\{111\}_{fcc}$ planes with respect to each other. When but a single array of partials appears in one area of a $(0001)_{hcp}$ interface, a built-in array of growth ledges is thus present. These misfit dislocation ledges are sessile with respect to growth normal to the plate broad faces but glissile from the standpoint of ledgewise growth. A similar situation evidently obtains during the fcc-to-hcp martensite transformation in cobalt and Co-Ni alloys, though in the absence of TEM observations, some preference was expressed for the dislocation pole mechanism.[79] However, the large compositional width of the $\alpha + \gamma$ region and its roughly vertical boundaries at low temperatures ensure that the aging temperatures and alloy compositions normally used to study the interfaces of γ $AlAg_2$ plates lie well above the T_0 temperature, the thermodynamic upper-limiting temperature for a shear transformation.[80] Hence, the Shockley partials at the broad faces of $AlAg_2$ plates cannot move by simple glide but can only be displaced by kink motion along these dislocation ledges at rates limited by the concurrent volume interdiffusion of aluminum and silver.[26,35]

From the standpoint of this conference, the interfacial

structure of the broad faces of ferrite (sometimes called bainite) plates is of especial importance. In 1979, Rigsbee and Aaronson[81] reported that ferrite plates formed at 450 °C and 475 °C in an Fe-0.62 wt pct C-2.0 wt pct Si alloy contain two types of linear misfit-compensating defects in their broad faces. One type is misfit dislocations with $\mathbf{b} = a/2\langle 110\rangle_{fcc} = a/2\langle 111\rangle_{bcc}$; these dislocations appear in a single parallel array and are spaced 1.5 to 2.5 nm apart. Their Burgers vector is parallel to the $\{111\}_{fcc}//\{110\}_{bcc}$ atomic terrace planes. Hence, these dislocations cannot produce an LID, and therefore, they must be sessile with respect to the shear mode of transformation. However, Christian privately points out an alternative interpretation. Because the slip plane of the Burgers vector of the misfit dislocations makes a small but nonzero angle with the counterpart {110} plane in ferrite formed martensitically, these dislocations could serve as a vehicle for growth by shear. It is not clear, though, that this mechanism will have sufficient driving force to be operative at temperatures above M_s.

Two years earlier, Bowles and co-workers[71,72] found a single set of screw dislocations at the broad faces of β $VH_{0.45}$ plates with just the spacing and the Burgers vector orientation needed to accomplish the LID by glide. This appears to be the only nonmartensitic transformation so far identified in which the interfacial structure is unequivocally appropriate for transformation by shear, even though hydrogen diffusion necessarily accompanies the shear process. Christian[28] has also discussed this situation.

Structural ledges

Hall *et al.*[82] predicted in 1972 the existence of structural ledges on the basis of hard-sphere modeling. They noted that at $\{111\}_{fcc}//\{110\}_{bcc}$, interfaces for orientation relationships from Kurdjumow and Sachs[83] to Nishiyama[84] and Wasserman,[85] introducing ledges from one to three atomic planes high causes coherent areas to be repeated immediately on the terrace of each successive ledge instead of relatively far-away, as at a planar interface. Shortly afterward, Russell *et al.*[86] recognized that the regions between parallel rows of coherent areas contain one more plane of atoms in one phase than in the other. These misfit dislocations complete construction of partially coherent interfaces. In 1979, Rigsbee and Aaronson[81] confirmed the structural ledge model by applying the weak-beam, dark-field, TEM technique of Cockayne[87] to study the broad faces of ferrite sideplates embedded in a retained austenite matrix in an Fe-0.6 wt pct C-2 wt pct Si alloy. A computerized model of structurally ledged interfaces and their attendant misfit dislocations was found to give a good quantitative accounting for the interdefect spacing and angle measurements made on the interfaces studied.[88] The facts that the misfit dislocations in these interfaces are either pure edge or contain an edge component and hence are unable to translate the terraces on which they reside in the normal direction and that the structural ledges are incapable of long-range movement were emphasized as inconsistent with a shear mechanism of growth. Purdy and co-workers[89,90,91] later confirmed that a Burgers vector lies in the plane of the broad faces of ferrite and bainite plates formed in alloy steels. Their studies were probably conducted on structural ledges, but it is possible that they were instead performed on misfit dislocations.

Recently, Furuhara *et al.* have used high-resolution transmission electron microscopy (HRTEM)[92] and computer modeling[93] to demonstrate the existence of structural ledges at bcc:hcp boundaries between proeutectoid α plate broad faces and their β matrix. It is now recognized that structural ledges may be formed on the basis of closely matched conjugate habit planes, which would serve as their terraces, whenever there is a change in stacking sequence across these planes.[88,93]

In 1984, Christian and Edmonds (CE)[94] published a valuably provocative critical review of the literature on the bainite reaction. During 1988 and 1989, Aaronson and co-workers[95,96,97] and Howe *et al.*[98] responded to this stimulation, and CE[99,100,101] provided suitable rejoinders. The issues raised by CE became key issues for discussion at the present conference. One was that CE questioned the significance of structural ledges. Christian[101A] continues to consider them as just "a necessary description of a high-index (in the limit, irrational) plane or interface.. . ." The Rigsbee–Aaronson[81] sessile interfacial structures were reinterpreted as glissile, composed only of one array of misfit dislocations lying nearly in screw orientation.[100]

The following response was given to this criticism. Both the deflection of extinction contours by successive structural ledges and the close agreement achieved between the detailed modeling of structural ledges and misfit dislocations and the TEM observations made upon these misfit-compensating defects strongly support the existence of structural ledges as a special, misfit-compensating structural entity.

Misfit-compensating ledges

In their 1991 articles, Furuhara and co-workers[92,93,102] reported the discovery of another type of linear misfit-compensating defect at hcp:bcc interfaces in their hypoeutectoid Ti-Cr alloy, which they named the "misfit-compensating ledge" because it shares the characteristics of structural ledges and misfit dislocations. They found that such ledges have only enough mobility to wriggle into occasional contact with one (or possibly both) of their nearest parallel neighbors. Mass movement of these ledges would require the same relatively long-term average concentration of kinks at each, a rather unlikely possibility.

Two types of linear misfit-compensating defects have thus been identified on the broad faces of hcp proeutectoid α plates in the bcc β matrix of a Ti-Cr alloy: structural ledges and misfit-compensating ledges. No misfit dislocations were identified.[92] Because neither is able to do more than locally wriggle, such boundaries are effectively sessile. Unlike fcc:bcc interfaces, however, wherein rotating the Burgers vector of the misfit dislocations would make the interface glissile,[103] no such flexibility exists at these bcc:hcp interfaces.

Structural dislocations

Although silicon crystals precipitated from aluminum-rich Al-Si solid solutions were first considered to be prototypes for incoherent nucleation and growth,[104] Westmacott and Dahmen[105] used a combination of HRTEM and modeling to show that four aluminum atoms on a {111} plane of the fcc Al-Si matrix match almost perfectly three silicon atoms in the conjugate plane of

the diamond cubic silicon lattice. Only the residual misfit is taken up by misfit dislocations. The dislocations formed by the 4 *vs* 3 configuration have been termed "structural dislocations"[106] because it appears impossible to eliminate them by any elastic shear strains parallel to their Burgers vector which both lattices could tolerate without plastic deformation and/or fracture. Because there are no coherent areas between these dislocations, in principle they should be able to undergo synchronous climb. Perhaps because this would require such a high vacancy concentration in the vicinity of the boundary, however, these boundaries appear to be flat on the two-dimensional atomic scale[105]—strongly suggesting that they are immobile in the absence of growth ledges.

c. *At plate edges*

Whereas the edges of a martensite plate are regarded, in effect, as no more than a mathematical curve because only the terrace plane of the (apparent) invariant plane can be present if strain energy minimization is to be achieved,[11,13] a distinctly different situation obtains in the case of plates formed by diffusional nucleation and growth. When the edges of precipitate plates and laths have been observed with optical microscopy[107] or with either replication[108] or transmission electron microscopy,[109] they have been found to exhibit small but nonetheless finite widths. The earliest report of characterized misfit-compensating defects at plate edges appears to be that of Matsuura and Koda[110] who found these defects on θ' Al-Cu plates to be misfit dislocations with $\mathbf{b} = a\langle 001 \rangle$. This finding was subsequently confirmed by Aaronson and Laird.[109] These dislocations girdle the perimeter of θ' plates. Laird and Aaronson[25] observed some form of linear defect at the edges of γ $AlAg_2$ plates. Howe *et al.*[111] subsequently characterized these defects as Shockley partials, arrayed as superledges; these partials are also present on the broad faces of these plates.[25,112] Lee and Aaronson[113] found what appeared to be uniformly spaced misfit dislocations and ledges (neither of which was characterized) at the leading edges of hcp α laths embedded in their bcc β matrix in a eutectoid Ti-Fe alloy. Again, the same defects were apparently present at the sides of the laths but now much more widely and irregularly spaced. Subsequently, Furuhara *et al.*[92] examined the edges of "normal α" plates in a Ti-Cr alloy. They identified misfit-compensating ledges and also another, unknown type of ledge at these edges. While the Burgers vector of the misfit-compensating ledges lies at an angle to the apparent habit plane of the edges, the simultaneous movement stricture should still apply to these ledges as well as to those at the plate broad faces. Hence even without consideration of the unknown ledges, the plate edges in this transformation should also be sessile.

d. *At grain boundary allotriomorphs*

On an early point of view, grain boundary allotriomorphs contain large areas of disordered or incoherent boundary.[77,114] Shewmon[115] appears to have been the first to recognize that this view may not be correct because of the extensive presence of facets on both sides of ferrite allotriomorphs, as discerned from published micrographs of these crystals.[77] However, because of the torturous path that faceted allotriomorph interfaces follow and the tendency of allotriomorphs to deviate from exact orientation relationships because of the need to compromise between the crystallographies of the two matrix grains forming a grain face and that of the critical nucleus,[116] not until recently was the first effort successfully made to examine the interphase boundary structure of this morphology. Furuhara and Aaronson,[102] again using the proeutectoid α reaction in the same hypoeutectoid Ti-Cr alloy, found that both the allotriomorph interfaces formed by rational or near-rational orientation relationships and those formed by clearly irrational orientation relationships contained misfit-compensating ledges. Comparisons with the parallel computer-modeling study[93] as well as faint, closely spaced (2 nm) lines in the interfaces indicated that structural ledges were probably present at many of these boundaries. Hence, consistently with the prediction of an analysis of the nucleation kinetics of proeutectoid ferrite at austenite grain faces in high-purity Fe-C alloys,[117] both interfaces of grain boundary allotriomorphs, almost certainly in alloy systems in general, can be considered as partially coherent, even though ferrite allotriomorphs do not display (at least at the magnifications provided by optical microscopy) an IPS or any other obvious type of "geometric" surface relief effect.[17]

3. *Extrinsic interphase boundary structure: Growth ledges and intruder dislocations*

a. *Intruder dislocations*

An intruder dislocation may have the wrong Burgers vector or the wrong orientation of its Burgers vector relative to the other members of an array of misfit dislocations. An intruder dislocation may also have the correct Burgers vector and orientation thereof but be too close to one or both of its nearest neighboring misfit dislocations. In the context of this section, a Burgers vector orientation out of the plane of the interface will produce a ledge and will therefore be considered further in the context of growth ledges. The other two deviations from equilibrium configuration need not yield a ledge structure and therefore do not lie within the present scope. References 118 and 119 include the early history and experimental evidence on intruder dislocations.

b. *The growth ledge concept*

Growth ledges were first conceived by Gibbs[120] for the growth of crystals from the vapor or liquid phase and described by him in remarkable detail. In 1962, growth ledges were proposed for solid-solid phase transformations.[77] This proposal was modified in important ways in 1970.[35] In the latter form, the terraces of growth ledges were conceived as formed by well-matched, closely packed, conjugate habit planes. Such interfaces may be fully coherent or may contain misfit dislocations or other misfit-compensating defects. However, the barrier to their movement is not the climb of the misfit dislocations or other defects but rather the immobility of the coherent areas between such defects during diffusional growth.[35] As long as there is a stacking sequence change across the boundary[88,93] differently structured planes must be created one atom at a time. Particularly, because closely packed planes are involved, this means that the reconstruction process must be begun interstitially, in sites which may not become substitutional until restructuring

is at least fairly well advanced. However, this type of process requires the creation of large numbers of interstitialcies. In more or less close-packed metals, especially, this is energetically infeasible when only thermal activation is available to create interstitialcies. Hence an alternative means of displacing such interfaces is required. By analogy to Gibbs' mechanism, growth ledges were proposed, whose risers were initially suggested to have a disordered structure.[77] It was soon realized, however, on the basis of kinetic evidence that the risers of growth ledges also tend to be fully or partially coherent.[35] Thus, a ledge-on-ledge, and ultimately, a kink-on-ledge mechanism is required in order to secure sufficient localized atomic disorder to permit passage of atoms through the risers without having to surmount an unduly high-strain energy barrier to movement. The first detailed proposal for escaping such a barrier was made by Howe *et al.*[121] and is described in Section 3–g.

Hence, the terrace-ledge (riser)-kink model of Burton *et al.*[122] is reproduced for solid-to-solid transformations. However, whereas in crystal growth into a fluid (vapor or liquid) matrix, kinks are the most favorable sites for atomic attachment because the number of bonds binding an atom from the matrix phase to the crystal is largest, the reverse situation can obtain when one crystal grows diffusionally into another where there is a significant difference in stacking sequence across closely matching interphase boundaries. During solid-solid phase transformations, as just described, kinks on the risers of growth ledges provide sites at which atoms can be *least* tightly bound to their neighbors and hence can pass through the interphase boundary and attach to or leave the precipitate without undue kinetic difficulty.[121]

The earlier suggestion that dislocation climb kinetics can determine the migration rate of a partially coherent interphase boundary[77] is thus applicable only when the crystal structures of the matrix and precipitate are essentially identical and also have the same spatial orientation. This situation ensures that no significant change occurs in stacking sequence across the interphase boundary.[123]

c. *Rules for observing growth ledges with TEM*

During the initial study of θ' plate-thickening kinetics,[109] when these rules were not yet available, no growth ledges were seen with TEM. All deductions about them had to be made indirectly, from measurements of thickening kinetics. After these rules became available, however, a realistic integration of intergrowth ledge spacings and measured thickening kinetics could be made upon the basis of direct measurements of these spacings.[124] These rules, devised on the basis of diffraction contrast theory by Weatherly and co-workers,[125–128] include the following: (1) if the precipitate has an ordered structure, a low-order superlattice reflection should be employed; (2) reflections with a short extinction distance should be favored for viewing ledges to improve resolution of small interledge spacings; (3) matrix reflections should not be excited, if possible, in order that bright ledges on precipitates may stand out against a dark matrix background; and (4) a small deviation from the exact Bragg condition will further diminish background intensity. The latter conditions are important parts of the weak-beam, dark-field technique introduced shortly afterward by Cockayne.[87]

d. *Sources of growth ledges*

The oldest proposed source of growth ledges, whether for the formation of crystals from a vapor, liquid, or crystalline matrix, the so-called "two dimensional" nucleation, is due to Gibbs.[120] Since even a monolayer nucleus can have a significant interfacial energy at its edges, however, a more appropriate term for the initiation of a new crystal with the same crystal structure and composition as its substrate crystal is sympathetic nucleation.[129,130] Use of this term not only recognizes that nucleation is occurring at a type of site where supersaturation driving nucleation is likely to be less than obtains far away from a previously formed crystal of the same phase but also that conventional heterogeneous nucleation is involved with a same-phase crystal:matrix interface replacing, say, a grain boundary in the matrix phase as the nucleation catalyst, and with a critical nucleus shape derived from an appropriate modification of the Wulff construction.[131–134]

Sympathetic nucleation of ledges at plate broad faces has been reported by Weatherly[127] in 1971 on Mg_2Si plates and by Sankaran and Laird[124] in 1974 on Θ' $CuAl_2$ plates. Aikin and Plichta[135] have analyzed the sympathetic nucleation of growth ledges on both γ $AlAg_2$ and Θ' $CuAl_2$ plates from the standpoint of interaction of the strain energy associated with the substrate plate with that of a disc-shaped critical nucleus formed atop the substrate. They concluded that this interaction favors γ $AlAg_2$ but not Θ' $CuAl_2$ nucleation at plate edges because the strain energy in the former case is primarily shear while that in the latter is predominantly dilatational. This analysis correctly predicts the observations Elangovan and Plichta[136] made during coarsening of γ and Θ' plates.

The classic site for the development of growth ledges is that predicted by Burton *et al.*[137] in 1949, *i.e.*, the emergence of a screw dislocation, or of a dislocation with a component normal to the surface or interface of a product phase crystal, through this interface. This geometry produces an unhealable spiral ledge requiring no nucleation step. Burton *et al.*[122] treated such growth in great detail in their wide-ranging 1951 article. Cahn *et al.*[138] rewrote this treatment for individual growth spirals operative during solidification and Aaronson *et al.*[35] modified the Cahn *et al.* treatment to apply to solid-solid transformations. Sankaran and Laird[124] clearly identified the operation of the spiral dislocation mechanism during the thickening of Θ' plates. When many rotations of the spiral are involved, this mechanism becomes the well-known pole mechanism. Although the pole mechanism has long been of interest in the context of clearly martensitic transformations, direct evidence for its participation in diffusional transformations is still lacking.

As indicated in Section D–2–a, during fcc-to-hcp transformations (and also during the reverse transformation), crystallography requires that adjacent Shockley partials be stepped down by two $\{111\}_{fcc}$ planes in order to maintain the correct stacking sequence. These built-in steps serve as growth ledges if only one Burgers vector of partials is present over an appreciable area of the fcc:hcp interface.[25,26] However, as the other Burgers

vectors of Shockley partials begin to appear in order to make misfit compensation more complete, the contribution to growth by this mechanism diminishes.[26] During coarsening, Aikin and Plichta[135] have shown that sympathetic nucleation near the edges of broad faces then becomes the predominant mechanism for growth ledge creation.

To use Ashby's felicitous term, the fillets where grain boundary allotriomorphs join secondary Widmanstatten sideplates break down into "geometrically necessary" growth ledges. Kinsman *et al.*[44] obtained evidence for such arrays of ledges at the fillets of secondary Widmanstatten ferrite sideplates in Fe-C alloys with replication electron microscopy.

Weatherly[127] and later Merle and Martin[139] and Sankaran and Laird[124] have shown that the lines of impingement where Θ' plates parallel to different $\{100\}_\alpha$ planes meet are particularly fertile sources of growth ledges. Rajab and Doherty[140] have found a similar effect on ledge formation at γ $AlAg_2$ plates. In neither transformation, however, have the mechanistic details of growth ledge formation been elucidated.

It has been suggested that the volume change associated with the formation of proeutectoid ferrite may cause local rotations of the austenite:ferrite boundaries.[44,77] Relaxation of the resulting strains could introduce high ledges in these boundaries. The large size of ferrite sideplates and intragranular plates in steels heat-treated to produce a coarse austenite grain size, relative to that of plates in substitutional alloys (particularly those with fcc and hcp matrices), would make plastic as well as elastic relaxation of the ferrite plates more likely than in substitutional alloys. However, no test of this mechanism has as yet been made, probably because martensite formation destroys the austenite matrix, in which slip might otherwise be detected, during quenching to room temperature.

Sankaran and Laird[124] have found that impurity precipitates in contact with the broad faces of Θ' plates can catalyze the formation of growth ledges. Whereas this particular type of catalyst may not be frequently operative, this observation does suggest that carbides at austenite:ferrite boundaries, or Θ plates formed at α:Θ' boundaries, may also encourage formation of growth ledges on ferrite and Θ plates.

A recently discovered source of growth ledges is the wriggling of adjacent, parallel, misfit-compensating ledges into contact with each other during the proeutectoid α reaction in a hypoeutectoid Ti-Cr alloy.[92] These contacts produce superledges with kinks where the wriggles remain and with the combined height of the component single misfit-compensating ledges. Although it is not known how many misfit-compensating ledges can ultimately combine into a growth ledge, the fact that only the combined ledges are able to migrate long distances has been confirmed by means of visual hot-stage TEM observations.[92]

e. *Intergrowth ledge spacings*

Data on this essential factor in the evaluation of ledgewise growth kinetics remain surprisingly scanty more than 25 years after their importance was recognized. The principal cause for this situation is that in a given specimen, these spacings exhibit such wide scatter that extensive data taking is required in order to secure meaningful results, as Elangovan and Plichta[141] have demonstrated.

The first directly measured data, obtained with TEM using the Weatherly techniques for imaging ledges, on the interledge spacing as a function of isothermal reaction time at a given temperature were reported by Sankaran and Laird[124] on Θ' plates. These data passed through a minimum at a relatively early reaction time and then rose slowly thereafter. While parallel measurements were not made by Laird and Aaronson[25,26] on γ $AlAg_2$, combining their qualitative observations with the earlier ones of Hren and Thomas[112] on γ' led to the conclusion that a similar time dependence of the interledge spacing obtains for these plates.[142] Kinsman *et al.*[44] back-calculated the average intergrowth ledge spacing on the broad faces of ferrite sideplates as a function of the isothermal growth time from the measured growth kinetics of these faces in Fe-C alloys. They also measured the average spacing as a function of growth time with replication electron microscopy. While the back-calculated spacings exhibited great scatter because of the widely variable growth kinetics, rough agreement was nonetheless obtained between these spacings and those measured directly during the early seconds of growth. At later reaction times, however, the back-calculated spacings were about an order of magnitude larger than the measured spacings, even though the latter did not increase markedly with growth time. This result suggested that the ferrite growth ledges had migrated into low-energy boundary orientations where the supply of kinks on their risers was inadequate to approximate the disordered riser kinetics assumed by most available analyses.

f. *Growth ledge heights*

In their previously mentioned 1984 critical review of the literature on the bainite reaction, Christian and Edmonds[94] questioned the existence of superledges on the ground that the very large Burgers vector required and the accompanying shape deformation would be energetically infeasible to create and maintain. Introduction of suitable misfit-accomodating defects in the risers would remove the shape discontinuity at the interface and thus prevent formation of the IPS surface relief.[28] However, using the techniques for imaging growth ledges which he and his colleagues had recently developed (Section II–D–3–c), Weatherly[127] reported in 1971 that a careful study of growth ledge heights on Θ' and Mg_2Si plates revealed ledges from 1- to 10-nm high, but not smaller, and especially not of one lattice parameter or less in height. Kinsman *et al.*[44] reported superledges up to a micron high on the broad faces of ferrite plates, using the much lower resolution technique of thermionic electron emission microscopy (THEEM). These plates did yield various surface relief effects involving planar tilts. At the other end of the resolution scale, Howe *et al.*[143] used two-dimensional atomic resolution TEM to show that superledges several lattice parameters high occur frequently on γ or γ' $AlAg_2$ plates. Sun and Hiraga[144] have recently shown with HRTEM that the interface between decagonal and icosahedral quasicrystals in an Al-20 at. pct Pd-10 at. pct Mn melt-spun alloy is ledged. The terraces of the ledges are formed by the tenfold plane in the decagonal phase oriented parallel

to the fivefold plane of the icosahedral phase; the misfit between these planes is only 3.3 pct. The risers of the ledges are approximately 2 to 4 repeat periods in height and are both rough in appearance and ill-defined, indicating that they are probably heavily kinked. Sun and Hiraga suggested that transformation from the decagonal to the icosahedral phase proceeds by the diffusion-controlled migration of ledges. In more general terms under appropriate conditions, diffusional interactions between adjacent growth ledges have been shown by various mathematical techniques to promote their consolidation.[161,163,165,166] Hirth[166A] recently informed us that he and S.V. Kamat have found that strain energy interactions can also promote formation of high superledges.

Returning to the criticism of Christian and Edmonds, this probably arose because they continued to use the original suggestion by one of us that the risers of growth ledges are disordered.[77] In view of the difficulty initially experienced in understanding how good matching could be devised between the fcc and bcc lattices even at $\{111\}_{fcc}//\{110\}_{bcc}$ interfaces,[145,146,147] the suggestion that interfaces approximately orthogonal to this one were disordered seemed quite appropriate at the time it was made. However, by 1968 sufficient TEM information on interphase boundary structures and growth kinetics had been accumulated so that the original suggestion was retracted.[35] The risers were now considered to be partially coherent and thus also to migrate by the ledge mechanism. The ledges on the risers were termed "kinks," in parallel with the nomenclature Burton *et al.*[122] applied to crystal growth from the vapor phase. The concept of a kink on a riser appears to have originated with Frenkel[148] in 1945.

g. *Kinks on the risers of growth ledges*

The first experimental evidence for the presence of kinks on the risers of growth ledges and for the essential role that they play during ledgewise growth appeared during the hot stage TEM study of γ $AlAg_2$ plate thickening by Laird and Aaronson.[26] Although the measurements of the kinetics of ledge motion made with this technique were generally consistent with control by long-range volume interdiffusion in the α Al-Ag matrix as already noted, closer examination showed that when the ledge direction approached $\langle 110\rangle_{fcc}$ the growth rate dropped off sharply. This led to the suggestion that the risers of the Shockley partial ledges must also be coherent (just as are the regions of the terraces between Shockley partials) and hence must themselves be displaced by means of the ledge mechanism. Because the risers of these ledges are so short, these ledges are probably best described as kinks.

The studies of Howe and co-workers on the interfacial structure of γ and γ' $AlAg_2$ plates, conducted almost a generation later, were made with much higher resolution TEM instruments. Their two-dimensional atomic resolution TEM micrographs of superledges on these plates displayed irregularities on the risers, similar to those more recently observed by Sun and Hiraga[144] at decagonal:icosahedral risers. These are again strongly indicative of the presence of kinks.[143] In the recent work of Prabhu and Howe,[149,150] very large kinks in the risers of ledges (which should doubtless be termed "superkinks") were identified as arising particularly from deviations from low-energy riser orientations and from intersections of Shockley partials but also possibly from thermally activated double kinks. Laird and Aaronson[26] had deduced the first two of these three origins but were of course unable to confirm their deductions with the electron microscopes then available.

Of particular importance among the contributions of Howe and co-workers is the model of a kink in a bilayer ledge at an fcc hcp interface reproduced in Figure 3.[121,143] This model is seen to incorporate, in the heavily outlined area of this figure, an elongated vacancy-like defect large enough so that it should make possible the energetically easy passage of an atom through the riser. Experimental testing of this model, and of counterparts yet to be devised for other pairs of crystal structures seems not to be feasible with presently available electron microscopes. However, such studies are clearly of crucial importance to providing a firm base for the atomic mechanism of ledgewise diffusional growth.

h. *Flatness of terraces on the atomic scale*

A central aspect of the theory of diffusional growth by the ledge mechanism described in Section II–D–3–b is that the densely packed terraces of growth and other types of ledges should be essentially impenetrable to atomic transport across them. Thus, such interfaces should be essentially flat on an atomic scale unless the supersaturation at them is high enough so that "homogeneous" sympathetic nucleation of new growth ledges can occur at them. Howe *et al.*[143] found the terraces of growth ledges on γ $AlAg_2$ plates to be atomically flat as imaged in two dimensions, as did Westmacott and Dahmen[105] of the facets on silicon crystals as previously noted in Section II–D–3–b. These results are very important. However, it is still necessary to examine the flatness of these interfaces with a yet-to-be-invented TEM technique which would permit atomic resolution in three dimensions in order that we may be certain that this "foundation stone" of ledgewise diffusional growth theory is indeed correctly described.

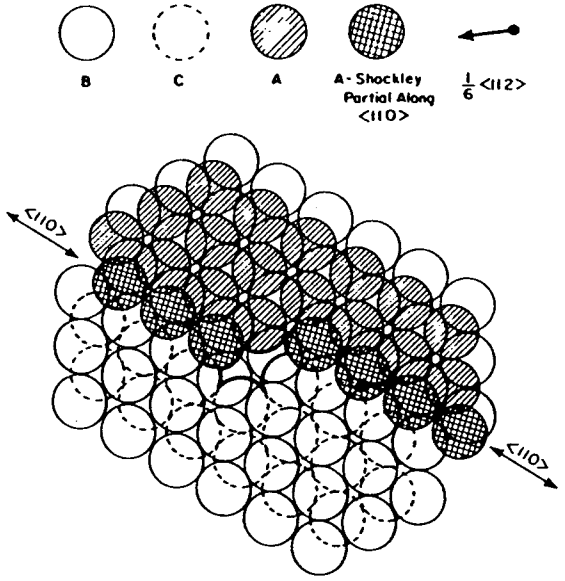


Fig. 3—A kink in the riser of a ledge two $\{111\}_{fcc}$ planes high, with the vacancy-like defect created at the kink outlined with heavy curves.[121]

4. *The fate of misfit-compensating defects overrun by growth ledges*

The third of three criticisms Christian and Edmonds[94] made of the mechanism of diffusional growth was as follows: particularly in interstitial alloys, misfit dislocations overrun by a growth ledge cannot climb up onto the terrace passing over them fast enough to make this mechanism feasible. Since no other plausible mechanism could be envisaged this consideration alone casts serious doubt upon the growth ledge mechanism itself. Except for the parenthetical expression, the response offered to this criticism may be summarized in the following manner. Although hot-stage TEM research will be required to determine the fate of overrun misfit dislocations in a particular transformation (see the article by Purdy and Zhang[151] in these proceedings to examine the results of just such an experiment), the evidence cited in References 26, 44, and 124 was proposed to be sufficient to demonstrate the existence of the growth ledge mechanism. The absence of evidence for a particularly low density of misfit dislocations or other misfit-compensating ledges in the immediate vicinity of the "upper" terrace of a growth ledge was taken to support the operation of mechanisms for resolving satisfactorily the fate of overrun misfit dislocations and maintaining the ambient density of misfit-compensating defects on the overrunning terrace.

While the question originally raised by Christian and Edmonds[94] was confined to the fate of overrun misfit dislocations, this question is clearly of sufficient importance to be generalized and thus to warrant inquiry as to what might happen to the other linear misfit-compensating defects described in Section II–D–2–b when they are overrun by growth ledges. The results of a recent attempt to do so on the basis of the nature and/or the relative energies of the various defects[5] will now be described together with the principal conclusions reached from consideration of the fate of overrun misfit dislocations.[98]

There should be no need to consider the transfer of structural dislocations from the overrun to the overrunning terrace because these defects seem unavoidable when their presence will lower the interfacial energy. Structural ledges do not incorporate an extra half plane in their terraces, though such can be periodically required in their risers. Because the ratio of the energy of a lattice misfit dislocation to that of a structural ledge is about 25 at bcc:fcc interfaces and about 80 at bcc:hcp interfaces, the ease of forming the structural ledges should be much greater.[152] It therefore seems likely that structural ledges will be able to form spontaneously on the overrunning terrace. Similarly, the lack of an extra half plane means that they should be overrun without leaving any trace of their presence in the overgrown region.

On the other hand, misfit-compensating ledges contain a partial dislocation,[92,93] and their total energy is not much less than that of a single misfit dislocation. Hence, it is not immediately clear what the fate of overrun linear defects of this type will be or how they will be reformed at or transmitted to the overrunning terrace.

In the case of misfit dislocations, Howe *et al.*[98] concluded that climb from the overrun to the overrunning terrace is the most probable mechanism in substitutional alloys. In interstitial alloys, however, expulsion of misfit dislocation loops into the matrix appears more likely on kinetic grounds. Purdy and Zhang[151] have tested the substitutional predictions in their contribution to the proceedings of this conference. The interstitial prediction remains to be examined with hot-stage TEM in an alloy system wherein thin-foil experimentation will not cause any substantial changes in mechanism relative to bulk specimens.

5. *Growth kinetics*

a. *Theoretical models*

By themselves, data on growth kinetics, even when taken at particular orientation of an interphase boundary, may mean little in the present context unless there is clear evidence for ledgewise growth. Even when discontinuities do appear in plots of half-length or half-thickness of plates *vs* isothermal reaction time, however, additional information can be extracted from the experimental data if there is a quantitative theoretical model with which the data can be compared. In the present subsection, the availability of such models is described, again in a historical context. In the next subsection, attention will be drawn to example tests of these models against experimental data.

Diffusional growth of planar disordered boundaries

Dube[153] wrote the first relationship for this situation as part of his Ph.D. thesis, based on the assumption of a diffusion coefficient in the matrix phase which is independent of composition. The following year, Zener[154] published derivations for the parabolic rate constant describing the growth kinetics of planar, spherical, and cylindrical interphase boundaries, all tacitly assumed to have a disordered structure not capable of affecting growth kinetics. Atkinson and co-workers[155,156] then incorporated a composition-dependent diffusivity in the analysis of planar boundary growth kinetics. These relationships have been repeatedly used to provide "baseline" plots of plate and allotriomorph half-thickness as a function of reaction time.

Diffusional lengthening kinetics of plates with disordered edges

Zener[80] wrote the first equation for these kinetics, albeit in incomplete form, in his famous 1946 article on austenite decomposition kinetics. As part of a major effort to understand quantitatively many aspects of the thermodynamics and kinetics of precipitate growth, Hillert[157] further developed and also modified Zener's relationship into what subsequently became known as the Zener–Hillert equation. Trivedi[158] then adapted the more exact Ivantsov[159] equation to include the effects of capillarity and uniform interfacial reaction control upon lengthening kinetics of parabolic cylinders.

Ledgewise growth kinetics of essentially planar boundaries

The first relationships for the growth kinetics of individual, diffusionally isolated ledges (assuming a composition-independent volume diffusivity in the matrix phase and uniform atomic attachment to the risers of the ledges) were written by Jones and Trivedi[160] in 1971. They subsequently extended this analysis to situations in which multiple ledges diffusionally interact.[161] Beginning in 1981, Atkinson[162,163] used a different

approach to both problems treated by his predecessors, securing results that differed significantly under conditions of low supersaturation and diffusional interactions among the ledges at a boundary. Having undertaken, with Atkinson, an analysis based upon the latter's approach to the problem of the growth kinetics of an interphase boundary containing a large number of growth ledges,[164] Enomoto decided to investigate ledgewise growth with a computerized finite difference approach rather than make a further effort to examine this very complicated problem analytically.[165,166] His results closely matched those of Atkinson. He also showed that growth kinetics faster than those allowed a planar disordered boundary can be achieved when a ledged boundary of finite length, at which the ledges are closely spaced and the diffusion distances of individual ledges are not too large, is studied during the early stages of growth. Further extending these considerations,[167] he also demonstrated that the growth kinetics of boundaries of effectively infinite length begin to match accurately those of planar disordered boundaries at short reduced growth times when the ratio of the intergrowth ledge distance to the growth ledge height is small, *e.g.*, 2, but do not do so until extremely long reduced growth times when this ratio is large, *e.g.*, 50.

Growth kinetics of a planar boundary by shear

Although a martensite plate can grow at near-sonic velocities,[21] growth rates many orders of magnitude slower have also been recorded, *e.g.*, References 168 through 170. These observations have inspired analyses of drag effects upon glissile interphase boundaries, which are taken as potentially applicable to the growth kinetics of plates by shear at reaction temperatures above M_d.

Owen and co-workers[169,171] drew attention in 1970 to the ability of forest dislocations interacting with a glissile interface to slow the motion of such an interface. In addition to the ambient dislocations already present, dislocations created by dilatational strains attending a martensitic transformation as well as shear strains not neutralized by formation of an adjacent martensite plate with an oppositely directed shear give rise to these dislocations. A 1985 article by Grujicic *et al.*[172] considered three additional sources of drag effects upon glissile boundaries. The Peierls force is overcome by the nucleation and growth of double kinks—a process that takes place rapidly except at low temperatures.[173] Phonon drag effects become important only at high temperatures and at growth velocities about half the velocity of sound or greater.[173,174] Finally, lattice invariant deformation dislocations that remain intact interact more strongly with obstacles than do interfacial dislocations performing this function, which become partials in the martensite and leave a trail of stacking faults or twins behind them in the transformation product.[172]

The limited ability of growth kinetics alone to define the atomic mechanism of transformation is shown by the finding of Schoen and Owen[169] that the growth kinetics of martensite in Fe-C-10 pct Ni alloys are less than those allowed by carbon diffusion control. Clearly, additional information, particularly on interphase boundary structure, must be available in order to assure a less ambiguous interpretation of data on growth kinetics.

b. *Experimental tests of the theoretical models*

Lengthening kinetics of plates

In 1962, Kaufman *et al.*[175] developed an elaborated version of the Zener–Hillert[80,157] equation for the lengthening rate of plates and compared its predictions with unpublished data obtained by Hillert[176] on the lengthening rate *vs* reaction temperature of ferrite and bainite plates in plain carbon steels. At most alloy compositions and temperatures, the measured kinetics were 5 to 10 times below those calculated. Kaufman *et al.* attributed the discrepancies to approximations in the Zener–Hillert equation and/or to an erroneous estimate of the interfacial energy of ferrite plate tips. During the same year, a comparison of the measured lengthening kinetics of proeutectoid α plates in a hypoeutectoid Ti-Cr alloy with those calculated using the elementary form of the Zener–Hillert equation and self-diffusivities rather than the more appropriate interdiffusivities in β Ti-Cr yielded a similar result: measured lengthening rates were about five-fold less than those calculated.[107] It is doubtful that use of interdiffusivities would have fully eliminated this discrepancy. Simonen *et al.*[108] measured lengthening rates of ferrite and bainite plates in high-purity Fe-C alloys and compared them with those calculated from the more sophisticated Trivedi[158] relationship. For the more plausible value of the edge interfacial energy used, full agreement between calculated and measured rates was obtained at 700 °C, the highest temperature studied, but the measured lengthening rate was about an order of magnitude smaller than that calculated at 450 °C, the lowest reaction temperature used.

Oblak and Hehemann[177] took a different view of the lengthening kinetics of ferrite and bainite plates in an influential symposium article published in 1967. Noting that upper and lower bainite plates tend to have lateral boundaries spaced along their length, they proposed that each boundary defines a component plate and that each of these plates grows by shear at a sufficiently high velocity to permit the ferrite thus formed to be supersaturated substantially with respect to carbon. The time that elapses between such shears was taken to be sufficient to produce the experimentally measured lengthening rates. Terming the product of each shear a "subunit," they noted that individual subunits have a lath morphology, and are approximately 1-μm thick, 5- to 10-μm wide and 10- to 50-μm long.[178] The limiting size of an individual subunit was proposed to result from accumulation of dilatational and shear strain energies.* Relaxation of these strain energies determines the time that elapses prior to the nucleation of the next subunit. The appearance of extensive sympathetic nucleation at and below the bay temperature of the TTT diagram in certain alloy steels, yielding heavily dislocated ferrite laths said to have been formed by shear, was emphasized. The inability of the ledgewise diffusional growth mechanism to explain the sudden increase in sympathetic nucleation kinetics was noted in the course of a debate with Kinsman and Aaronson on the bainite reaction.[179] (Reynolds

*Olson has privately questioned this reasoning on the ground that these strain energies are appropriately considered on a per unit volume basis and hence should be approximately independent of both plate size and reaction time.

et al.[179A] later did explain this phenomenon on the basis of diffusional growth by proposing that when the solute drag-like effect slows growth sufficiently to develop a critical level of supersaturation of austenite at austenite:ferrite boundaries with respect to further ferrite nucleation, sympathetic nucleation will begin abruptly to appear frequently as a consequence of the very strong dependence of nucleation kinetics upon the supersaturation driving nucleation.)

In one of their presentations during "the bainite debate,"[179] Kinsman and Aaronson showed plots of subunit length and of thickness *vs* reaction time, measured from motion pictures that Rouze and Grube[180,181] took during bainite sheaf growth in an Fe-0.66 wt pct C-3.32 wt pct Cr alloy reacted at 400 °C by means of THEEM. No evidence was found for subunit lengthening or thickening more rapid than that of the sheaf as a whole, whose lengthening kinetics were accepted as approximately those allowed by the diffusion of carbon in austenite. In his response, Hehemann[179] pointed out that frame-by-frame measurements of the Rouze–Grube films (rather than measurement of every fourth frame, as Kinsman and Aaronson found to be sufficient) revealed occasional growth rates at least one order of magnitude higher than average. He concluded that while such observations do not prove the operation of either high-velocity shear or ledgewise diffusional growth, they do indicate the need for further studies. Insofar as we are aware, no reports of such efforts have appeared.

In two substitutional alloy transformations, α_1 Ag-Cd[55] and α_1 Cu-Zn[182] plates were both deduced to lengthen more rapidly than the calculated rates by factors of 180 and up to 20 (again at low temperatures), respectively. In the former transformation, the result was ascribed to a shear mechanism of growth, while in the latter, the difference between calculated and measured rates was explained in terms of nonequilibrium precipitate compositions. However, a detailed critique of both comparisons indicates that neither is presently acceptable.[1] Similarly, Bhadeshia[183] and Ali and Bhadeshia[184] claimed that, upon re-evaluation of published data on the lengthening kinetics of ferrite plates on the basis of the Trivedi[158] equation for this process in carbon and in alloy steels, respectively, the measured lengthening rate became significantly higher than that calculated. However, this result was mainly the result of incorporating an arbitrary and unjustified value of the volume strain energy attending transformation in their analysis.[1]

Thus, the reliable comparisons that have so far been made indicate that the lengthening kinetics of plates are reduced by an interfacial reaction process. This process has been proposed to be the ledge mechanism.[108] Transmission electron microscopy evidence for ledges at the edges of plates has been secured for the proeutectoid α reaction in Ti-Cr[92] and Ti-Fe[113] alloys and less certainly for α_1 Cu-Zn plates.[185] Hot-stage TEM evidence has demonstrated directly that γ $AlAg_2$ plates lengthen in ledgewise fashion.[26] And very limited evidence has been reported for a possible ledge structure at the edges of proeutectoid α plates.[186] Transmission electron microscopy observations on these edges would be very welcome, particularly in a steel alloyed so that the adjacent austenite can be retained in thin foils at room temperature. Had these transformations proceeded by shear, no such interfacial structure evidence could have been secured since, as already noted, on this mechanism the edges of plates can only be a very thin sandwich (perhaps a few atomic planes in height, Olson suggests) between the two ledged interfaces whose terraces are the invariant plane of the transformation.[187,188]

Thickening kinetics of plates

Overall thickening kinetics of plates

Measurement of thickening kinetics represents a problem in quantitative metallography complicated by stereology and also by the thinness of precipitate plates during their early stages of development, thereby requiring exceptional accuracy of measurement. The first experimental evidence for ledgewise thickening of plates, in particular ferrite sideplates formed in an Fe-C alloy was reported in 1962.[77] This evidence consisted of only a single plot of plate half-thickness *vs* reaction time, showing a single marked discontinuity such as would be expected from the passage of a superledge through the fiducial line selected for measurement. These data were obtained from motion picture film taken of sideplates during growth at temperature by means of THEEM. The first detailed data set appeared 6 years later and consisted of plots of the maximum half-thickness of Θ' plates as a function of reaction time in an Al-4 pct Cu alloy.[109] All measurements were made at room temperature on isothermally reacted and quenched specimens. Transmission electron microscopy permitted accurate measurement of the thickness of these thin plates. Use of the microscope tilting facility in conjunction with the $\{100\}_\alpha$ habit plane of these plates readily permitted the true thickness to be measured, sans correction, by viewing the plates edge-on. Although this method of measurement could not detect the contributions of individual ledges, both the time law observed (linear instead of parabolic) and growth kinetics slower than allowed by volume diffusion control eliminated growth by uniform atomic attachment. The following year, application of hot-stage TEM to the thickening kinetics of γ $AlAg_2$ plates permitted easy detection of superledges.[26] Relatively prolonged periods during which no growth occurred, even though ample driving force was available, were frequently observed. Overall, growth kinetics were usually significantly less than but occasionally, during the early stages of growth, somewhat greater than those allowed by volume diffusion control with uniform atomic attachment to a planar disordered interphase boundary.[26] These data favor diffusional ledgewise growth unless much ingenuity is exercised in applying the various drag effects which can hamper the motion of a glissile interphase boundary.

Five years later, Sankaran and Laird[124] published decisive evidence for ledgewise control of the thickening of Θ' plates in Al-4 wt pct Cu. Measuring the average intergrowth ledge spacing on the broad faces of Θ' plates with the recently published techniques of Weatherly and co-workers[125–128] as well as the half-thickness *vs* time of these plates, they employed the Zener–Hillert equation and the Jones–Trivedi[160] analysis of ledgewise growth to back-calculate the interdiffusivity in the α matrix operative at the risers of growth ledges as a function

of reaction temperature. These diffusivities were in excellent agreement with those extrapolated from conventional interdiffusion measurements made in the α Al-Cu region and extrapolated down to the reaction temperatures used.

In the same year (1974), Kinsman *et al.*[44] published the results of their extensive THEEM measurements of ferrite sideplate thickening kinetics in three Fe-C alloys, conducted as a function of reaction temperature. Variations in the half-thickness *vs* time plots at a given temperature and alloy composition were as large as those encountered over the entire temperature-composition range used. Time laws varied from linear to parabolic to mixtures thereof to grossly irregular. Again, in the absence of growth ledges, no thickening at all took place for relatively long reaction time intervals despite the fact that adequate driving force for continued growth remained available. As Kinsman *et al.* noted, these kinetic behaviors are not easily reconciled with a shear mechanism of growth.

Growth kinetics of individual ledges

Almost necessarily, all of the data on this subject are derived from hot-stage forms of microscopy. Although the measurements of the kinetics of ledge motion on the broad faces of γ $AlAg_2$ plates in an Al-Ag alloy were generally consistent with control by long-range volume interdiffusion in the α Al-Ag matrix, closer examination showed that when the ledge direction approached $\langle 110 \rangle_{fcc}$, the growth rate dropped off sharply.[26] This led to the suggestion that the risers of the Shockley partial ledges must also be coherent (just as are the regions of the terraces between Shockley partials) and hence must themselves be displaced by means of the ledge mechanism.

Very high (approximately 1 μm) individual superledges on the broad faces of ferrite sideplates could sometimes be resolved with THEEM.[44] Their lengthening kinetics were approximately those allowed by the diffusivity of carbon in austenite. However, comparison of intergrowth ledge spacings back-calculated from the overall thickening kinetics of these side plates with those measured on quenched specimens with replication electron microscopy revealed good agreement only during the early stages of growth. In the later stages, the back-calculated spacings were an order of magnitude larger, indicating that also during this transformation the growth ledges had migrated extensively into energy cusp orientations wherein kink formation at their risers took place much less readily.

Thickening kinetics of grain boundary allotriomorphs

When the matrix has an fcc crystal structure and both matrix and precipitate have substitutional solid structures, the "collector plate" mechanism accelerates both thickening and lengthening of grain boundary allotriomorphs by means of grain boundary and interphase boundary diffusion when the reaction temperature is less than 0.85 to 0.90 T_m, where T_m is here the solidus temperature of the alloy.[189,190] As anticipated the collector plate mechanism is less effective when the matrix has a bcc structure but is nonetheless still operative.[191,192] Only when the matrix and product phases have interstitial structures is the collector plate mechanism largely suppressed.[156] Hence, consideration of allotriomorph thickening kinetics will be confined to ferrite allotriomorphs in Fe-C alloys, where thickening kinetics will more directly reflect the interfacial structure of the allotriomorph broad faces.

The stereological problem involved in obtaining the true thickness of allotriomorphs is more difficult than the parallel problem encountered in plate thickening studies, since the broad faces of allotriomorphs are roughly parallel to the more or less random and often changing average plane of the grain boundaries along which they preferentially grow. The solution found to this problem is to austenitize thin specimens until the grain boundaries can be shown, by orthogonal sectioning, to be nearly perpendicular to the intended plane of polish.[193,194,195] The thickness of the thickest allotriomorph (which is usually the longest as well) is then plotted as a function of the square root of the isothermal reaction time. The slope of this plot yields the parabolic rate constant, customarily designated as α. While counterpart plots taken from motion pictures of the growth of a single ferrite allotriomorph in a similar specimen observed with THEEM display less scatter about the correlation line[193] the values of α obtained from a single specimen can easily vary by a factor of 2 or even more from one allotriomorph to the next, thereby inhibiting efforts to compare the variation of α with temperature in different alloys.[196]

The α data obtained from specimens measured at room temperature thus represented (if the effects of tilting of the austenite grain boundaries away from perpendicularity to the specimen surface studied can be ignored) the smallest value of the ratio of the intergrowth ledge spacing to the growth ledge height at each reaction time. Even with this bias, however, the ratio of the measured α's to those calculated from the Atkinson *et al.*[156] analysis for a planar boundary (wherein the effect of carbon concentration upon the diffusivity of carbon in austenite is taken into account) was significantly less than unity. The combination of this result and the scatter in the THEEM data on α indicate that ferrite allotriomorphs also have partially coherent interphase boundaries and thicken by the ledge mechanism. Although examination of the interphase boundary structure of ferrite allotriomorphs has so far been prevented by transformation of the unreacted austenite to martensite during quenching to room temperature, nucleation theory indicates[117] that there can be little doubt that the observation of partial coherency on both the rationally and the irrationally oriented interfaces of proeutectoid α allotriomorphs in β Ti-Cr alloys[102] is of general importance.

6. *The invariant line concept and its role in the shear vs diffusion controversy*

The most important contribution toward resolving the shear *vs* diffusional growth controversy in a positive fashion has undoubtedly been the extraction of the invariant line strain concept from the PTMC and its application to accounting for the orientation relationships and the habit plane or direction of precipitate plates or needles in numerous alloy systems by Dahmen and his co-workers[197–201] since 1981 and by Luo and Weatherly[202,203] since 1987. The Invariant Line (IL) is

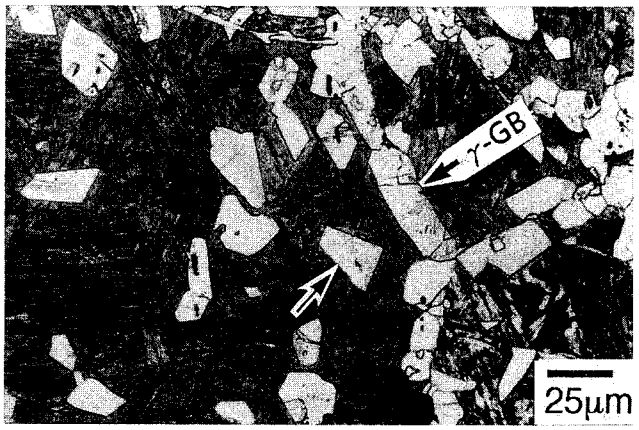


Fig. 4—Intragranular ferrite idiomorphs nucleated at VN or NbN precipitated on MnS inclusions in 0.23 wt pct C, 1 52 Mn steel quenched from 640 °C after cooling from 1250 °C.[207]

one of zero misfit. Because this is usually irrational, slight deviations of the lattice orientation relationship from rationality and then the breakdown of the IL into coherent terraces and coherent risers are required to achieve the zero misfit and, thus, the zero strain energy condition. Inasmuch as the macroscopic broad faces of martensite plates must also contain an IL,[204] a link has thus been established between the PTMC and the transformation crystallography of precipitate plates and needles formed entirely by diffusional processes. Among example transformations previously discussed (Section II–B–2), differences in composition and/or in long-range order between these precipitates and their matrix phase necessarily establish the transformation mechanism as diffusional, at least in the case of substitutional alloys. Whereas martensite plates have long been recognized as containing an IL in their broad faces in order to minimize their transformation shear strain energy, the suggestion was recently made[1] that the rate of formation and/or lateral movement of growth ledges should often also be a minimum at the interface containing the IL of the transformation. This suggestion has yet to be tested by experiment.

An IL alone cannot, of course, define a habit plane. A variety of other criteria has been proposed to establish a second, intersecting line.[190] Because, however, the habit plane does not appear to be particularly sensitive to which criterion is used for this purpose, this lacuna seems relatively unimportant, albeit for reason(s) not yet understood.

7. *Product phase morphology*

As previously stated, Christian[182,183] has emphasized that minimization of transformation shear strain energy requires that a martensite plate consist solely of two IPS interfaces. Each interface is composed, at the atomic scale, of terrace planes containing LID dislocations (if required) and risers serving as transformation dislocations. While the broad faces of a precipitate plate or lath formed diffusionally can consist predominantly of IL interfaces, the edges of these plates have a finite width and their own interfacial structure. Further, the products of ledgewise diffusional growth have considerably greater flexibility of morphology than narrow plate and lath edges imply.[205] As long as there are two or more crystallographically distinct boundary orientations at which the ratio of the intergrowth ledge spacing to the growth ledge height are comparable, a wide range of morphologies can be developed. These include grain boundary allotriomorphs and intragranular idiomorphs.[206] An example of the latter, kindly furnished by Mr. Fusao Ishikawa of Nippon Steel Corporation,[207] is shown in Figure 4. The planar facets easily visible on these ferrite idiomorphs formed in a low-carbon industrial steel can probably be safely regarded as partially coherent and growth ledged, even though an interfacial structure study of them has yet to be reported.

III. A COMPARISON OF SHEAR AND DIFFUSIONAL MECHANISMS FOR FORMATION OF PLATES AT TEMPERATURES ABOVE M_d

A. *Comparison of Ledgewise Growth by Shear and Diffusion*

As previously stated,* growth by shear evidently proceeds by dislocation glide whereas diffusional growth may be described as diffusional random walk biased toward the product phase by the free energy change driving growth. Both types of growth now appear usually, and perhaps always, to take place by means of the ledge mechanism. As schematically illustrated in Figure 5, however, there is a major difference in the way in which these two types of growth use this mechanism. In the general case (from the standpoint of growth by shear) of non-fcc/hcp transformations, both the risers and the terraces are mobile. The risers are formed by transformation dislocations, and the terraces incorporate LID dislocations. Both sets must glide in order to complete the transformation process. The essence of diffusional

*In order to provide a coherent presentation of section, it will be necessary to repeat some statements made in Section II on the history of the shear *vs* diffusion controversy.

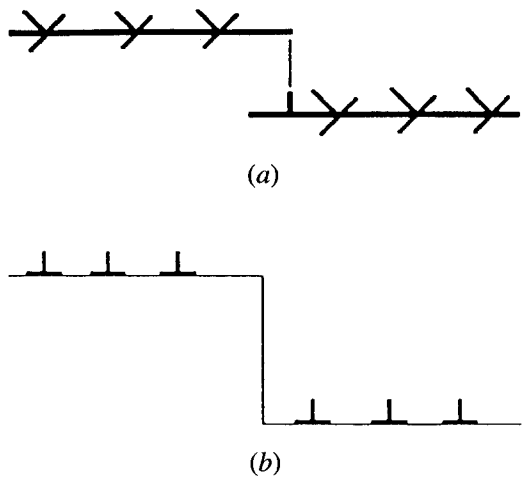


Fig. 5—Schematic sketches of single ledges capable of growth (*a*) by shear and (*b*) by diffusion, showing their respective dislocation structures. In part (a), the dislocation structure applies to non-fcc/hcp transformations.

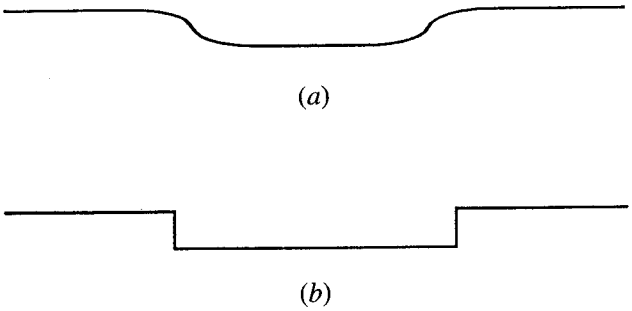


Fig. 6—(*a*) Diffuse kinks, characteristic of shear transformations, and (*b*) sharp kinks expected in diffusional transformations.

growth by the ledge mechanism, on the other hand, is that the terraces must remain entirely immobile, whatever the details of their interfacial structure. Boundary migration is now restricted not just to the risers of growth ledges but still further to atomic-scale kinks on these risers. For generally similar reasons, glide of both transformation and LID dislocations during shear is probably also restricted to atomic-scale kinks in these dislocations.[1,5]

Howe *et al.*[121] have modeled a bilayer high riser of a growth ledge for the fcc/hcp transformation in aluminum-rich Al-Ag alloys, as illustrated in Figure 3. The heavily outlined vacancy-like defect in the kink should be large enough, as already remarked, to permit passage of atoms through it on their way from the matrix to the precipitate (and *vice versa*) without the need to overcome an activation barrier larger than that required for the usual vacancy-atom interchanges. Although similar models have yet to be constructed for fcc/bcc, bcc/hcp, and other structure-change transformations proceeding diffusionally, it seems likely at this time that the principle evoked, namely, the development of a well-defined opening in the otherwise closely matched and densely packed conjugate habit planes comprising the riser, is likely to carry over into these crystallographically more complex phase transformations.

During growth by shear, which can take place at high velocities at low temperatures,[208] on the other hand, the kinks are typically diffuse and remain coherent, thus permitting their formation with minimum thermal activation, *e.g.*, as described by Hirth and Lothe[209] for kink formation at low temperatures in dislocations within a single phase. Figure 6 quite schematically illustrates the difference between sharply defined kinks appropriate to diffusional transformations and the diffuse kinks appropriate for growth by shear.

At the still larger scale of interphase boundary structure, fcc/hcp transformations of both the shear[210–213] and the diffusional[26,112] types take place by means of arrays of Shockley partials with the same Burgers vector. As discussed in Section II–D–2–b, maintenance of the correct stacking sequence (without stacking fault formation) requires that adjacent partials in this situation be stepped down with respect to one another.[26,35] This system of built-in ledges provides the transformation mechanism until such time as the proportion of the interfacial area containing only one array of ledges becomes small as a consequence of the formation of arrays with one or both of the other two equivalent variants of the Burgers vector. Then sympathetic nucleation of (apparently multilayer) ledges near the edges of plates provides the remaining growth ledges needed.[135] An essentially parallel situation should obtain during the built-in ledge stage of growth by shear. A replacement suited to shear transformations for the second or sympathetic nucleation stage of growth ledge formation, on the other hand, is not apparent.

During diffusional fcc/bcc transformations, growth of the broad faces of bcc plates occurs at interfaces with the type of structure illustrated in Figure 7.[92] Two types of linear misfit-compensating defects are present at this interface: the misfit dislocation and the structural ledge. As discussed in Section II–D–2–b, the Burgers vector of both defects is parallel to the terrace plane of the structural ledges, and the structural ledges themselves appear to have negligible mobility at ranges greater than

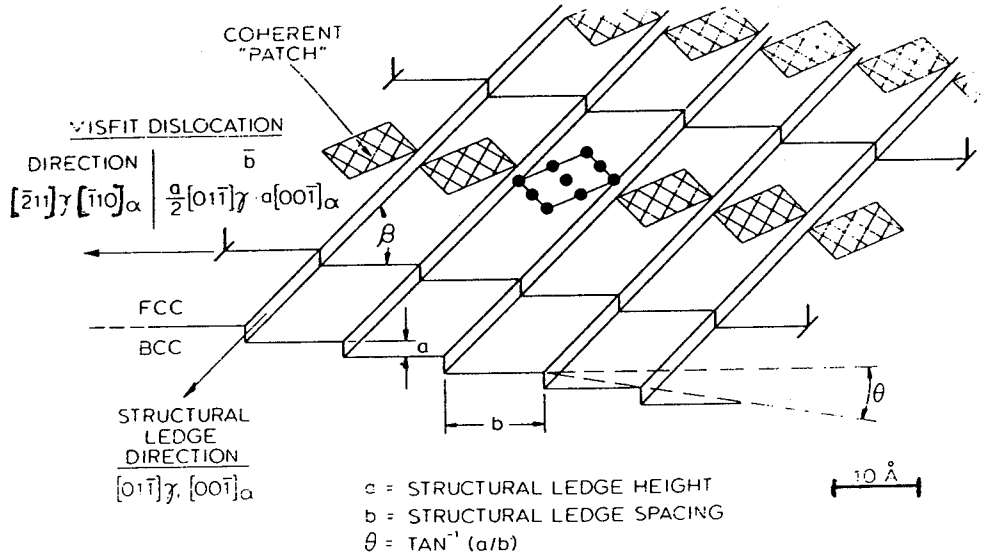


Fig. 7—Schematic of partially coherent fcc:bcc interface, showing sessilely oriented misfit dislocations and structural ledges.[88]

the spacing between adjacent parallel structural ledges. Hence, this interfacial structure is sessile with respect to growth by shear. However, Sandvik and Wayman[214] have deduced that the misfit dislocations at this type of interface are in pure screw orientation at the broad faces of martensite laths in an Fe-20 pct Ni-5 pct Mn alloy. Hence, the glide of these dislocations provides the needed LID. *Some other source needs to be discerned, however, for transformation dislocations at fcc/bcc interfaces.* During diffusional growth, superledges serving as growth ledges during diffusional transformations (only) are readily generated at the broad faces of ferrite plates and especially at the junctions between secondary sideplates and the grain boundary allotriomorphs from which they develop (Section II–D–3–d). Growth ledges operative during diffusional transformations in general must be extensively kinked in order to provide viable atomic attachment sites. They should usually be higher (or at least have a different height) than any intrinsic, *e.g.*, structural or misfit-compensating ledge in order to avoid partial mutual cancellation of strain fields such as would tend to trap the growth ledges.[5] Growth ledges appropriate to diffusional growth are unlikely to serve this function during growth by shear, mainly because of the structure of their kinks but, normally, also because of their height.

Figure 8 illustrates the interfacial structure deduced from TEM and computer modeling to be operative at the broad faces of hcp plates formed in a bcc matrix.[92] Two different kinds of linear misfit-compensating defects are present at this boundary: structural ledges and misfit-compensating ledges. Figures 9(a) and (b) are atomic models sectioned normal to the interphase boundary of these defects at bcc:hcp interfaces. Whereas there is no extra half plane in the terraces of structural ledges, a partial dislocation is present in misfit-compensating ledges. Otherwise, the misfit-compensating ledges also exhibit the coherency "pattern advance"[215] characteristic of structural ledges.[82] While the latter ledges can wriggle sideways sufficiently to merge with their neighbors and thereby form growth ledges (Section II–D–3–d), they do not appear to have any of the long-range mobility (relative to the spacing between adjacent misfit-compensating ledges) characteristic of growth ledges.

Ko and Cottrell[16] have proposed that in an interstitial alloy, the substitutional atoms can effect the crystal structure transformation by shear whereas the interstitial atoms can produce diffusionally the necessary composition change. However, this mechanism would require that, even if only one interstitial atom were to be present at each of some proportion of the kinks on transformation and LID dislocations, all such atoms would have to jump across (or away from) the interface simultaneously in order to allow that interface to move without a possibly sharp rise in the free energy of the product phase. Even if this condition had to be met only along the moving perimeter of a single transformation dislocation, it is statistically grossly improbable that such simultaneous jumping would happen if more than a very few interstitial atoms were to be involved.

B. *Comparison of Nucleation in Shear and Diffusional Transformations*

Critical nuclei develop during diffusional transformation through atom-by-atom additions during a purely random walk process.[216] The work, which must be accomplished by statistical fluctuations to form a critical nucleus, may not exceed the approximately 60 kT upper limit[217] for detectable nucleation kinetics. With the possible exception of the low strain energy fcc/hcp transformations,[218] however, the transformation shear strain energy subtraction from the volume free energy change driving nucleation is too large to avoid greatly exceeding 60 kT during attempts to form a martensite nucleus by statistical fluctuations.[219] Hence, some type of preformed nucleus is needed to exceed the critical size at the transformation temperature *ab initio*. At the present time, the Olson–Cohen[220] models for such nuclei seem the most appropriate available. For the simple and apparently noncontroversial fcc/hcp case, Figure 10[220] illustrates their model. A sub-boundary several dislocations high is seen to have decomposed into pairs of Shockley partials, separated by regions of hcp stacking fault tantamount to an hcp nucleus. The height of this nucleus is large enough so that sufficient volume free energy change is available, even in the absence of a composition change, to provide both the interfacial energy and the volume strain and dilatational strain energies needed. Note, in particular, that the Shockley partials are all in glissile orientation even in the as-formed state of this supracritical nucleus. Attempts to convert a sessile into a glissile interphase boundary structure at fcc:bcc interfaces (as suggested by Bhadeshia[221]) have been shown by detailed analyses to be energetically or kinetically futile.[222] Olson and Cohen[223] have evolved a more elaborate model of supracritical nuclei for fcc/bcc shear transformations which may also be appropriate for the purpose.

In the case of nuclei formed during diffusional transformations, the transformation strain energy can be much less, thereby permitting critical nuclei to form with detectable kinetics at moderate undercoolings. However, this can happen only if low energy interfaces enclose more or less completely the critical nuclei.[117] These interfaces may be fully coherent, or may include structural ledges and perhaps also misfit-compensating ledges[5] but not misfit dislocations.[224–227] When the nuclei have progressed sufficiently far into growth so that misfit dislocations could be accommodated with a decrease in the total free energy of the nuclei, the shear strains operative at coherent interphase boundaries will tend to cause the formation of misfit compensating dislocations (and possibly other defects) whose Burgers vector is parallel to the interfaces, *i.e.*, in sessile rather than in glissile orientation. However, the erratic nature of misfit dislocation acquisition can readily cause the first dislocations acquired to appear in glissile orientation.[35] As long as the nuclei continue to differ in composition from the surrounding matrix appreciably beyond their critical size, interdiffusion between the new crystal and the matrix is likely to continue unless growth by shear can develop rapidly enough to prevent further diffusional transformation.

C. *Comparison of Morphologies Producible during Shear and Diffusional Transformations*

As Christian and co-workers[11,13,187,188] have shown, the requirement that the apparent martensite habit plane

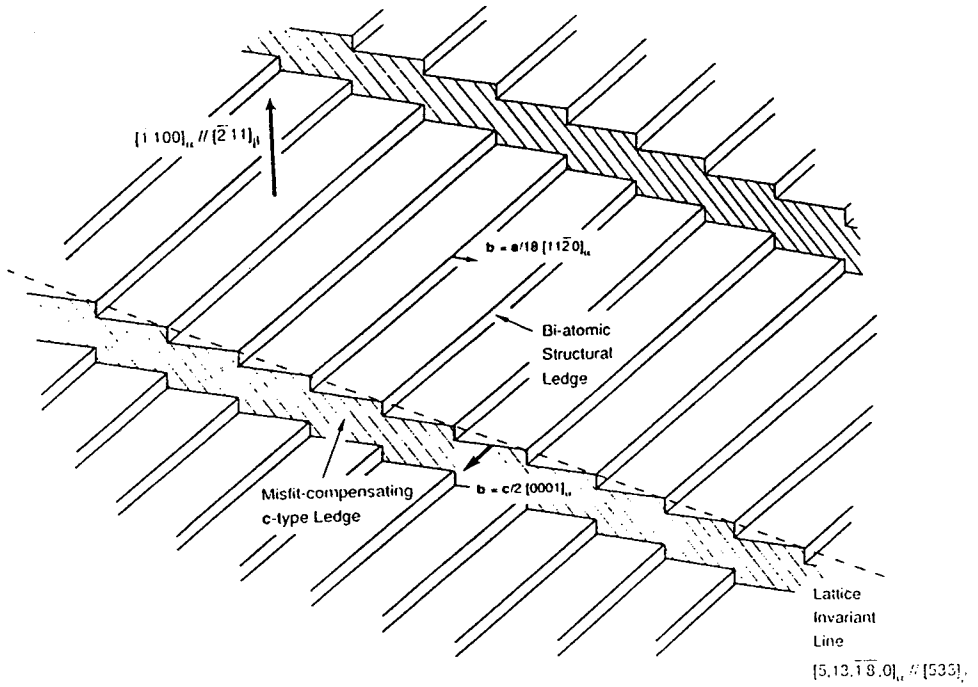


Fig. 8—Schematic of partially coherent bcc:hcp interface, showing structural ledges and misfit-compensating ledges.[92]

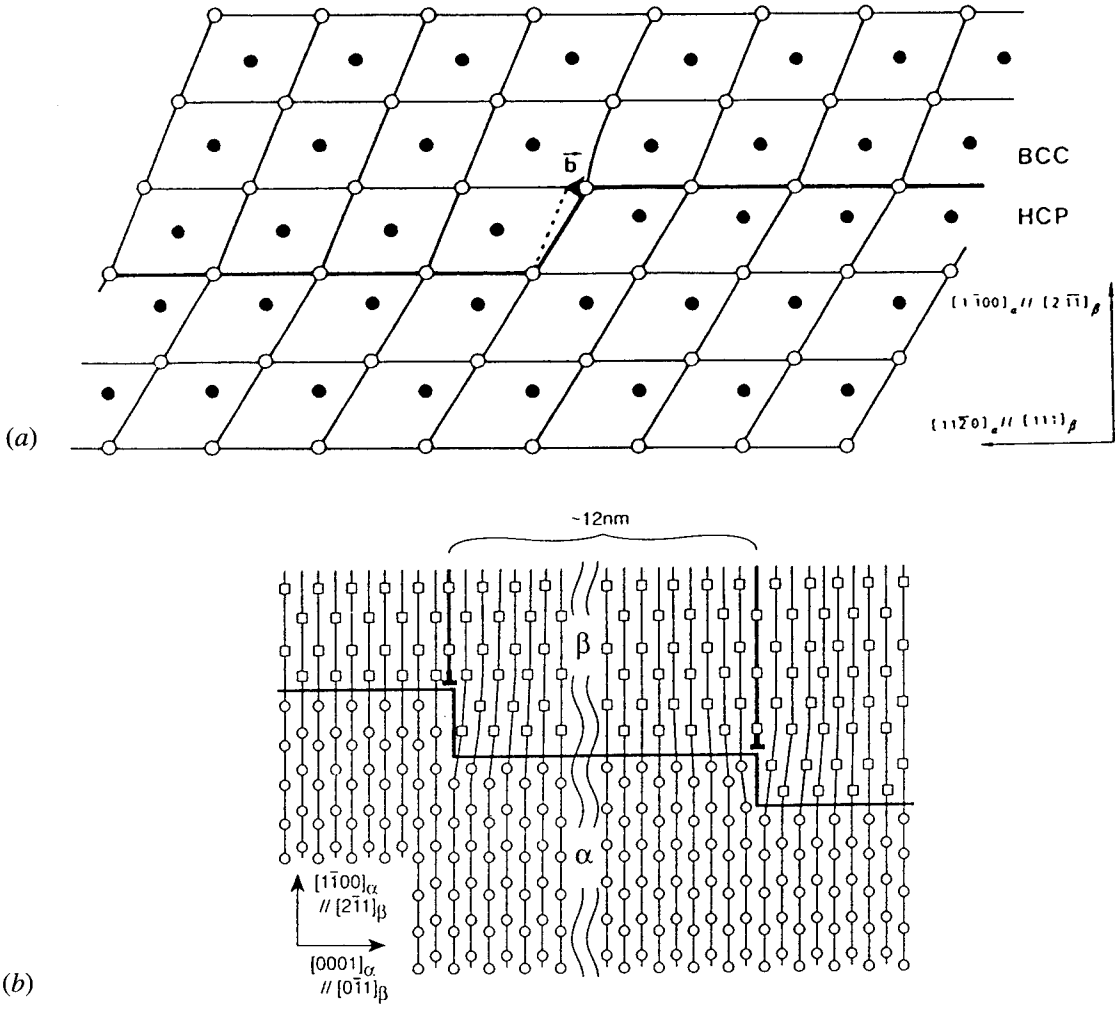


Fig. 9—(*a*) Atomic model of a structural ledge at a bcc:hcp interface.[93] (*b*) Atomic model of a misfit-compensating ledge at a bcc:hcp interface (T. Furuhara: private communication, 1991).

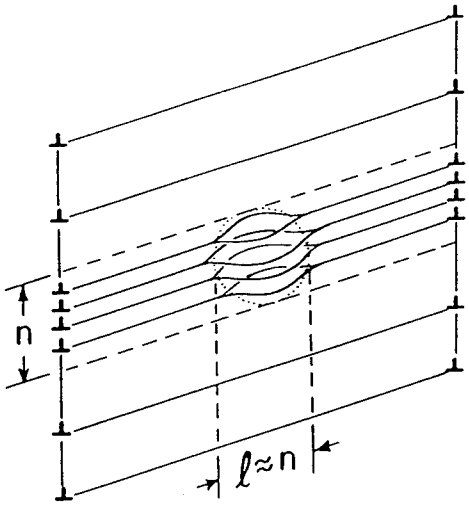


Fig. 10—Olson–Cohen model of an hcp nucleus in an fcc matrix.[207]

be undistorted and unrotated demands that both sides of a martensite plate be invariant planes and that this be the only (average) boundary orientation present. Hence, the morphology of martensite is restricted to that of plates. The findings of Dahmen and co-workers[197–201] and of Luo and Weatherly[202,203] that the broad faces of a precipitate plate or lath and the cylindrical or long planar faces of a precipitate rod or needle contain the IL of an IPS, however, do not similarly restrict the morphology of the products of diffusional transformations. It has been postulated[1] that such interfaces occupy a prominent facet or boundary orientation because this is the interface at which the kinetics of formation and/or lateral movement of growth ledges pass through a well-defined minimum. As long as the ratio of the intergrowth ledge spacing (λ) to the growth ledge height (h) at one or more different boundary orientations is not too much smaller than that at the IL interface, however, the other boundary orientation will also be represented in the "kinetic" morphology of the precipitate. Thus, the edge of a diffusionally formed precipitate plate or lath is not vanishingly thin, as in the case of a martensite plate but has instead a perceptible width as has been illustrated for a proeutectoid α lath in a Ti-5.2 wt pct Fe alloy.[113]

If several distinct boundary orientations have similar values of λ/h, then a faceted, roughly equiaxed morphology, known as an idiomorph,[206] will develop. Examples of this morphology developed during the proeutectoid ferrite reaction in a low-carbon commercial steel are shown in Figure 4.[207] Luo and Weatherly[202,203] have studied the interfacial structure of chromium-rich laths in a Ni-Cr alloy. They found different interfacial structures at the various boundary orientations where interfaces prominent in the lath morphology developed. It would now be of great interest to measure and compare λ/h of growth ledges at these interfaces.

D. *The PTMC as a Tool for Distinguishing between Shear and Diffusional Transformation Mechanisms*

The "formal" requirements of the PTMC that must be fulfilled by a shear transformation include the following: an IPS surface relief effect, a lattice correspondence, an irrational orientation relationship and an irrational habit plane (for non-fcc/hcp transformations), and a LID (also in non-fcc/hcp transformations), all of which must be quantitatively consistent with all other formal requirements.[12,15] The "derivative" requirements of the PTMC include an interphase boundary structure capable of supporting growth by shear and, at least in substitutional alloys, the absence of both a composition change and a change in long-range order during growth. In Section II–C, several well-documented examples were given in which the formal requirements are accurately fulfilled but one or more of the derivative requirements are not. Among substitutional alloys, which presently pose the strictest test of a shear or martensitic transformation mechanism, these include Cu-Au II (drastic change in long-range order), α_1 Cu-Zn (composition change), and γ $AlAg_2$ (composition change). Among interstitial alloys, proeutectoid ferrite plates and the ferritic plate component of microstructurally defined bainite often fail to obey both the formal requirements and the interfacial structure derivative requirement. They also do not obey the no composition change requirement. We maintain that the latter requirement is also applicable, because shear and diffusion are held to be incompatible on the ground that the former produces atomic transport by directed glide to pre-destined sites (within the radius of a shuffle) in the product phase whereas the latter does so by means of biased random walk, to which predestination is inapplicable—unless the bias is truly extreme. However, the formation of β $VH_{0.45}$ plates was noted to fulfill all formal and the applicable derivative requirements because only the interfacial structure requirement is applicable if one *does* accept simultaneous shear and diffusion. This is the only non-martensitic, plate-forming, nucleated transformation found to date in which so many PTMC requirements are fulfilled—and even this transformation is eliminated if the Ko–Cottrell shear-cum-diffusion mechanism is not accepted. Hence, the conclusion is reached that obedience to the formal requirements of the PTMC is a necessary but not a sufficient condition for identification of the atomic transformation mechanism as one of shear. However, particularly if simultaneous shear of substitutional atoms and diffusion of interstitial atoms is rejected, fulfillment of *all* of the derivative requirements as well as self-consistent satisfaction of all of the formal requirements does appear at this time to constitute a stringent test of the atomic mechanism for the formation of plate-shaped transformation products above M_d.

IV. SUMMARY

For some time, the principal source of the now more than 60-year-old controversy as to the roles of shear and diffusion in the formation of plate-shaped transformation products at reaction temperatures above the martensite range appears to be the failure of the advocates of the shear or martensitic mechanisms to heed fully all of the "formal" and the "derivative" requirements that must be met before plates can be validated as the product of a shear transformation. These requirements are again spelled out by Wayman[12] in the preceding article in these conference proceedings, and have been described previously

in ample detail by Clark and Wayman[15] and in the reviews by Christian and co-workers.[13,14,24,187,188] A secondary but nonetheless persistent source of argumentation has been the use of terminology, particularly when referring to events and processes evocative of the shear mechanism, by advocates of the diffusional ledge mechanism which those preferring shear find misleading or simply inaccurate. An example of such an "event" is the surface relief effect. An example of a "process" in this context is the use of the term shear to describe both an atomic process and the geometric relationship between matrix and product phase lattices.

Perhaps one reason why failures to obey exactly the formal requirements of the PTMC[18,19] by transformation products formed above the martensite range, and particularly above the M_d temperature, have not been taken too seriously by advocates of shear is that significant failures have been cataloged for transformation products that are unquestionably martensite. Wayman[228] has straightforwardly described these for ferrous martensites, but has emphasized that they tend to be associated with volume changes during transformation in excess of about 2.5 pct. Such changes presumably lead to additional plastic deformation for which the PTMC was not designed to account. A more disturbing trend, however, involves transformations in which the formal requirements of the PTMC are accurately obeyed but in which one or more derivative requirements are not. Good examples of these in substitutional alloys are the formation of AuCu II,[23,29] α_1 Cu-Zn,[38] and γ $AlAg_2$[26] plates. AuCu II plate formation is accompanied by the replacement of a disordered matrix with a long-range ordered product phase; the formation of plates in the other two reactions involves an evidently *ab initio* change in composition with respect to the matrix phase. Both violations of important derivative requirements of the PTMC are now clear-cut, though not until recently was the composition change accompanying α_1 Cu-Zn plate formation during the early stages of growth identified with full certainty.[38] Perhaps a broader acceptance of all three of these example transformations as fully diffusional processes by many leading members of the concerned community will date from the time of this conference.

The atomic mechanisms of the ferrite and bainite reactions in steel have long posed a special set of problems. The rapid reaction rates permitted, even in the absence of shear, by the high diffusivities of carbon in austenite, coupled with the accelerated kinetics provided by the point effect of diffusion at plate edges gave early investigators (Section II–A) the pronounced impression that high-velocity shear is the atomic mechanism of growth of bainite plates. In a more modern context, however, the Ko–Cottrell[16,17] combination of a shear transformation of the fcc iron lattice to the bcc iron lattice whose kinetics are controlled by the diffusion of carbon atoms through the interstices of austenite appeared to merge "the best of both worlds" when these investigators also found, by means of hot-stage optical microscopy, that the bainite plates they studied really did grow slowly. The influence of this 40-year-old proposal was clearly evident in many of the presentations made on the bainite reaction in steel during the conference. Hence, when Rigsbee and Aaronson[81] found by a combination of TEM and computer modeling that the structure of the broad faces of ferrite sideplates is sessile, and hence could not support a shear transformation mechanism, their results were (and continue to be) greeted by a mixture of skepticism and nonreferencing. Structural ledges, an integral component of this interfacial structure, though deduced by Hall *et al.*[82] 20 years ago, have only recently begun to gain acceptance. And the earlier work of Kinsman *et al.*[44] demonstrating complex surface relief effects in association with single ferrite plates and clear evidence for the operation of a ledge mechanism (particularly through superledges) during the thickening of sideplates, presaged the reception giver Rigsbee and Aaronson by making negligible impact upon advocates of a shear mechanism of transformation.

There can be little doubt but that the Invariant Line (IL) concept, extracted from the PTMC by Dahmen and co-workers[103,197–201] and repeatedly used to show that a wide range of precipitate plates contain an IL in their broad faces, has been the principal factor leading to a change of direction in the thinking of some of the participants in the shear *vs* diffusion controversy, as seemed particularly evident to us during this conference. Instead of "accentuating the negative," as the senior author has admittedly done throughout a career's participation in this controversy, the IL concept provides a still not quite complete but nonetheless "positive" reason why the formal components of the PTMC should be as well satisfied by diffusionally formed as by martensitically produced plates even though their atomic mechanisms are as different as physical metallurgy allows.

One of the subtler benefits derived from this conference appears to us to be the increasing acceptance of the "derivative" requirements of the PTMC as a tool for distinguishing between the two atomic mechanisms from observations made at much lower levels of resolution. Of course, Wayman said this and more about the PTMC's requirements for qualification as the product of a shear transformation 25 years ago,[15] and he has further and yet more clearly made this point in his introductory article at this conference.[12] Hopefully, this lesson will become more widely utilized through the publication of these conference proceedings.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Professor J.W. Christian, Oxford University, for his detailed comments on an earlier version of this article, for providing a copy of both the original and final versions of his conference article, and, especially for an extensive and stimulating interchange of FAXes about the principal issues being debated. We are also grateful to Professor J.P. Hirth, Washington State University, for his comments on the manuscript, for extensive discussions by mail and telephone about this controversy, and for providing a preprint of his article with S.V. Kamat on strain energy interactions between adjacent ledges. We thank Professor G.B. Olson, Northwestern University, for a very useful review of this article and for a personal discussion of many questions central to the debate. Mr. Fusao Ishikawa, Nippon Steel Corp., is thanked for an original print of Figure 4.

REFERENCES

1. H.I. Aaronson, T. Furuhara, J.M. Rigsbee, W.T. Reynolds, Jr., and J.M. Howe: *Metall. Trans. A*, 1990, vol. 21A, pp. 2369-409.
2. H.I. Aaronson, W.T. Reynolds, Jr., G.J. Shiflet, and G. Spanos: *Metall. Trans. A*, 1990, vol. 21A, pp. 1343-80.
3. H.I. Aaronson, T. Furuhara, J.M. Howe, W.T. Reynolds, Jr., and G. Spanos: *Interfaces, Structure and Properties*, Oxford and IBH Publishing Co. Pvt. Ltd., New Delhi, 1992, p. 169.
4. W.T. Reynolds, Jr., H.I. Aaronson, and G. Spanos: *Mater. Trans. J. Inst. Mater.*, 1991, vol. 32, p. 737.
5. H.I. Aaronson: *Metall. Trans. A*, 1993, vol. 24A, pp. 241-76.
6. H.I. Aaronson, G. Spanos, and W.T. Reynolds, Jr.: *Metal. ABM*, in press.
7. J.M. Robertson: *J. Iron Steel Inst.*, 1929, vol. 119, p. 391.
8. E.S. Davenport and E.C. Bain: *Trans. AIME*, 1930, vol. 90, p. 117.
9. S.L. Hoyt: *Trans. AIME*, 1930, vol. 90, p. 149.
10. R.F. Mehl and C.S. Barrett: *Trans. AIME*, 1931, vol. 93, p. 78.

10A. J. Hengstenberg and G. Wassermann: *Ztsch. F. Metalkunde*, 1931, vol. 23, p. 114.

10B. M. Straumanis and J. Weerts: *Ztsch f. Physik* 1932, vol. 78, p. 1.

10C. G. Kurdiumow: *Trans. AIME*, 1933, vol. 105, p. 253.

11. B.A. Bilby and J.W. Christian: *J. Iron Steel Inst.*, 1961, vol. 197, p. 122.
12. C.M. Wayman: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1787-95.
13. B.A. Bilby and J.W. Christian: *The Mechanism of Phase Transformations in Metals*, Institute of Metals, London, 1956, p. 121.
14. J.W. Christian: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 371.
15. H.M. Clark and C.M. Wayman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 59.
16. T. Ko and S.A. Cottrell: *J. Iron Steel Inst.*, 1952, vol. 172, p. 307.
17. T. Ko: *J. Iron Steel Inst.*, 1953, vol. 175, p. 16.
18. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. AIME*, 1953, vol. 194, p. 1503.
19. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 129, 138, and 224.
20. D. Hull: *Bull. Inst. Met.*, 1954, vol. 2, p. 134.
21. R.F. Bunshah and R.F. Mehl: *Trans. AIME*, 1953, vol. 197, p. 1251.
22. F.C. Frank: *Acta Metall.*, 1953, vol. 1, p. 15.
23. R. Smith and J.S. Bowles: *Acta Metall.*, 1960, vol. 8, p. 405.
24. J.W. Christian: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, p. 129.
25. C. Laird and H.I. Aaronson: *Acta Metall.*, 1967, vol. 15, p. 73.
26. C. Laird and H.I. Aaronson: *Acta Metall.*, 1969, vol. 17, p. 505.
27. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, p. 845.
28. J.W. Christian: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1821-39.
29. J.S. Bowles and W.J. McG. Tegart: *Acta Metall.*, 1955, vol. 3, p. 590.

29A. H. Warlimont: *Physical Properties of Martensite and Bainite*, The Iron and Steel Institute, London, 1965, p. 100.

30. R.D. Garwood: *Physical Properties of Martensite and Bainite*, The Iron and Steel Institute, London, 1965, p. 90.
31. P.E.J. Flewitt and J.M. Towner: *Acta Metall.*, 1966, vol. 8, p. 1013.
32. E. Hornbogen and H. Warlimont: *Acta Metall.*, 1967, vol. 15, p. 943.
33. I. Cornelis and C.M. Wayman: *Acta Metall.*, 1974, vol. 22, p. 301.
34. P.E.J. Flewitt and J.M. Towner: *J. Inst. Met.*, 1967, vol. 95, p. 273.
35. H.I. Aaronson, C. Laird, and K.R. Kinsman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 313.
36. I. Cornelis and C.M. Wayman: *Scripta Metall.*, 1973, vol. 7, p. 579.
37. G.W. Lorimer, G. Cliff, H.I. Aaronson, and K.R. Kinsman: *Scripta Metall.*, 1975, vol. 9, p. 271.
38. Y. Hamada, M.H. Wu, and C.M. Wayman: *Metall. Mater. Trans. A*, in press.
39. H.I. Aaronson and K.R. Kinsman: *Acta Metall.*, 1977, vol. 25, p. 367.
40. J.S. Bowles and C.M. Wayman: *Acta Metall.*, 1979, vol. 27, p. 833.
41. C.M. Wayman: *Proc. Int. Conf. on Solid-Solid Phase Transformations*, TMS, Warrendale, PA, 1981, p. 1119.
42. C.M. Wayman and G.R. Srinivasan: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, p. 310.
43. H.I. Aaronson: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, p. 270.
44. K.R. Kinsman, E. Eichen, and H.I. Aaronson: *Metall. Trans. A*, 1975, vol. 6A, pp. 303-17.
45. J.D. Watson and P.G. McDougall: *Acta Metall.*, 1973, vol. 21, p. 961.
46. K.R. Kinsman, R.H. Richman, and J.D. Verhoeven: *Mater. Sci. Symp. Abs.*, ASM, Metals Park, OH, 1974, p. 41.
47. M.G. Hall and H.I. Aaronson: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1923-31.
48. M.G. Hall, H.I. Aaronson, and G.W. Lorimer: *Scripta Metall.*, 1975, vol. 9, p. 533.
49. H.J. Lee and H.I. Aaronson: *Acta Metall.*, 1988, vol. 36, p. 787.
50. G.C. Weatherly, P. Humble, and P. Borland: *Acta Metall.*, 1979, vol. 27, p. 1815.
51. M.G. Hall and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, p. 1409.
52. K.R. Kinsman: Pittsburgh, PA, private communication, Intel Corp., 1988.
53. H.K.D.H. Bhadeshia: *Bainite in Steels*, Institute of Materials, London, 1992, p. 14.
54. A.B. Greninger and V.G. Mooradian: *Trans. AIME*, 1938, vol. 128, p. 337.
55. M.M. Kostic and E.B. Hawbolt: *Metall. Trans. A*, 1979, vol. 10A, pp. 165-76.
56. H.M. Otte and T.B. Massalski: *Acta Metall.*, 1958, vol. 6, p. 494.
57. H. Hu and C.S. Smith: *Acta Metall.*, 1956, vol. 4, p. 638.
58. R.F. Mehl, C.S. Barrett, and H.S. Jerabek: *Trans. AIME*, 1934, vol. 113, p. 211.
59. C.S. Barrett, A.H. Geisler, and R.F. Mehl: *Trans. AIME*, 1941, vol. 143, p. 134.
60. S. Takeuchi and T. Honma: *Sci. Rep. Res. Inst. Tohoku Univ.*, 1957, vol. 9, pp. 492 and 508.
61. B.C. Muddle, J.F. Nie, and G.R. Hugo: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1841-56.
62. J.S. Bowles and N.F. Kennon: *J. Aust. Inst. Met.*, 1960, vol. 5, p. 106.
63. G.V. Smith and R.F. Mehl: *Trans. AIME*, 1942, vol. 150, p. 211.
64. G.R. Srinivasan and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, p. 609.
65. G.R. Srinivasan and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, p. 621.
66. P.G. McDougall: Orlando, FL, private communication, University of New South Wales, Australia, Oct. 1986.
67. C.M. Wayman: *Metallography*, 1975, vol. 8, p. 105.
68. H.I. Aaronson, M.G. Hall, D.M. Barnett, and K.R. Kinsman: *Scripta Metall.*, 1975, vol. 9, p. 705.
69. S. Hoekstra: *Acta Metall.*, 1980, vol. 28, p. 507.
70. B.P.J. Sandvik: *Metall. Trans. A*, 1982, vol. 13A, pp. 777-87.
71. J.S. Bowles, B.C. Muddle, and C.M. Wayman: *Acta Metall.*, 1977, vol. 25, p. 513.
72. M.P. Cassidy, B.C. Muddle, T.E. Scott, C.M. Wayman, and J.S. Bowles: *Acta Metall.*, 1977, vol. 25, p. 829.
73. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. London A*, 1949, vol. 198A, p. 205.
74. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. London A*, 1949, vol. 198A, p. 216.
75. J.W. Matthews: *Phil. Mag.*, 1961, vol. 6, p. 1347.

76. P. Delavignette, J. Tournier, and S. Amelinckx: *Phil. Mag.*, 1961, vol. 6, p. 1419.
77. H.I. Aaronson: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 387.
78. J.L. Whitton: *J. Nucl. Mat.*, 1964, vol. 12, p. 115.
79. P. Gaunt and J.W. Christian: *Acta Metall.*, 1959, vol. 7, p. 529.
80. C. Zener: *Trans. AIME*, 1946, vol. 167, p. 550.
81. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, p. 365.
82. M.G. Hall, H.I. Aaronson and K.R. Kinsman: *Surf. Sci.*, 1972, vol. 31, p. 257.
83. G.V. Kurdjumow and G. Sachs: *Z. Phys.*, 1930, vol. 64, p. 325.
84. Z. Nishiyama: *Sci. Rep. Tohoku Univ.*, 1934, vol. 23, p. 638.
85. G. Wassermann: *Arch. Eisenhüttenwes*, 1933, vol. 16, p. 647.
86. K.C. Russell, M.G. Hall, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1974, vol. 5, pp. 1503-05.
87. D.J.H. Cockayne, *J. Microsc.*, 1973, vol. 98, p. 116.
88. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, p. 351.
89. Chergji Li and G.R. Purdy: *Proc. 4th Int. Congr. on Heat Treatment of Materials*, Berlin, 1985, p. 32.
90. Gary R. Purdy: *Scripta Metall.*, 1987, vol. 21, p. 1035.
91. C. Li, V. Perovic, and G.R. Purdy: *Phase Transformations '87*, Institute of Metals, London, 1987, p. 326.
92. T. Furuhara, J.M. Howe, and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, p. 2873.
93. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, p. 2857.
94. J.W. Christian and D.V. Edmonds: *Phase Transformations in Ferrous Alloys*, TMS, Warrendale, PA, 1984, p. 293.
95. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1988, vol. 22, p. 567.
96. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1988, vol. 22, p. 575.
97. H.I. Aaronson, J.M. Rigsbee, T. Furuhara, N. Prabhu, W.T. Reynolds, Jr., and J.M. Howe: *Scripta Metall.*, 1989, vol. 23, p. 279.
98. J.M. Howe, T. Abinandanan, C.S. Chiang, T. Furuhara, N. Prabhu, and H.I. Aaronson: *Scripta Metall.*, 1987, vol. 21, p. 1639.
99. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1988, vol. 22, p. 573.
100. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1988, vol. 22, p. 577.
101. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1989, vol. 23, p. 285.

101A. J. W. Christian: private communication, 1993.

102. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, p. 2887.
103. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, p. 1029.
104. H.S. Rosenbaum, D. Turnbull, and E.I. Alessandrini: *Acta Metall.*, 1959, vol. 7, p. 678.
105. K.H. Westmacott and U. Dahmen: *Proc. 40th Annual EMSA Meeting*, Claitor's Publishing Division, Baton Rouge, LA, 1982, p. 620.
106. H.I. Aaronson, G. Spanos, E.S.K. Menon, M.G. Hall, W.F. Lange III, and K. Chattopadhyay, *Structure and Deformation of Boundaries*, TMS, Warrendale, PA, 1986, p. 3.
107. H.I. Aaronson: *Trans. TMS-AIME*, 1962, vol. 224, p. 693.
108. E.P. Simonen, H.I. Aaronson, and R. Trivedi: *Metall. Trans.*, 1973, vol. 4, pp. 1239-45.
109. H.I. Aaronson and C. Laird: *Trans. TMS-AIME*, 1968, vol. 242, p. 1437.
110. K. Matsuura and S. Koda: *J. Phys. Soc. Jpn.*, 1963, vol. 18, suppl. 1, p. 50.
111. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, p. 639.
112. J.A. Hren and G. Thomas: *Trans. TMS-AIME*, 1963, vol. 227, p. 308.
113. H.J. Lee and H.I. Aaronson: *J. Mater. Sci.*, 1988, vol. 23, p. 150.
114. C.S. Smith: *Trans. ASM*, 1953, vol. 45, p. 533.
115. P.G. Shewmon: *Trans. TMS-AIME*, 1965, vol. 233, p. 736.
116. P.R. Howell and R.W.K. Honeycombe: *Proc. Int. Conf. on Solid-Solid Phase Transformations*, TMS, Warrendale, PA, 1982, p. 399.
117. W.F. Lange III, M. Enomoto, and H.I. Aaronson: *Metall. Trans. A*, 1988, vol. 19A, p. 427.
118. C. Laird and H.I. Aaronson: *Trans. TMS-AIME*, 1968, vol. 242, p. 1393.
119. R. Sankaran and C. Laird: *Metall. Trans.*, 1974, vol. 5, p. 1795.
120. J.W. Gibbs: *The Scientific Papers of J. Willard Gibbs, Vol. 1: Thermodynamics*, Dover Publications, New York, NY, 1961, p. 325.
121. J.M. Howe, U. Dahmen, and R. Gronsky: *Phil. Mag. A*, 1987, vol. 56, p. 31.
122. W.K. Burton, N. Cabrera, and F.C. Frank: *Phil. Trans. R. Soc. London*, 1951, vol. 56, p. 31.
123. H.I. Aaronson: *Phase Transformations*, Institution of Metallurgists, Chameleon Press, London, 1979, vol. 1, p. II-1.
124. R. Sankaran and C. Laird: *Acta Metall.*, 1974, vol. 22, p. 957.
125. G.C. Weatherly: *Can. Metall. Q.*, 1969, vol. 8, p. 105.
126. G.C. Weatherly and C.M. Sargent: *Phil. Mag.*, 1970, vol. 22, p. 1049.
127. G.C. Weatherly: *Acta Metall.*, 1971, vol. 19, p. 181.
128. G.C. Weatherly and T.D. Mok: *Surf. Sci.*, 1972, vol. 31, p. 335.
129. H.I. Aaronson and C. Wells: *Trans. AIME*, 1956, vol. 206, p. 1216.
130. E.S.K. Menon and H.I. Aaronson: *Acta Metall.*, 1987, vol. 35, p. 549.
131. D.W. Hoffman and J.W. Cahn: *Grain Boundaries and Interfaces*, North-Holland Pub. Co., Amsterdam, 1972, p. 368.
132. J.W. Cahn and D.W. Hoffman: *Acta Metall.*, 1974, vol. 22, p. 1205.
133. J.K. Lee and H.I. Aaronson: *Acta Metall.*, 1975, vol. 23, p. 799.
134. J.K. Lee and H.I. Aaronson: *Acta Metall.*, 1975, vol. 23, p. 809.
135. R.M. Aikin, Jr. and M.R. Plichta: *Phase Transformations '87*, Institute of Metals, London, 1988, p. 504.
136. S. Elangovan and M.R. Plichta: *Scripta Metall.*, 1986, vol. 20, p. 575.
137. W.K. Burton, N. Cabrera, and F.C. Frank: *Nature*, 1949, vol. 163, p. 398.
138. J.W. Cahn, W.B. Hillig, and G.W. Sears: *Acta Metall.*, 1964, vol. 12, p. 1421.
139. P. Merle and J.W. Martin: *Acta Metall.*, 1981, vol. 29, p. 1919.
140. K.E. Rajab and R.D. Doherty: *Proc. Int. Conf. on Solid-Solid Phase Trans.*, TMS, Warrendale, PA, 1982, p. 555.
141. S. Elangovan and M.R. Plichta: *Scripta Metall.*, 1985, vol. 20, p. 575.
142. H.I. Aaronson: *Trans. Indian Inst. Met.*, 1979, vol. 32 (1), p. 1.
143. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, p. 649.
144. W. Sun and K. Hiraga: *Phil. Mag. Let.*, 1993, vol. 67, p. 159.
145. R.F. Mehl, C.S. Barrett, and D.W. Smith: *Trans. AIME*, 1933, vol. 105, p. 215.
146. H. Brooks: *Trans. ASM*, 1952, vol. 44A, p. 20.
147. H.W. Paxton: *J. Chem. Phys.*, 1957, vol. 26, p. 1769.
148. J. Frenkel: *J. Phys. USSR*, 1945, vol. 9, p. 392.
149. J.M. Howe and N. Prabhu: *Acta Metall.*, 1990, vol. 38, p. 881.
150. N. Prabhu and J.M. Howe: *Acta Metall.*, 1990, vol. 38, p. 889.
151. G.R. Purdy and W.H. Zang: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1875-83.
152. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., John Wiley, New York, NY, 1982, p. 271.
153. C.A. Dube: Ph.D. Thesis, Carnegie Institute of Technology, Pittsburgh, PA, 1948.
154. C. Zener: *J. Appl. Phys.*, 1949, vol. 20, p. 950.
155. C. Atkinson: *Acta Metall.*, 1968, vol. 16, p. 1019.
156. C. Atkinson, K.R. Kinsman, H.B. Aaron, and H.I. Aaronson: *Metall. Trans.*, 1973, vol. 4, pp. 783-9.
157. M. Hillert: *Jernkontorets Ann.*, 1957, vol. 141, p. 757.
158. R. Trivedi: *Metall. Trans.*, 1970, vol. 1, pp. 921-27.
159. G.P. Ivantsov: *Growth of Crystals*, Consultants Bureau, New York, NY, 1960, vol. III.
160. G.J. Jones and R. Trivedi: *J. Appl. Phys.*, 1971, vol. 42, p. 4299.
161. G.J. Jones and R. Trivedi: *J. Cryst. Growth*, 1975, vol. 29, p. 155.
162. C. Atkinson: *Proc. R. Soc. London*, 1981, vol. A378, p. 351.

163. C. Atkinson: *Proc. R. Soc. London*, 1982, vol. A384, p. 107.
164. M. Enomoto, H.I. Aaronson, J. Avila, and C. Atkinson: *Proc. Int. Conf. on Solid-Solid Phase Trans.*, TMS, Warrendale, PA, 1982, p. 567.
165. M. Enomoto: *Acta Metall.*, 1987, vol. 35, p. 935.
166. M. Enomoto: *Acta Metall.*, 1987, vol. 35, p. 947.
166A. J.P. Horth: private communication, 1993.
167. M. Enomoto and H.I. Aaronson: *Scripta Metall.*, 1989, vol. 23, p. 1983.
168. R.H. Goodenow and R.F. Hehemann: *Trans. TMS-AIME*, 1965, vol. 233, p. 651.
169. F.J. Schoen and W.S. Owen: *Metall. Trans.*, 1971, vol. 2, pp. 2431-42.
170. M. Grujicic, G.B. Olson, and W.S. Owen: *Metall. Trans. A*, 1985, vol. 16A, pp. 1723-34.
171. W.S. Owen, F.J. Schoen, and G.R. Srinivasan: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 157.
172. M. Grujicic, G.B. Olson, and W.S. Owen: *Metall. Trans. A*, 1985, vol. 16A, pp. 1713-22.
173. U.F. Kocks, A.S. Argon, and M.F. Ashby: *Prog. Mater. Sci.*, 1975, vol. 19, p. 1.
174. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., John Wiley, New York, NY, 1982, p. 202.
175. L. Kaufman, S.V. Radcliffe, and M. Cohen: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 313.
176. M. Hillert: Swedish Institute for Metals Research, Stockholm, Sweden, unpublished research.
177. J.M. Oblak and R.F. Hehemann: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, p. 15.
178. R.F. Hehemann: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 397.
179. R.F. Hehemann, H.I. Aaronson, and K.R. Kinsman: *Metall. Trans.*, 1972, vol. 3, pp. 1077-94.
179A. W.T. Reynolds, Jr., F.Z. Li, C.K. Shui, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1433-63.
180. W.L. Grube and S.R. Rouze: *High Temperature–High Resolution Metallography*, Gordon and Breach, New York, NY, 1967, p. 313.
181. S.R. Rouze and W.L. Grube: General Motors Research Laboratory, Warren, MI, private communication, 1971.
182. E.P. Simonen and R. Trivedi: *Acta Metall.*, 1977, vol. 25, p. 945.
183. H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1985, vol. 1, p. 497.
184. A. Ali and H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1989, vol. 5, p. 398.
185. K. Chattopadhyay and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, p. 695.
186. E. Eichen, H.I. Aaronson, G.M. Pound, and R. Trivedi: *Acta Metall.*, 1964, vol. 12, p. 1298.
187. J.W. Christian and K.M. Knowles: *Proc. Int. Conf. on Solid-Solid Phase Trans.*, TMS, Warrendale, PA, 1982, p. 1185.
188. J.W. Christian: *Dislocations and Properties of Real Materials*, Institute of Metals, London, UK, 1985, p. 94.
189. H.B. Aaron and H.I. Aaronson: *Acta Metall.*, 1968, vol. 16, p. 789.
190. J. Goldman, H.I. Aaronson, and H.B. Aaron: *Metall. Trans.*, 1970, vol. 1, pp. 1805-10.
191. E.S.K. Menon and H.I. Aaronson: *Metall. Trans. A*, 1986, vol. 17A, pp. 1703-15.
192. M. Enomoto, H.I. Aaronson, and T. Furuhara: *Metall. Trans. A*, 1991, vol. 22A, pp. 1341-48.
193. K.R. Kinsman and H.I. Aaronson: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, p. 39.
194. J.R. Bradley and H.I. Aaronson: *Metall. Trans. A*, 1977, vol. 8A, pp. 317-22.
195. J.R. Bradley, J.M. Rigsbee, and H.I. Aaronson: *Metall. Trans. A*, 1977, vol. 8A, pp. 323-33.
196. K.R. Kinsman and H.I. Aaronson: *Metall. Trans.*, 1973, vol. 4, pp. 959-67.
197. U. Dahmen: *Scripta Metall.*, 1981, vol. 15, p. 77.
198. U. Dahmen: *Acta Metall.*, 1982, vol. 30, p. 63.
199. U. Dahmen and K.H. Westmacott: *Proc. Int. Conf. on Solid-Solid Phase Trans.*, TMS, Warrendale, PA, 1982, p. 433.
200. U. Dahmen, P. Ferguson, and K.H. Westmacott: *Acta Metall.*, 1984, vol. 32, p. 803.
201. U. Dahmen and K.H. Westmacott: *Acta Metall.*, 1986, vol. 34, p. 475.
202. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, p. 1963.
203. C.P. Luo and G.C. Weatherly: *Phil. Mag. A*, 1988, vol. 58, p. 445.
204. C.M. Wayman: *Introduction to the Crystallography of Martensitic Transformations*, Macmillan Co., New York, NY, 1964.
205. H.I. Aaronson and W.T. Reynolds, Jr.: *Phase Transformations '87*, Institute of Metals, London, 1988, p. 301.
206. C.A. Dube, H.I. Aaronson, and R.F. Mehl: *Rev. Met.*, 1958, vol. 55, p. 201.
207. F. Ishikawa, T. Takahashi, and T. Ochi: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 929-36.
208. S.A. Kulin and M. Cohen: *Trans. AIME*, 1950, vol. 188, p. 1139.
209. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., John Wiley, New York, NY, 1982, p. 532.
210. J.W. Christian: *Proc. R. Soc.*, 1951, vol. A206, p. 51.
211. E. Votava: *Acta Metall.*, 1960, vol. 8, p. 901.
212. E. Votava: *J. Inst. Met.*, 1961, vol. 90, p. 129.
213. J.A. Venables: *Phil. Mag.*, 1962, vol. 7, p. 35.
214. B.P.J. Sandvik and C.M. Wayman: *Metall. Trans. A*, 1983, vol. 14A, p. 823.
215. Jan. H. van der Merwe, Gary J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, p. 1165.
216. K.C. Russell: *Adv. Colloid Interfacial Sci.*, 1980, vol. 13, p. 205.
217. J.W. Cahn and J.E. Hilliard: *J. Chem. Phys.*, 1959, vol. 31, p. 539.
218. J.B. Hess and C.S. Barrett: *Trans. AIME*, 1952, vol. 194, p. 645.
219. L. Kaufman and M. Cohen: *Prog. Met. Phys.*, 1958, vol. 7, p. 165.
220. G.B. Olson and M. Cohen: *Metall. Trans. A*, 1976, vol. 7A, pp. 1897-1904.
221. H.K.D.H. Bhadeshia and D.V. Edmonds: *Acta Metall.*, 1980, vol. 28, p. 1265.
222. J.M. Rigsbee, E.S.K. Menon, H.J. Lee, and H.I. Aaronson: *Scripta Metall.*, 1983, vol. 17, p. 1465.
223. G.B. Olson and M. Cohen: *Metall. Trans. A*, 1976, vol. 7A, pp. 1905-14.
224. H.I. Aaronson and K.C. Russell: *Proc. Int. Conf. on Solid-Solid Phase Trans.*, TMS, Warrendale, PA, 1982, p. 371.
225. J.H. van der Merwe: *J. Appl. Phys.*, 1963, vol. 34, p. 117.
226. J.H. van der Merwe: *J. Appl. Phys.*, 1963, vol. 34, p. 123.
227. J.H. van der Merwe: *J. Appl. Phys.*, 1963, vol. 34, p. 3420.
228. C.M. Wayman: *Adv. Mater. Res.*, 1968, vol. 3, p. 147.

Crystallographic Theories, Interface Structures, and Transformation Mechanisms

J.W. CHRISTIAN

The structure and properties of an idealized planar interface that traverses a single crystal of the parent phase are first discussed. A (macroscopic) coherent interface is defined in terms of a relatively coarse-grained Larché-Cahn network related to the observed shape deformation. Reconstructive and displacive transformations are distinguished, and a new category of "diffusional-displacive" transformations is introduced. The crystallographic theory of martensite requires that the habit plane interface is atom conserving (or "glissile"), but nonconservative ("epitaxial") interfaces may form in some diffusional-displacive transformations. A modified Eshelby procedure is used to discuss the strain energy of particles of a new phase forming, by any mechanism, inside a constraining matrix. It is shown that the effective Burgers vector of a step (or ledge) in a fully or partly coherent interface is dependent on the parameters of the shape deformation and increases with the ledge height. Multiple height ledges ("superledges") should only be observed if their fields have been effectively neutralized, either by averaging over displacement directions that are spatially distinct but crystallographically equivalent or by combining with lattice dislocations through processes essentially equivalent to emissary slip or climb. In the latter case, the shape discontinuity is effectively transferred from the interface into the matrix or to a surface. The use of invariant line theories and the concepts of growth, structural, and misfit ledges are also examined.

I. INTRODUCTION

THIS article is an attempt to establish common ground among scientists interested in the growth of plate-shaped particles during reconstructive or displacive phase transformations in the solid state. Attention is focused on the nature of the interface in relation to the phenomenological theory of martensite crystallography (PTMC), transformation mechanisms, and observable shape changes. Many controversies in this field reflect linguistic rather than scientific difficulties, and to avoid this, I have inserted in the text a series of interlocking, dogmatic statements, each of which may be accepted or rejected by the reader. Most of these statements are results or theorems that, subject to specified approximations, are derivable from the basic assumptions of Euclidian geometry and Newtonian mechanics. The statements also include some definitions of terms to which other authors may have given different meanings. Thus, I hope that the statements contain only well-defined concepts; many are "obvious" results, but others may be challenged.*

*The statements are here given in a slightly different order, and some have been slightly amended since the conference. A reader who immediately accepts (or rejects) the whole list as obvious (or trivial) will not need to read the intermediate material.

I initially called these statements "axioms," intending this simply to mean statements that command general assent, but in mathematics and logic an axiom is an *unverifiable*, although widely accepted, initial assumption, almost synonymous with a postulate. The only true axioms in the present treatment are those that lead to the laws of classical physics, *i.e.*, to a Euclidian-Newtonian world. Gödel's (mathematical) proof that "no finite set of axioms can answer all the questions it raises" has recently been given some prominence,[1] and this emphasizes some of the difficulties of the present approach. However, I hope that this highlighting of certain results from geometrical, crystallographic, and elastic theories will help us to understand genuine scientific problems and to separate them from linguistic trivia.

Concepts such as homogeneity, coherency, deformation, correspondence, internal stress, *etc.* that are important in theories of phase transformation have to be examined over different distances before a complete description can be given. For example, the lattice correspondence often does not describe the relations between macroscopic vectors, and the concept of partly coherent interfaces that is useful at the atomic level has no meaning in a more macroscopic description. The important scaling factors are the interatomic distance or the unit cell size, the repeat distance in an interface with a periodic structure, the average separation of linear discontinuities that form arrays in some interfaces, and the average combined width of two adjacent, mutually twinned regions in duplex products.

II. PROPERTIES OF FINITE HOMOGENEOUS DEFORMATIONS

A homogeneous deformation **E** converts the vectors **u** into new vectors **v** = **Eu** where **E** is a second-rank tensor. In a coordinate system, **u** and **v** may be written as column matrices of components, and **E** is then a 3 × 3

J.W. CHRISTIAN, Emeritus Professor, is with the Department of Materials, University of Oxford, Oxford OX1 3PH, UK.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

matrix. A formal factorization separates **E** into a "pure" strain, **P**, preceded (or followed) by a pure rotation, **R**. The three principal axes are not rotated by the pure strain but change their lengths in ratios λ_i:1 where λ_i are the principal deformations, and $\lambda_i - 1$ are the principal strains. The tensor **E** is an invariant line strain (ILS) if a particular vector is unchanged in direction and magnitude and an invariant plane strain (IPS) if all vectors in a particular plane are invariant. When **E** is an IPS, there is one other undistorted (but rotated) plane, which becomes invariant if instead of **R**, a different rotation $\bar{\mathbf{R}}$ is combined with **P**. (In the special case when **E** is a simple shear, $\bar{\mathbf{R}} = \mathbf{R}^{-1}$.) The new IPS is $\bar{\mathbf{E}} = \bar{\mathbf{R}}\mathbf{P}$, where, in the terminology of deformation twinning, $\bar{\mathbf{E}}$ is the "conjugate" IPS.

If each λ_i is either greater or smaller than unity, there are no invariant lines (ILs) or invariant planes (IPs). At the other extreme, if there are two nonparallel IPs, all planes and vectors are invariant and there is no deformation, *i.e.*, $\mathbf{v} = \mathbf{Iu} = \mathbf{u}$ and each $\lambda_i = 1$. A single IP is obtained with either of two conjugate deformations if $\lambda_1 = 1$, $\lambda_2 < 1$ and $\lambda_3 > 1$. Finally, if $(\lambda_i - 1)$ is finite but does not have the same sign for all λ_i, an ILS, but not an IPS, may be obtained. Thus, the following are the first two dogmatic statements:

(1) A homogeneous deformation may include 0, 1, or an infinite number of invariant planes; there are no other possibilities.

(2) A general IPS may always be analyzed into a simple shear on the IP and a uniaxial expansion or contraction normal to the IP.

It follows that the matrix form of an IPS is

$$\mathbf{E} = \mathbf{I} + g\mathbf{e}\mathbf{n}' \qquad [1]$$

where g specifies the magnitude of the deformation, and **e** and **n**′ are column and row matrices giving the contravariant and covariant components of unit vectors in the displacement direction and normal to the IP, respectively. The shear and normal components, s and δ, are parallel to unit vectors **l** and **n** where:

$$g\mathbf{e} = s\mathbf{l} + \delta\mathbf{n}; \qquad \mathbf{n}'\mathbf{l} = 0 \qquad [2]$$

Consideration of macroscopic vectors now reveals the following statement:

(3) An ideal planar interface traversing a single crystal of the parent and having no long-range elastic field produces a macroscopic shape deformation that is an IPS for all transformations.

This is a trivial consequence of the need to maintain the continuity of matter at the interface. The impossibility of violating this condition is here emphasized only because it is often stated or implied that an IPS is characteristic only of a martensitic transformation.

In an isothermal, isobaric assembly, free from external deviatoric stresses, the idealized planar interface must separate two phases that are in equilibrium. Since there is no elastic energy, the equilibrium will not vary with the nature of the interface, whereas Larché and Cahn,[2] followed by many others,[3–7] have shown that for small embedded particles, the equilibrium may be critically dependent on whether or not the interface is coherent. A slowly moving planar interface in a slowly cooled or heated specimen with an imposed temperature gradient is not in equilibrium, but regions adjacent to it may be close to equilibrium. This configuration, familiar in experiments on directional solidification, is rarely obtained in solids, but it is useful to consider it before the more complex case of the inclusion. The IPS condition may be modified if small changes in length or orientation of interface vectors can be accommodated elastically, but the interface then has a long-range stress field.

The equation $\mathbf{v} = \mathbf{Eu}$ represents a physical deformation if displacement of the interface produces the finite strains of **E**. However, the transmission of shear stresses and strains implies some coherency at the interface; incoherent interfaces (like fluids) are considered to be "freely sliding;" although in practice, this may only be true at temperatures where thermally activated atomic migration becomes rapid. (Aaronson [private communication] has questioned whether incoherent interfaces ever form between two solid phases.) While the concept of coherency may be intuitively obvious, it is difficult to give a precise structural definition. In the theory of cubic grain boundaries, significance is commonly attached to coincident site lattice (CSL) orientations and especially to boundaries that contain a high density of coincident sites. This leads to definitions of coherence based on periodic boundary structures or the continuity of lattice planes and directions across the interface. However, the K_1 interfaces of deformation twins that are clearly coherent may also be irrational and hence nonperiodic, and across a rational K_1 interface, the lattices may be relatively translated so that there are no coincident sites and strictly no continuity of planes and directions. The concepts of continuity and coherence are nevertheless closely linked, and a definition of "macrocoherence," based on an imaginary common network, is contained in the next statement. The distinction between "fully" coherent (f-c) and "partly" (partially or semi) coherent (p-c) interfaces is made on an atomic level, and the structural implications are discussed later.

Larché and Cahn[2] introduced the concept of a network for two phases in coherent contact with each other. The network is "a certain identity . . . embedded in the solid [that] permits the definition of a displacement and hence a strain." In this definition, the specific structure of the interface is not considered, and in the cited examples, the crystal lattice usually defines the network. The network does not forbid diffusion of individual atom species, nor deformation by dislocation glide or climb, but it is disrupted at places such as free surfaces or dislocation cores.

Larché and Cahn[2] did not consider the effects of a periodic array of interface dislocations, but it is postulated here that local disruptions of a lattice network may still allow the network concept to be applied on a more macroscopic scale. The network is now defined, not by the crystal lattice, but by a "shape lattice," the primitive vectors of which have a length sufficient to average the local lattice deformations. The deformation **E** of the network changes, in general, the shape of the region transformed and implies a correspondence between large vectors that may differ from the correspondence-relating vectors of the two lattices. Figures 1 and 2 illustrate the

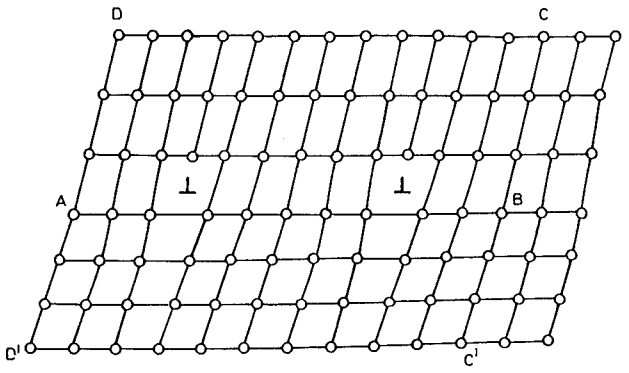


Fig. 1—Epitaxial p-c interface. The misfit dislocations must climb with the interface, so the number of atoms is not conserved. The shape network is represented by ABCD and ABC'D'.

"shape network" for an "epitaxial" and a martensitic p-c interface. Hence,

(4) Incoherent interfaces cannot transmit shear strains and thus do not cause measurable shape changes. Macroscopically coherent interfaces may be defined by the continuity of a coarse-grained Larché-Cahn network, and their displacements produce tilting of free surfaces unless the shear component is close to zero.

The previous results apply to any homogenous deformation whether of atomic or of macroscopic regions. However, **E** will always represent the macroscopic shape deformation, averaged when necessary over rather large distances, and a lattice deformation will be labelled **S**, and its principal deformations η_i.

III. LATTICE CORRESPONDENCE AND LATTICE DEFORMATION

Solid-state phase transformations are frequently classified as either reconstructive or displacive, depending on whether or not individual atoms are mobile. In one limit, the average atom has moved several thousand interatomic distances during growth, whereas in the other, an atom possesses insufficient thermal energy to make a single diffusive jump. In a chemical bond model, a pure displacive change occurs by stretching, compression, or bending of individual bonds, whereas a reconstructive change involves breaking and reforming of bonds. The two classes of transformation may also be labeled nucleation and growth and martensitic (probably the oldest terms used), diffusional and shear, or civilian and military.

In an entirely displacive transformation, each atom moves, relative to its neighbors, by less than an interatomic distance. This leads to a rather obvious but very important conclusion:

(5) The number of atoms must be conserved in any incremental volume transformed from the parent phase to the product phase by an entirely displacive (i.e., martensitic) mechanism.

Although St. (5) may also apply to many reconstructive transformations, its validity cannot then be assumed, since material transport (a counterflow of atoms and vacancies) may be possible when the atomic mobility is large.

The distinction between macroscopically coherent and incoherent interfaces is often assumed to correlate with the reconstructive-displacive classification. However,

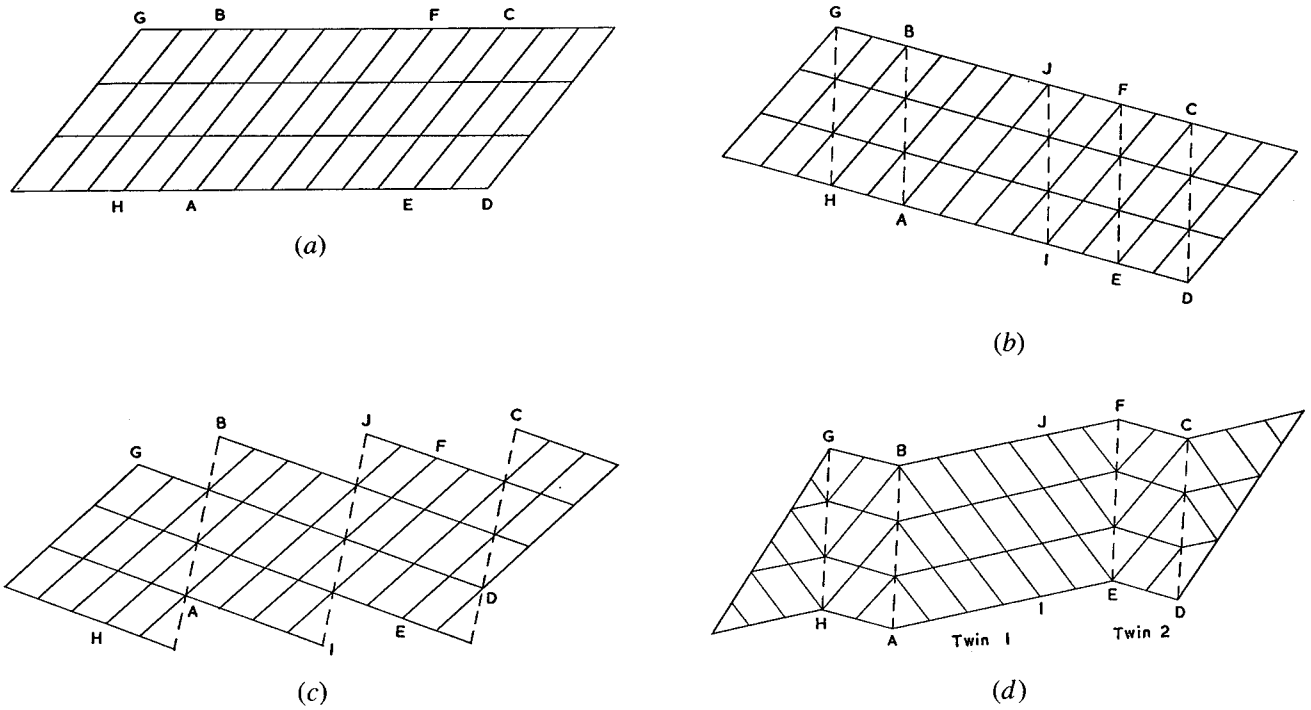


Fig. 2—Martensitic p-c interface.[8] In (*a*) and (*b*), habit plane sections of the parent and product structures related by the lattice correspondence are shown, while (*c*) and (*d*) give the actual product structure for single crystal and twinned martensite, respectively. The figure is not a true representation, since a rational habit plane is assumed, but the shape network that appears in (a), (c), and (d) as either ABCD or EFGH will be present in the more general case.

experiments[9–14] show that even when the atomic mobility is high, some ordered products or precipitates exhibit shape changes with appreciable shear components, and it seems inappropriate to class these changes as simply reconstructive. An early suggestion[15] that material transport should remove the shape change during growth was tentatively withdrawn in 1968 in view of some contrary experimental evidence, and more recent experiments support the conclusion that composition changes do not exclude concomitant shape changes, at least in a few cases. One possible reason for the failure of the original conclusion is that no account was taken of the theoretical difficulty that atom/vacancy sources and sinks might not be available, even under the high driving force of the shear in a constrained particle. It is tentatively suggested that transformations of this type might be called "diffusional displacive," where

(6) Reconstructive transformations have incoherent interfaces, but diffusional-displacive and pure-displacive transformations have coherent interfaces.

Any lattice may be generated from any other lattice by a homogeneous deformation **S**; *i.e.*, a new set of lattice points **y** may be derived from an original set **x** by a relation $\mathbf{y} = \mathbf{Sx}$. An unrelaxed model of the interface then has the points **Sx** and **x** on opposite sides of a plane through the common origin. There is an infinity of possible (although mainly physically implausible) alternatives to **S**. For a stationary interface, these notional deformations all represent the same unrelaxed structure and are thus equally valid, but if the transformation is displacive and the transformed volume is unconstrained, one particular deformation **S** gives the observed shape change. If this physically significant deformation is an IPS, and the interface plane is the IP, then the two structures match at the interface that is described as f-c and is characterized by $\mathbf{E} = \mathbf{S}$. An f-c interface may be either rational (*e.g.*, type I deformation twins and face-centered cubic-hexagonal close-packed (fcc-hcp) martensites[16]) or irrational (type II twins and certain Ti-Ta alloy martensites[17]).

A possible model of a rational, f-c interface has a lattice plane with coincident sites (atoms or centers of symmetry) common to the two crystals. Early theories of grain boundary structure emphasized orientations with a high-density CSL, but the actual structure may require a nonlattice translation **t** away from coincidence, so that

$$\mathbf{v} = \mathbf{Su} + \mathbf{t} \qquad [3]$$

is a more general expression for the lattice sites of the product structure. Nonzero values of **t**, have been predicted by computer simulations for K_1 interfaces in hcp[18] and even body-centered cubic (bcc) twins.[19] Thus, coherence on an atomic scale must allow a translation or, if the interface is diffuse, a curvature to break the lattice continuity of the interface. Hence

(7) An f-c, stress-free, planar interface between two single crystals is defined, irrespective of the detailed atomic structure, by the condition that S in Eq. [3] is an IPS.

In all displacive transformations, there exists, at least locally, a lattice correspondence that specifies how "labeled" lattice points (and hence vectors, planes, and unit cells) of one structure become *correspondingly* labeled points *etc.* of the other. The correspondence defines the pure lattice strain **B**, and *corresponding* cells contain equal numbers of atoms. For two single-lattice structures, the correspondence can (and usually does) relate all the atom sites. Otherwise, atomic "shuffles," which produce no macroscopic effects, are required to complete the structural change. Complex or lengthy (more than about one-half atom diameter) shuffles are improbable without thermal activation, so that the observed correspondences in both deformation twins and martensites either avoid shuffles or have simple shuffles. In martensites, the correspondence applies, within the uncertainty of possible shuffles, to labeled atoms, but in diffusive-displacive transformations, atomic migration prevents such a correlation. These two possibilities have been loosely described as atomic and lattice-site correspondences, respectively, but the imprecisely defined term "atomic correspondence" is probably best avoided. In summary,

(8) A lattice correspondence defines a one-to-one relationship between selected vectors, planes, and cells of the two lattices. A lattice correspondence defines the principal axes and principal trains of S but does not specify R.

IV. THE LATTICE INVARIANT DEFORMATION AND THE PTMC

Many measurements have confirmed that any macroscopic change of shape is an IPS, or some self-accommodating combination of different IPSs. Except for deformation twinning, f-c structures are very rare, because

(9) It is not possible, in general, to select a correspondence for which S is an exact or approximate IPS because of the restrictive condition that one principal strain of S must be (exactly or approximately) zero.

Thus, for the vast majority of phase transformations, the macroscopic change of shape is quite different from that of a unit cell. In a martensitic transformation, the lattice shape change may be modified into a macroscopic IPS either by averaging over the two lattice deformations of a finely twinned product or by forming an interface containing an array of so-called "misfit," "interface," or "anticoherency" dislocations, each with a lattice Burgers vector. (In a few martensitic products, faults of low energy are tolerated, so the interface dislocations may have partial Burgers vectors.) The array has a quasiuniform spacing and is displaced along planes **m** of the parent as the product grows, or along corresponding planes $\mathbf{m}^M$ of the product if the product shrinks. The planes **m** and $\mathbf{m}^M$ meet edge to edge in the interface along the invariant line of **S**. Such an array has been called "glissile" in previous publications,[20,21,38,39] but the term here has an entirely geometrical meaning, namely that the number of atoms is conserved.

The simple shear **G** produced by the array of anticoherency dislocations is a lattice invariant deformation (LID) that is homogeneous only on a macroscale, but it modifies the lattice shape change **S** into a macroscopic

shape change $\mathbf{E} = \mathbf{SG} = \mathbf{RBG}$, which is close to an IPS. (An exact IPS is calculated in the formal PTMC theories,[22,23] but the required shear seldom corresponds to a uniform spacing of the dislocations.) Hence,

(10) In a transformation to a single crystal of martensite, the LID of the original PTMC is a simple shear produced by a single set of discontinuities in the interface. These discontinuities are glissile (anticoherency) dislocations parallel to the IL of the lattice deformation.

The conclusion that a martensitic interface must contain a glissile dislocation array does not imply that a diffusion-controlled reaction may not also incorporate a similar interface. Thus,

(11) Planar interfaces containing a single-glissile array of anticoherency dislocations may also occur in certain diffusional transformations in which the growth rate is quite slow.

In the original PTMC theories, the dislocation and twin models were shown to be equivalent if K_1 corresponds to $\mathbf{m}^M$, $\boldsymbol{\eta}_1$ is parallel to the Burgers vector, and the magnitude of the LID shear is a positive fraction of the twinning shear. The twinned product is observed in many martensites.

The PTMC equations give, in the most general case, four independent solutions for the orientation relations, habit plane, and shape deformation, but certain symmetry conditions reduce the solutions to crystallographically equivalent variants of either two solutions or of a unique solution for each set of input data. The choice of the K_1 plane and $\boldsymbol{\eta}_1$ direction is severely restricted if the two lattice correspondences of a twinned product are equivalent variants. Either the parent plane $\mathbf{m}$ must be a mirror plane or the $\boldsymbol{\eta}_1$ direction must be derived from a two-fold axis of the parent matrix. Moreover, when either K_1 ($= \mathbf{m}^M$) or $\boldsymbol{\eta}_1$ is chosen, the other crystallographic parameter is fixed, and

(12) The input data for the PTMC are the structures and lattice parameters of the two phases and the plane and direction of the LID. If the product is twinned, the possible choice is limited to either twin plane or twin direction (not both) and is very severely restricted.

The predictions specify not only the orientation relations and the parameters of the shape and LID, but also the specific variants of each. So

(13) Experimental tests of the PTMC must include measurements of associated specific variants of orientation relations, habit planes, and shape deformation so that many independent parameters are compared with theoretical predictions. Agreement of all parameters thus indicates a high degree of probability that the PTMC theory is at least approximately correct.

Elaborations of the original PTMC theory include interfaces with small deviations from the IPS condition, *e.g.*, the Bowles-Mackenzie[23] dilatation parameter, or more than one set of anticoherency dislocations. If there is not an exact IPS, the interface is strained elastically and has a long-range stress field. Two or more sets of dislocations may still produce a simple shear LID if they have in common either the Burgers vector or the slip plane, but the resultant plane or vector may then have very high indices.

Hirth[24] has proposed a generalization of the basic PTMC equation $\mathbf{E} = \mathbf{RBG}$ to include a uniform dilatation and an additional lattice rotation that relaxes the stress field of an array of steps in an original (assumed rational) interface, as discussed in Section VI. The dilatation term might similarly be relaxed, at least partly, by misfit dislocations. Hirth thus suggests that such rotational and translational relaxations, which are not subject to any symmetry constraints, permit a range of possible orientation relationships, rather than the unique prediction of the PTMC. Each orientation has an "equilibrium" interface, *i.e.*, one without a long-range stress field. In a formal sense, the additional rotation in Hirth's Eq. [4] may be incorporated into $\mathbf{R}$, and the equation is then identical with the Bowles-Mackenzie formulation, except that the dilatation is not necessarily accommodated elastically. The equation may also be seen as a generalization of the assumption in PTMC that the LID is a simple shear, but it is important to note that any model which includes misfit of epitaxial type cannot be valid for martensitic growth because of St. (5) (see also Sts. (22) through (23) shown later). Finally, the most general form of the transformation matrix has to allow for two principal misfits within the interface plane, so that instead of Hirth's Eq. [4], the general transformation matrix becomes

$$\mathbf{A} = \mathbf{R}_1\,\mathbf{RBG}\boldsymbol{\delta}_1\,\boldsymbol{\delta}_2 \qquad [4]$$

where $\mathbf{E} = \mathbf{RBG}$ is an IPS, $\mathbf{R}_1$ is the additional rotation, and $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are uniaxial misfits in the plane of the plate. If $\mathbf{A}$ also represents an IPS (on a different plane), the equation may be rewritten in the standard form $\mathbf{A} = \mathbf{R}_2\mathbf{BG}_2$ where the additional rotation and misfit have been incorporated into the lattice relationship and the LID, respectively.

V. ENCLOSED PARTICLES CONSTRAINED BY A MATRIX

A very small enclosed particle of a second phase, in or near equilibrium, has a morphology that is mainly determined by the anisotropic interfacial free energy. If the misfit is appreciable, however, the strain energy will become increasingly important as the particle size increases. The strain energy is due to the resistance of the matrix to the change of shape that the particle is trying to impose on it, and strain energy dominance for large sizes arises because this energy is proportional to the volume of the particle, rather than to its surface area. The following is a rather obvious conclusion:

(14) The minimum strain energy for an isolated constrained particle is obtained when the minimum linear misfit is parallel to a long dimension of the particle and the maximum misfit is parallel to a short dimension.

Thus, the preferred morphology for a large one-dimensional misfit, is likely to be plates with a habit normal to the misfit direction, for a two-dimensional misfit, to be needles or laths with the long axis parallel

to the direction of best fit, and for misfits of comparable magnitude in all three dimensions, to be plates, needles, or equiaxed shapes with small dimensions only.

The strain energy may also strongly influence the microstructures of grains containing a number of particles, since it then includes contributions from their mutual interactions. An arrangement in which the strain field of each individual particle is largely cancelled by the fields of other particles may lead to a substantial decrease in the overall free energy. Interaction effects can be quite dramatic, leading, for example, to aligned rows of cuboidal precipitates and to the reversal of Ostwald ripening. The influence of strain energy on microstructure was first treated in detail by Khachaturyan[25] and Roitburd,[26] and there are now many articles dealing with this important subject. Only individual, isolated particles will be considered here.

A nucleus or a subcritical embryo will generally be elastically strained into full coherence with the matrix, because the interfacial energies are lower than those of p-c or incoherent interfaces. The matrix and particle form a self-stressed assembly, and the shape-dependent strain energy is proportional to the volume of the particle. In the exceptional case when **S** is an IPS, a plate-shaped particle parallel to the IP will have a low strain energy, because of the good fit along all major dimensions. However, steps or facets inclined to the IP are unavoidable in an enclosed particle and are responsible for the stress field in the assumed absence of accommodating plastic deformation in the matrix.

Essentially similar conclusions follow for particles in which **E**, but not **S**, is an exact or approximate IPS. An enclosed plate with the IP as habit will again have a stress field attributable to the forced shape compatibility at the steps or other facets. A p-c particle may, therefore, be treated as an f-c particle by replacing **S** by **E**. This equivalence breaks down at very small sizes where p-c particles are not possible. Hence,

(15) An enclosed (partly or fully) coherent particle has a (shape or lattice) network in common with the surrounding matrix. The continuity of the network implies that all interfaces are either naturally coherent or forced elastically into coherence.

It follows as a corollary that

(16) It is not possible for some interfaces or parts of an interface (*e.g.*, ledges or kinks) to be incoherent (*i.e.*, disordered) while others (*e.g.*, terraces) are (partly or fully) coherent.

Finally, even an enclosed incoherent particle has a strain field if the volume change cannot be eliminated by material flow in the time scale of an experiment. This strain energy will be small if atomic migration can ensure that the whole of the misfit is normal to a plate-shaped particle. Thus, in all types of interface, if the stress-free lattice or shape deformation approximates to an IPS, the strain energy is likely to be least for a plate shape with its plane parallel to the IP. If, however, **S** represents a pure expansion or contraction, the (isotropic) elastic energy of a particle with the *same elastic stiffnesses* as the matrix is independent of its shape, and

(17) If the stress-free deformation is a uniform expansion or contraction, the strain energy in isotropic linear theory is independent of the particle shape (Crum's theorem). If the misfit is large and has the same sign in all three dimensions, only small particles can form unless the misfit can be reduced or eliminated by material flow. Large equiaxed particles cannot form unless the net shape change is very small.

More exact calculations utilize the virtual operations introduced by Eshelby.[27,28] Eshelby's procedure was to remove some volume from the matrix interior, allow it to transform freely and displacively to a new phase by a lattice deformation **S**, then apply tractions to its surfaces to return it elastically to its original size and shape, reinsert it in the hole, and finally remove the surface forces and allow the self-stressed composite to relax elastically, obtaining the constrained strains in the matrix and particle and the strain energy. The finite strains of **S** define a symmetric elastic strain tensor with components $[(1/2)(S_{ij} + S_{ji}) - \delta_{ij}]$, and these "stress-free" strains were used in the linear elastic calculations.

This calculation thus treats both the matrix and the misfitting particle as linear elastic objects even though **S** or **E** are finite deformations. Eshelby considered only an f-c particle, and he obtained analytical results for particles of varying shape, specified by ellipsoids of differing aspect ratios. (Within an ellipsoid, the strain is uniform in both isotropic[28] and anisotropic[30] elastic treatments.) The theory was later generalized to include anisotropic effects[31] and reformulated by Khachaturyan[25,32] to separate the shape function from the elastic term. The calculation was adopted to p-c martensitic transformation[34] by defining the stress-free strains in terms of **E** rather than **S**.

These calculations may be used to justify the assumption of the PTMC that the habit plane is the IP of the shape deformation.[33,34] For an oblate-spheroidal martensite plate, the strain energy is linear in the aspect ratio (thickness:length), if **E** is an IPS on the plane of the plate. Thus, the strain energy can be made indefinitely small by reducing the aspect ratio, whereas in other conditions (**E** not an IPS or habit not parallel to IP), the energy increases with the dimensions of the plate. Hence,

(18) The strain energy of an enclosed plate-shaped particle can be decreased indefinitely by decreasing its aspect ratio if and only if its stress-free shape deformation is an IPS with the IP parallel to the habit plane. The assumptions of the PTMC are thus consistent with minimization of the strain energy of an enclosed plate.

The symmetrical strains used in linear elastic theory have the disadvantage that rotations and, hence, orientation relations cannot be included, and incorrect predictions may be made. In deformation twinning, for example, the linear theory predicts the equivalence of K_1 and the plane normal to η_1. Thus, although there are linearized versions of PTMC,

(19) Great care is needed if crystallographic results are obtained from the linear elastic theory. Spurious symmetries may be introduced, especially in linear elastic versions of the PTMC.

More general operations on the freely transformed particle can be envisaged[35] to follow **S** in order to

simulate nonmartensitic p-c interfaces or even incoherent interfaces. Figure 3 shows the Eshelby operations (a), (b), and (e), while (c) maintains coherency but allows growth to include a change in composition or long-range order. Further modifications are made if a (d) operation follows (b) or (c). In (d_1), the shape change is eliminated by the transfer of atoms to, from, and around the particle, and this requires high lattice diffusivity. As there is no strain energy, the particle may well be equiaxed. The operation (d2) is similar, but the atomic flux is only around the particle, so the number of atoms is conserved. This might occur if an incoherent interface is able to provide a rapid diffusion path; a similar flow around a coherent particle would require rapid lattice rather than interface diffusion and also suitable atom/vacancy sources and sinks.

Operation (d_3) allows regularly spaced dislocations to pass through the particle on their glide planes, thus producing an IPS shape change by displacement of a glissile interface as in the PTMC. Finally, (d_4) envisages the addition or removal of regularly spaced atomic planes, giving a uniaxial LID to compensate a misfit in a rational direction. This will also require rapid lattice diffusion unless the excess atoms can be inserted or removed by prismatic punching. Figure 1 shows why such an epitaxial interface cannot migrate conservatively. In summary,

(20) Plate shapes give the smallest strain energy for coherent particles with stress-free shape strains that approximate to an IPS and also for incoherent particles that grow conservatively with an appreciable stress-free volume change. With coherent particles, particular orientation relations and habit planes are necessary to obtain low energy

and

(21) The energy is reduced in diffusive displacive transformations if material flow is able to eliminate the shear component of the IPS shape change.

Interface misfits in two perpendicular directions may be corrected by cross grids of misfit dislocations, and various other possibilities have been extensively discussed.[36,37,85] If the principal misfits have opposite signs, much of the material flow may be conservative (*i.e.*, around the particle).

(22) The growth of partly coherent particles with epitaxial interfaces is, in general, non conservative and requires long-range material flow from or to the misfit dislocations which must "climb" in order to remain in the interface when it is displaced

and

(23) Long-range transport (St. (17) and (21)) may be rate limiting in some transformations unless there are slower processes needed for growth.

VI. THE FRANK-BILBY AND BOLLMANN EQUATIONS

The Frank-Bilby equation[51,20,34,38–40] for the dislocation content of a general grain boundary or interphase

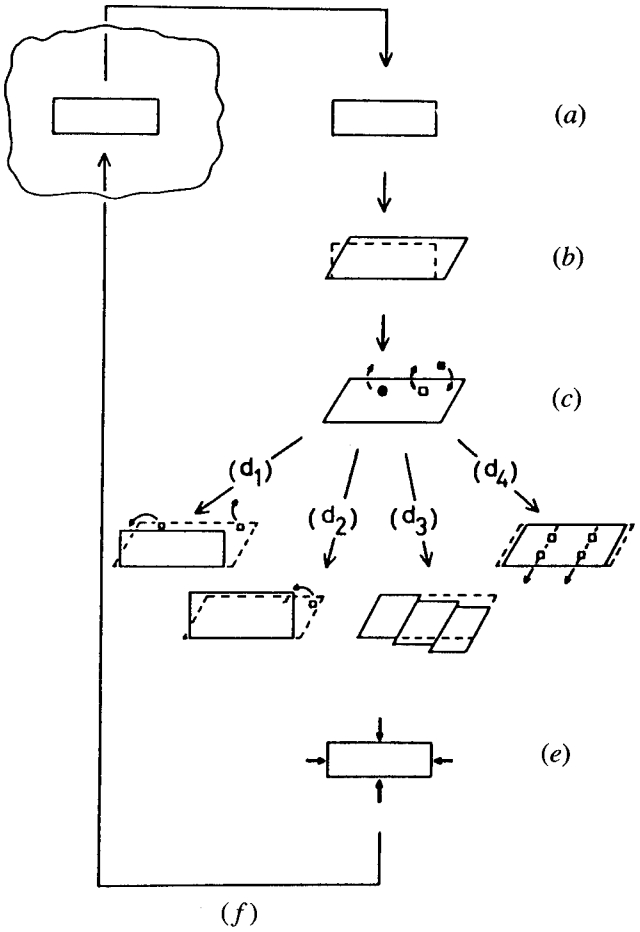


Fig. 3—Modifications of Eshelby's virtual operations to allow p-c and incoherent interfaces to be treated within the same (linear elastic) framework as the f-c particles originally considered.[35] For details, see text.

interface has the same form as Bollmann's O-lattice equation[34,41] and may be written as

$$\mathbf{b}_p = (\mathbf{S}^{-1} - \mathbf{I})\mathbf{p} \qquad [5]$$

here $\mathbf{b}_p$ and $\mathbf{Sb}_p$ give the net Burgers vector (referred to the parent and product lattices, respectively) of interface dislocations crossing any vector $\mathbf{p}$ in the interface. This equation is simply a measure of the net misfit in any direction $\mathbf{p}$ of a given interface plane $\mathbf{n}$ (*i.e.*, $\mathbf{p} \cdot \mathbf{n} = 0$) for two lattices connected by $\mathbf{S}$. The equation may be interpreted in various ways, the most usual of which is to regard $-\mathbf{b}_p$ as the net effect of discrete interface dislocations crossing $\mathbf{p}$. These dislocations form part of one or more parallel arrays and have Burgers vectors that are repeat vectors of the lattice in which $\mathbf{b}_p$ is defined. (A slightly more general equation defines $\mathbf{b}_p$ in any reference lattice.) Bollmann's O-lattice[41] uses an equivalent quantized approach, and the resulting dislocation array models are generalizations of the original low-angle boundary or epitaxial-misfitting boundaries. The dislocations are "primary" dislocations, referred to a single-crystal reference state, and they modify the lattice deformation to give an average shape deformation that is an IPS on the plane $\mathbf{n}$. Thus, as shown in a famous article by Bullough and Bilby,[42] Eq. [5] leads directly

to the PTMC and enables **n** and the rotational part of **S** to be determined, whereas if **S** and **n** are specified in advance, the misfit dislocations will not, in general, all be glissile.

An alternative interpretation of Eq. [5] is that it gives the total density of virtual (or coherency) dislocations in an interface **n** across which the phases are forced elastically into coherence. Christian and Knowles[43] emphasized that the dislocation content of an interface can be defined only in relation to a reference state of agreed zero density. There are two possible reference states, namely either full coherence as shown by a continuous lattice network, or the absence of a long-range stress field. The definitions coincide only if **S** is an IPS; otherwise, according to the network definition, the p-c boundary has a discrete dislocation content specified by the Frank-Bilby equation, while, with the stress definition, the f-c boundary has a quasicontinuous array of stress-generating ("coherency") dislocations with fractional Burgers vectors, and the p-c boundary has superimposed upon this the discrete anticoherency dislocation array to give a net Burgers vector content of near zero. This latter interpretation is largely due to Olson and co-workers,[44,45,46] who suggested the terms coherency and anticoherency, and to Bonnet and co-workers,[47–50] who regards the continuous array as a Somigliana dislocation and the discrete array as Volterra dislocations. Hence,

(24) The Frank-Bilby-Bollmann equation gives the total Burgers vector of the stress-generating coherency dislocations in an interface across which the two phases are elastically strained into full coherence, or equivalently the density of discrete anticoherency dislocations with lattice Burgers vectors that constitute a p-c interface with no long-range stress field.

The Frank-Bilby equation is concerned with compatible lattice and shape deformations. In Bilby's original derivation,[51] he considered a three-dimensional compatible continuous distribution of dislocations, specified by Nye's tensor dislocation density,[52] which he then confined to a thin layer and so defined a "surface dislocation." If this surface density is uniform, it exactly cancels the strain field of the interface (or of the coherency dislocations), but there is no network continuity. The interface is now incoherent with a high surface energy but no strain energy. Thus,

(25) If the density of anticoherency dislocations given by the Frank-Bilby equation is uniformly distributed along the interface, all network continuity is destroyed, and the interface has become incoherent.

There are two different approaches to interface structure. The first, much used in grain boundary theory, examines a selected CSL interface of given **S** and **n**, the energy of which is then minimized by atomic relaxation using interatomic force laws (usually pair-wise interactions). Adjacent interfaces and orientation relations are obtained by introducing "defects" (*e.g.*, displacement shift complete (DSC) dislocations) into this reference interface. Although **n** and the orientation relation are also usually assumed to be rational for an interphase interface, this will not now be an IP, and a limited variation of **S** or **n** may be permitted in order to lower the energy (*e.g.*, "structural ledges" may be introduced to rotate the interface plane). In the other approach, which stems mainly from the PTMC, the aim is to find an optimum interface, and the anticoherency dislocations or other defects are selected and used to minimize the strain energy and, thus, to determine **S** and **n**. In this approach, rational orientations or habit planes are exceptional.

As a p-c interface is displaced, the interface dislocations must glide or climb, and the net effects of these displacements produces the LID. In an entirely displacive transformation, only glide is permissible, but in a diffusive-displacive transformation, some climb of the anticoherency dislocations might be possible. Thus, it is not obvious whether a rational habit plane requiring material transport for its motion or an irrational plane able to move purely displacively will be preferred.

(26) Partly coherent interfaces normally have irrational (or high index) habits and orientation relations if the transformation is martensitic (*i.e.*, entirely displacive) but may have either irrational or rational habits in diffusive-displacive transformations.

VII. STEPS, LEDGES, AND INTERFACE DISLOCATIONS

The literature[29,63–68,76–78,81] contains many references to ledges, steps, terraces, and risers, qualified by adjectives such as growth, structural, misfit-compensating, *etc*. Step and ledge are used interchangeably in this article, but step will usually denote a minimum change of interface level, and larger changes will be called ledges, superledges, or facets. Terraces will denote the major areas of a stepped interface and risers the minor areas. (Riser is an area in this article, not a vector.)

The Russian physicist Vladimirskij[53] and (independently) Frank and Van der Merwe[54] first defined a "twinning dislocation" as a step in a rational coherent K_1 interface, and Bilby[55] called the corresponding step in an f-c interphase interface a "transformation dislocation." Vladimirskij's twinning dislocation is shown in

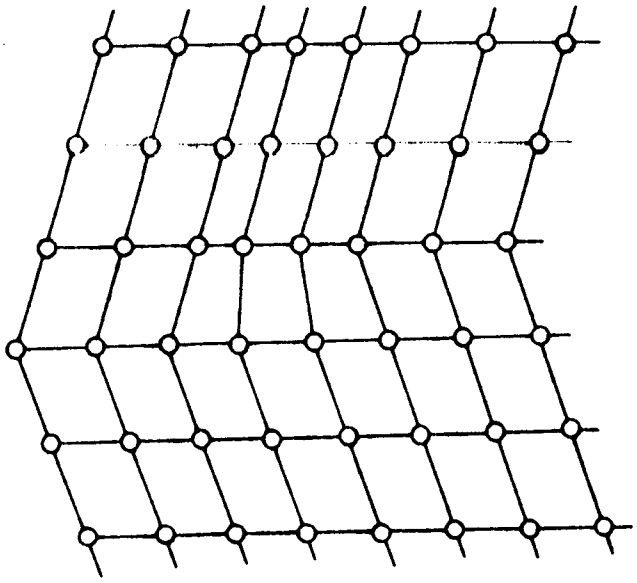


Fig. 4—Vladimirskij's twinning dislocation.[53]

Figure 4, and Figure 5 demonstrates that a step of height h in an f-c interface with $\mathbf{S} = \mathbf{E} = \mathbf{I} + g\mathbf{en}'$ has a Burgers vector;

$$\mathbf{b}_t = hg\mathbf{e} = h[\delta\mathbf{n} + s\mathbf{l}] \qquad [6]$$

The angle between the step direction and $\mathbf{l}$ determines the edge, screw, or mixed character of the step, and $\mathbf{b}_t$ is usually smaller than the smallest lattice vector. Both the direction and the magnitude of $\mathbf{b}_t$ may be irrational, except in cubic twinning. The next statement is essentially a definition.

(27) A step in a rational, f-c interface is a twinning or transformation dislocation with edge, screw, or mixed character and a Burgers vector in accordance with Eq. [6].

Note that although the plane of the riser may be inclined at any angle to the interface, h must be measured normal to the terrace interface.

A step normally separates two levels of the interface with identical structures so that h must be an integral multiple of the spacing d of lattice planes parallel to the interface. When two structures have different numbers of atoms in their primitive unit cells, h is a multiple of $d' = kd$ where d' defines "equivalent" cells with equal numbers of atoms. A parameter q gives the minimum h in type I twinning;[39] if $h_{\min} > d$, the step is called a zonal twinning dislocation, and a (possibly unstable) step of height d is an elementary twinning dislocation. Elementary steps are stable in all single-lattice twinning modes but not in some hcp modes. There is a similar but more complex condition[39] for a zonal transformation dislocation.

A zonal step may dissociate into elementary steps that mutually repel, but their separation will be limited by the extra energies of the interface structures between the elementary steps. As in the dissociation of a lattice dislocation, the equilibrium separation may be negligible, appreciable, or very large on an atomic scale, so that a zonal step may be unextended, extended, or nonexistent. In the later case, the different interface structures are energetically degenerate, and elementary steps are stable. It is also possible, in principle, for two different interface structures to have nearly equal energies and to be separated by defects with smaller Burgers vectors than those of elementary transformation dislocations and step heights less than the lattice spacing. An elementary dislocation may then either extend or dissociate completely. Atomistic calculations[56,57] indicate that in some bcc twins, for example, structures with $\mathbf{t} = 0$ and $\mathbf{t} = (1/12)\langle 11\bar{1}\rangle$, separated by a $(1/2)d$ step, have nearly equal energies.

A transformation dislocation, despite the normal component of its Burgers vector, may glide along the terrace plane, thereby conserving the number of atoms but not the volume. Since its motion displaces the interface through h or $-h$, it is also a growth ledge, as usually defined.[29,24,63,66] If the transformation is entirely displacive, the mobility of the step is limited only by its own oscillating self-energy, which leads to a "frictional" resistance or "Peierls-Nabarro force" opposing the motion. Steps are usually modeled as rather sharp changes, indicating high Peierls-Nabarro forces, but some atomistic calculations[56,57] predict diffuse steps and, thus, indicate a lower lattice resistance to the spontaneous nucleation and growth of a new layer than that given by a bulk elasticity calculation. In a diffusive-displacive transformation, the step's mobility may be determined mainly either by the interface kinetics of atom attachment or by any necessary ordering or long-range diffusion. Hence,

(28) Twinning or transformation dislocations are coherent growth ledges. They may move rapidly along the interface during displacive changes, but in diffusional-displacive transformations, their velocities may be limited by ordering, chemical diffusion, or long-range atomic transport.

The theory of transformation dislocations has largely been developed for rational, stress-free, f-c interfaces, but steps in interfaces in forced elastic coherence or rational p-c interfaces have similar properties. Figure 6

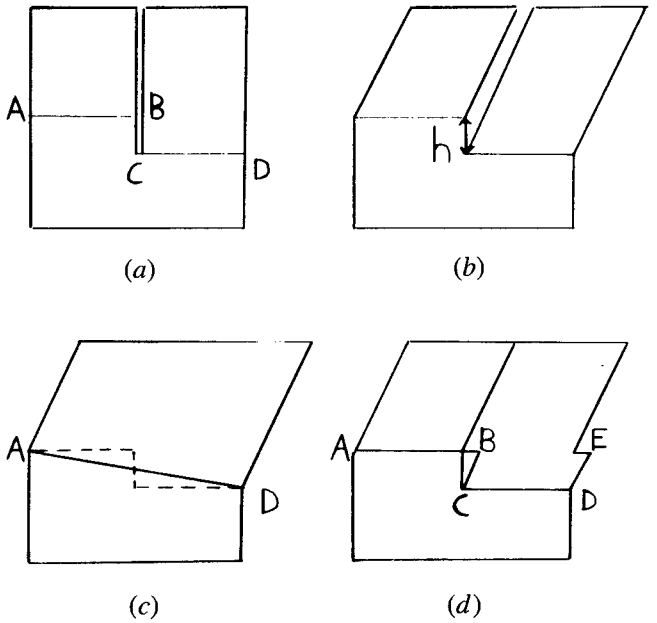


Fig. 5—Burgers vector of a transformation dislocation. A cut ending on the step in (*a*) becomes a fissure after transformation (*b*) with effective Burgers vector proportional to h. In (*c*), a step becomes an intrinsic part of the macroscopic interface, while (*d*) illustrates that the shear discontinuity across a step might be largely eliminated by a lattice dislocation of opposite sign.

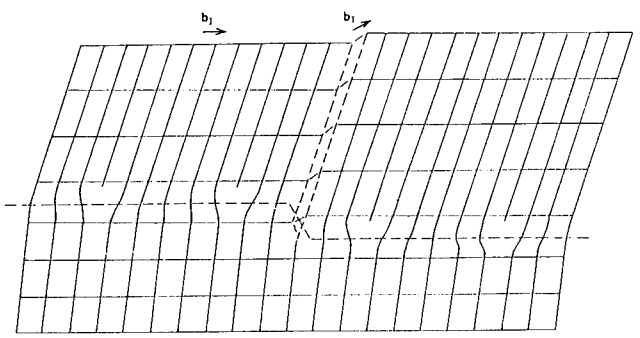


Fig. 6—A step in the p-c interface of Fig. 1.[39]

shows a step that has been formed in the p-c interface of Figure 1 by an operation equivalent to that in Figure 5. The effective Burgers vector is seen to be given by Eq. [6] where the parameters no longer relate to **S** but still specify the shape deformation **E**. The reason for this is that dislocation descriptions of steps involve long-range fields, and these are clearly dependent on the average shape deformation, rather than the local deformation. There are, however, some additional problems in p-c interfaces at an atomic level, especially the interaction between a step and the misfit dislocations of Figures 1 and 6. The remainder of this section will be mainly concerned with f-c structures, since this is how the theory has developed, but most of the conclusions are applicable also to p-c structures, since

(29) A coherent step in a rational p-c or f-c interface is a transformation dislocation/growth ledge with a Burgers vector given by Eq. [6] where the parameters are those of the shape deformation.

Note that since $\mathbf{E} = \mathbf{S}$ for an f-c interface, St. (29) subsumes St. (27).

Since the energy of a coherent step varies as $|\mathbf{b}_t|^2$, *i.e.*, as h^2, n steps of height d will have a lower energy than a "superledge" of height nd. Thus, with large s or δ, f-c superledges should dissociate spontaneously into elementary steps. This agrees with observations of elementary steps on bcc, fcc, and hcp twins and on some fcc-hcp martensitic interfaces. However, a superledge need not be f-c, since the shape discontinuity represented by the Burgers vector $n\mathbf{b}_{t,\min}$ may be reduced by adding suitable anticoherency dislocations, or in special cases, by averaging over shear directions that are crystallographically equivalent but spatially distinct. Such superledges may be described as p-c, whereas St. (30) is restricted to f-c superledges.

(30) An f-c, multiple height step (or superledge) in a rational f-c interface is unstable, and unless the lattice (and shape) deformation is very small, only elementary or zonal twinning or transformation dislocations of minimum permissible height should be observed.

The dependence of the energy on h^2 also explains why facets do not generally form on f-c interfaces and why a lenticular plate may have a lower energy than a rectangular parallelepiped. Figure 7(b) illustrates how the individual steps lose their identity as they are aligned to form a superledge or small facet. The array of virtual coherency dislocations defines the stress field of the faceted plate, just as the separated dislocations of Figure 7(a) define the field of the lenticular particle. Similar representations of cracks as virtual dislocation arrays are often used.

Equation [6] may also be interpreted in terms of the Frank-Bilby equation. If **S** is an IPS, Eq. [5] gives $\mathbf{b}_p = g\mathbf{e}(\mathbf{n}'\mathbf{p}) = hg\mathbf{e} = \mathbf{b}_t$ for any vector joining the top and bottom edges of the riser (see Figure 8). An f-c superledge of height nd or nd' may thus be regarded either as a single dislocation, with $\mathbf{b}_t = n\mathbf{b}_{t,\min}$, or as an array of n coherency dislocations, while a p-c superledge containing m anticoherency dislocations, with lattice Burgers vectors $\mathbf{b}_L$, is equivalent to a single dislocation of Burgers vector $\mathbf{b}_t - m\mathbf{b}_L$. (Exact cancellation to give a pure step will usually be impossible unless $\delta = 0$ and $s\mathbf{l}$ is an integral fraction of a lattice vector, a condition met only in cubic twins.) Any discrepancy, $\Delta h = h\delta$, in step heights will probably be taken elastically, since if Δh is compensated by misfit dislocations, the superledge must reject or absorb atoms as it migrates across its terrace. A superledge that tolerates Δh elastically is a glissile growth ledge, moving conservatively along the interface and producing a small volume change. This is

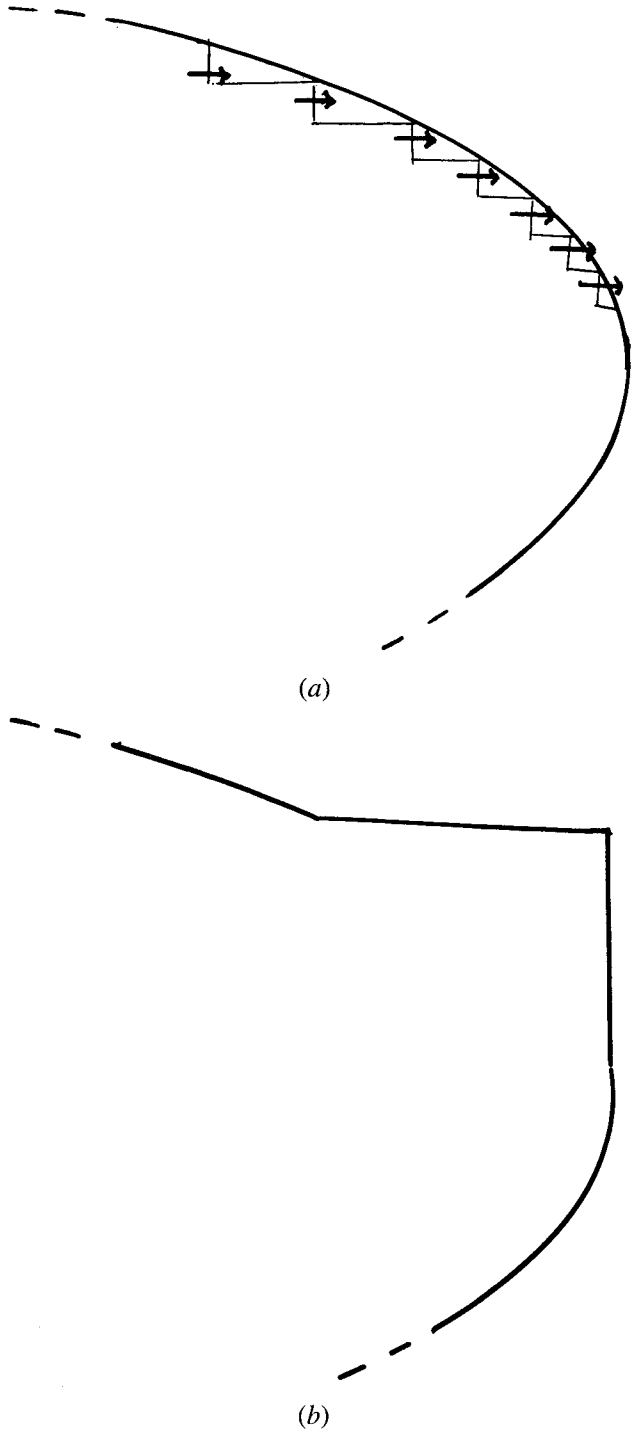


Fig. 7—(*a*) Step model of the tip of an enclosed, f-c, lenticular particle. In (*b*), the hypothetical formation of facets by alignment of steps in (a) is shown.

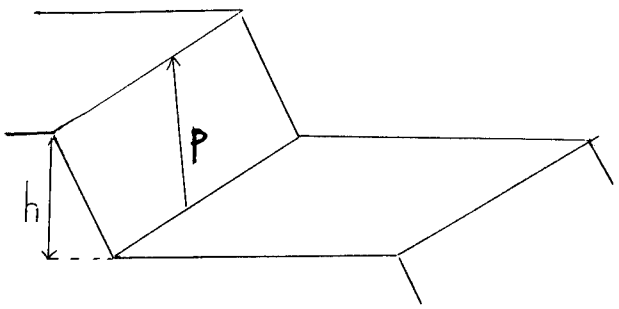


Fig. 8—The geometry of a superledge.

consistent with St. (1), since there are now two IPs, one exact and one approximate. With this model, the minimum p-c ledge height contains a single anticoherency dislocation and, thus, may be appreciably larger than the minimum height of an f-c ledge. However, ledges higher than the minimum will still be unstable unless the residual strain is very small.

Classification of the possible steps in a reference interface, and their associated Burgers vectors, is a general problem of considerable interest and is independent of the mechanism of formation. Lattice dislocations restore the crystal structure; similarly, perfect interface dislocations reproduce the interface structure, and a (geometrically) complete set of Burgers vectors is obtained for any step by combining $\mathbf{b}_t$ with any lattice repeat vector. Dislocations in a crystal are characterized by translational symmetry operations whereas in interfaces they depend on combined symmetry operations, one from each crystal. Two grains in a coincident site orientation and position, have interface Burgers vectors that are vectors of Bollmann's DSC lattice[41] and are not dependent on **t**. In most f-c interfaces, all dislocations will have DSC vectors, but Pond and co-workers[58–61] have recently developed a powerful general theory of interface defects that predicts additional dislocations in some bicrystals. In Pond's formulation, the two crystal lattices ("black" and "white") interpenetrate to form a dichromatic pattern, and the two structures similarly form the dichromatic complex. The DSC dislocations arise from the broken translational symmetry of the interface and may be introduced by a Volterra cut along the slip plane. Translational symmetry operations in the white and black crystals are denoted by vectors $\mathbf{w}(\boldsymbol{\lambda})$ and $\mathbf{w}(\boldsymbol{\mu})$ so that the step on each separate interface reproduces the interface configuration of its half crystal. When the crystals are rewelded, the mismatch gives for the Burgers vector of the step

$$\mathbf{b}_s = \mathbf{w}(\boldsymbol{\lambda}) - \mathbf{S}^{-1}\mathbf{w}(\boldsymbol{\mu}) \qquad [7]$$

The Burgers vector $\mathbf{b}_s$ is thus the difference between the two step vectors expressed in the same (white) coordinate frame, and the set of possible vectors are those that join white sites to black sites in the dichromatic pattern. Equation [7] may be regarded as a form of the Frank-Bilby equation.

Pond and co-workers[58–61] have also shown that in nonsymmorphic crystals with mirror-glide planes or screw-rotation axes, interface dislocations may also arise from combinations of aligned-point symmetry operations with either unequal or nonparallel intrinsic-glide components. These "supplementary displacement dislocations" have Burgers vectors that are modified by any nonlattice displacement **t** and usually form the junction between two (equivalent) facets, but they can form in a single interface in special conditions. An example is a twinning dislocation in an hcp $\{11\bar{2}1\}$ twin interface.

When possible, the atomic configuration at a step will minimize its Burgers vector, and this means that steps of minimum height with Burgers vectors other than $\mathbf{b}_t$ are improbable. However, superledges of height zd that are not required to maintain the lattice correspondence will generally have Burgers vectors smaller than $z\mathbf{b}_{t,\min}$. It follows that

(31) A superledge may be treated as a p-c facet with a small residual Burgers vector density or as a single dislocation with a particular Burgers vector, which is generally a minimum for a superledge of given height. The equivalence of the Frank-Bilby equation with a Volterra procedure defining $\mathbf{b}_t$ is formally valid even for small steps.

(32) Partly coherent superledges are unable, in general, completely to correct the misfit in their riser planes, and they form only if the shape deformation is small. The ledge will be (geometrically) glissile if the volume change of the transformation arises from a uniaxial strain, normal to the habit plane of the plate, and accommodated elastically at the ledge.

The operations just described have reduced the shape deformation for a fixed lattice change, but the model now implies a lattice correspondence that relates points only on every z planes where zd is the superledge height. This means that unless z is small, a majority of the atoms must shuffle, and this is improbable (and contrary to observation) in displacive transformations. Formally, the principal strains of **S** can always be reduced by choosing larger corresponding cells, but in practice, the avoidance of complex shuffles seems more important. However, in diffusive-displacive transformations, shuffles and p-c superledges may be more likely. Hence,

(33) Growth by p-c superledges requires an increase in the proportion of atoms that shuffle. In displacive transformations, this is improbable, but it might occur in some diffusive-displacive transformations.

In special cases (cubic twinning, fcc-hcp, and related "stacking transformations" without a volume change), pure steps may form. A pure step of fixed height is possible only if the set of vectors for that step includes a null vector. Expressed otherwise, a pure step has a riser plane that is a lattice plane of a CSL, *i.e.*, both δ and s in Eq. [6] must be either zero or an integral fraction of a lattice vector. There may, for reasons of symmetry, be more than one possible Burgers vector for a step of minimum height; *e.g.*, in an fcc deformation twin, $\mathbf{b}_t$ will generally be parallel to $\boldsymbol{\eta}_1$, but steps in annealing twins may have Burgers vectors in any of the three $\langle 11\bar{2}\rangle$ directions in the $\{111\}$ plane. If these three steps lie above one another on successive $\{111\}$ planes, they form a pure step, and repetition of this alternation will lead to a facet across which there is zero shape change, thus permitting any morphology.

The self-accommodation of the three shear directions

has interesting consequences in fcc-hcp and related transformations. Apart from a small change in the spacing of the {111}//{00.1} interface planes, the structural change may be obtained by the passage of a $(1/6)\langle 11\bar{2}\rangle$ dislocation (or growth ledge) along every second atomic {111} plane, whereas in fcc twinning, the same dislocation moves through every plane. In a deformation twin, or a martensitic transformation under an applied stress, successive planes will be displaced in the same direction, the steps will all be of minimum height (St. [29]), and there will be a macroscopic shape shear of $2^{-1/2}$ and $8^{-1/2}$ for the twin and the martensite, respectively. The same shear may be observed in the absence of an applied stress if growth utilizes a dislocation pole or similar mechanism.[38] However, if new ledges form readily, the large strain energy of the shear can be avoided by the previous alternation. The shape change would be exactly zero for a hypothetical hcp phase with ideal axial ratio, but in actual alloys, the step will have a small residual strain field because of the discrepancy in the spacings of the {111} and {00.1} planes.

A pure step in either a bcc or an fcc twin interface may also be obtained from a pileup of three identical twinning dislocations. In bcc metals and alloys, $\mathbf{b}_t = (1/6)\langle \bar{1}\bar{1}1\rangle$, and three identical steps on adjacent {112} planes may be changed into a pure step by addition of a lattice Burgers vector to give a net zero vector. Repetition of this configuration gives a superledge with zero net dislocation content. A similar result is obtained in an fcc twin by adding a lattice vector, $(1/2)\langle \bar{1}\bar{1}2\rangle$, or equivalently $(1/2)[0\bar{1}1] + (1/2)[\bar{1}01]$, to three twinning dislocations with $\mathbf{b}_t = (1/6)\langle 11\bar{2}\rangle$. If each of the $(1/2)\langle 0\bar{1}1\rangle$ vectors is added to one of the original $\mathbf{b}_t$ vectors, the pure step consists of the three $(1/6)\langle 11\bar{2}\rangle$ types discussed earlier; *i.e.*, the pure step is entirely equivalent to that obtained by shear alternation.

In both fcc and bcc structures, it is thus possible in principle to form a twin by migration of pure steps or larger superledges. The twin does not change its shape so that it need not be a plate and might form during annealing or recrystallization. The ledge is a p-c interface in which the anticoherency dislocations are so closely spaced that their cores overlap.

(34) The formation of a pure step or a pure superledge on a p-c or f-c plate is only possible in a few cases, notably cubic twinning, where there is a three-dimensional CSL. Otherwise, a p-c superledge may be formed in order to cancel all or most of the shear component of the IPS.

The physical formation of a pure step on an *existing* lenticular deformation twin requires a lattice dislocation to glide on a {112} plane from the matrix into the twin interface where it forms a three-layer pure step, and a succession of lattice dislocations every three planes will give a pure superledge. The apparent inconsistency between the discontinuous shape shear around the remainder of the twin interface and the zero shear across the facet (Sts. (15) and (16)) disappears when the lattice dislocations are included in the final description. In principle, these may have nucleated spontaneously as dipoles or loops, one-half of which glide into the interface. The residual segments may all remain in the matrix, or they may glide in the opposite direction to a free surface or a high-angle grain boundary. Either the matrix array or the surface is now the limit of a slipped region, *i.e.*, a shear discontinuity without structural change, which has replaced the shape change across the original f-c facet (Figures 9 and 5(d)). The same final result is obtained by emissary slip, as shown in Figure 10 and first considered by Sleeswyk.[62] Indeed, the two hypotheses are indistinguishable.

The bcc twin is a special case, inasmuch as {112} is both the (K_1) plane and an observed slip plane, and the Burgers vector of the cancelling dislocation is that of a stable lattice dislocation. Emissary slip can thus exactly accommodate the shear discontinuity at the edge of a constrained twin. In fcc metals, since $(1/2)\langle 11\bar{2}\rangle$ dislocations are not stable, $\{111\}\langle 1\bar{1}0\rangle$ slip on two intersecting planes must be invoked to cancel the twinning shear. In practice, more complex slip is likely to accommodate the very high stresses near the edge of a displacive product. When plastic accommodation is less regular, it is difficult to measure accurately the shape deformation.

(35) Emissary slip from the superledge (or equivalent slip originating in the matrix) may transfer the shape discontinuity to a limiting surface in the matrix or relax it at a free surface. Small deviations from a CSL mean that steps or ledges with very small residual Burgers vectors may be formed in a similar manner.

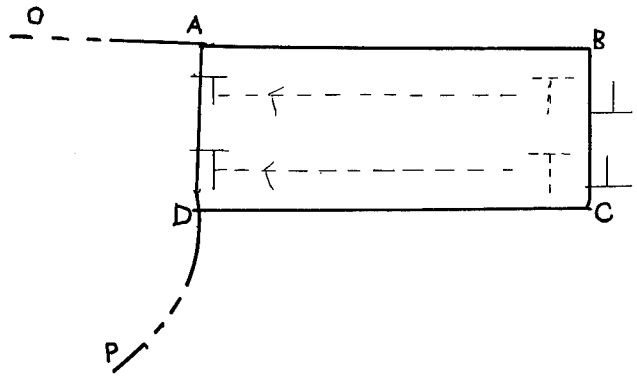


Fig. 9—To illustrate that if a superledge, AD, becomes p-c, in order to reduce its Burgers vector, the shape discontinuity of the particle, OADP, is transferred to the particle + slipped area, OABCDP, where B and C represent limiting dislocation sources. If B and C are on a free surface, the slip offset between them will equal the original stress-free shear strain across AD.

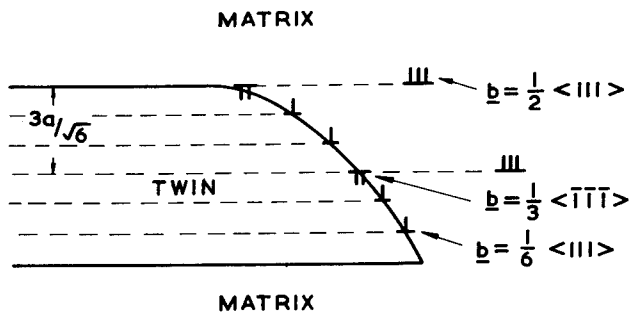


Fig. 10—Sleeswyk's model of emissary slip in bcc.[62]

VIII. IRRATIONAL STEPS AND INTERFACES

According to Eq. [6], steps in the rational K_1 interfaces of many noncubic, type I twins have Burgers vectors in irrational directions, while in type II twins the habit plane interface itself is irrational. It is not obvious how to specify a step when the spacing of lattice planes parallel to the interface is indefinitely small. The problem is compounded in many p-c martensitic interfaces that have both interface plane n and Burgers vector direction **e** irrational.

A high-index (or, in the limit, irrational) step on a rational interface may be modeled as a long period (or, in the limit, a nonperiodic) set of roughly equispaced kinks in a well-defined, rational step. For example, a first approximation to a particular step might have kinks every 8, 7, and 7 lattice repeat distances, and the next approximation might have a slightly longer repeat pattern . . . 8777 . . . while a better (but still not exact) model could have a longer repeat of . . . 8777877787787778777 . . .. In the limit, when the period has become infinite, the step is truly irrational, but the kink pattern remains ordered. The step may now be regarded as an example of a one-dimensional quasicrystal.

A similar approach may be made to the conceptual difficulties of an irrational interface. On an atomic scale, a high-index or irrational interface must consist of faceted sections of relatively close-packed (low index) planes. At least three differently orientated facets are required unless, as in type II twinning, the interface contains a (low index) rational direction **c**, in which case two alternating facets, intersecting along **c**, may represent the atomic structure. Presumably, these facets will be the two most close-packed planes in the zone of **c**, and the larger facets will be those nearest in orientation to the mean interface. If the smaller facets (H) are all simply steps between adjacent lattice planes of the larger facets (Q), and are spaced along Q at multiples of the spacing of lattice planes parallel to H in an ordered pattern with a very long or infinite period, as in the previous kink pattern, then the average interface will have high indices and will be irrational in the limit.

The steps would be ordinary transformation or twinning dislocations if the IP were parallel to the larger facets rather than to the irrational plane. For a matching irrational plane, however, they are an essential element in the structure of the boundary and have been called intrinsic dislocations;[39,47,48,82] they are also equivalent to structural ledges.

(36) A high-index or irrational interface containing a low-index direction may be modeled as an adjacent close-packed interface with an ordered sequence of elementary or zonal steps parallel to this direction. The steps are intrinsic elements of the structure if the interface plane is an IP, but the periodicity of the pattern of separation increases to infinity if the interface is truly irrational.

Since, with the correct step density, a planar, irrational IP boundary across a single crystal has no long-range stress field, there will be no net force on the individual steps in the presence of a chemical or mechanical driving force on the interface. Nevertheless, the interface will be displaced if the steps are able to move together as an array along their respective close-packed planes. With a planar interface, the correct spacing would have to be maintained by forming new steps at one edge of the interface at a rate that matches the vanishing rate at the other edge. Comparable growth in an enclosed particle requires a mechanism for interior nucleation of new steps. A step parallel to **c** is an intrinsic screw dislocation in type II twinning, where **c** is parallel to $\boldsymbol{\eta}_1$, but in general, it has mixed character. If, as previously suggested, intrinsic steps glide along their rational terraces as the interface moves forward, they are acting as transformation dislocations and may have local screw, edge, or mixed orientations. There may also be fluctuations in the local step density and the excess or deficit steps are then extrinsic transformation dislocations that will respond to an external stress in the same way as steps in rational K_1 interfaces. Figure 11 shows that these extrinsic dislocations are steps in the macroscopic irrational interface so that their motion displaces this interface. The magnitude of the effective Burgers vector is given by Eq. [6] with h equal to the height of the step measured normal to the irrational plane, *i.e.*, some multiple of the interplanar spacing of the Q lattice planes resolved in the direction of the normal **n** to the irrational habit plane. However, the structure of the Q facets on the two sides of a step of minimum height may not be identical, and then the equivalent of the zonal dislocation may have to be defined.

It is not clear whether or not the intrinsic step structure of an irrational coherent interface obviates the need for extrinsic steps in growth. In the analogous situation in crystal growth from the vapor, stepped interfaces are generally assumed to grow freely by migration of the intrinsic steps, but both the formation of new steps and their migration may be more difficult in a solid-solid interface, and there is some experimental evidence that structural ledges are relatively immobile. When elementary steps give different structures to successive Q facets, it need not follow that the intrinsic steps of the irrational interface will be zonal steps. It is conceivable, for example, that an alternation of the configuration at the Q facets has a lower total energy than a repetition. The following is the final assertion:

(37) The intrinsic steps of a high-index or irrational interface might possibly act also as transformation dislocations or growth ledges if new steps can form at a rate that will maintain the required density. An alternative growth mechanism utilizes excess steps

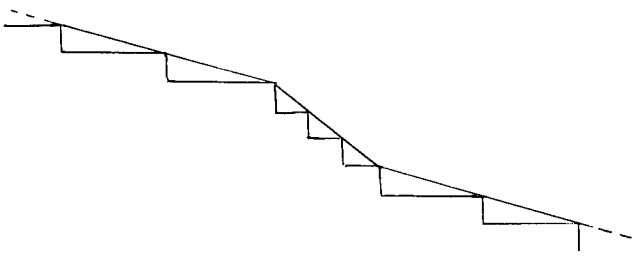


Fig. 11—A local variation in the density of intrinsic steps (structural ledges) may be regarded as an extrinsic step in the macroscopic interface.

that are extrinsic to the structure and that constitute transformation dislocations or growth steps in the macroscopic interface.

The foregoing treatment applies only to interfaces containing a suitable low-index direction; the structure and growth mechanism for the more complex general high-index or irrational interface will presumably be similar, but the facet intersections have not been discussed in detail. The use of structural features of an interface as a growth mechanism will probably be much more difficult in such cases, since there are then two or more sets of intrinsic steps that form a network in the equilibrium interface.

An equivalent formal treatment of a faceted interface is to use the Frank-Bilby equation to calculate separately the formal dislocation content required to correct the misfit along each of the facets H and Q. This corresponds to $-\mathbf{b}_T$ for an individual step facet P but sums to zero over a sufficient number of facets if the relative lengths of Q and H perpendicular to **c** give the exact IP interface. More generally, a faceted high-index plane that is not an IP has a net (coherency) dislocation content of

$$\mathbf{b}_p = (\mathbf{I} - \mathbf{S}^{-1})\mathbf{p} = (\mathbf{I} - \mathbf{S}^{-1})\mathbf{q} + (\mathbf{I} - \mathbf{S}^{-1})\mathbf{h} \quad [8]$$

where p is an average interface length made up of facets of lengths q and h (Figure 12). Various ambiguities in dislocation descriptions of stepped interfaces again depend on how zero content is defined.[43–47,20,61,82]

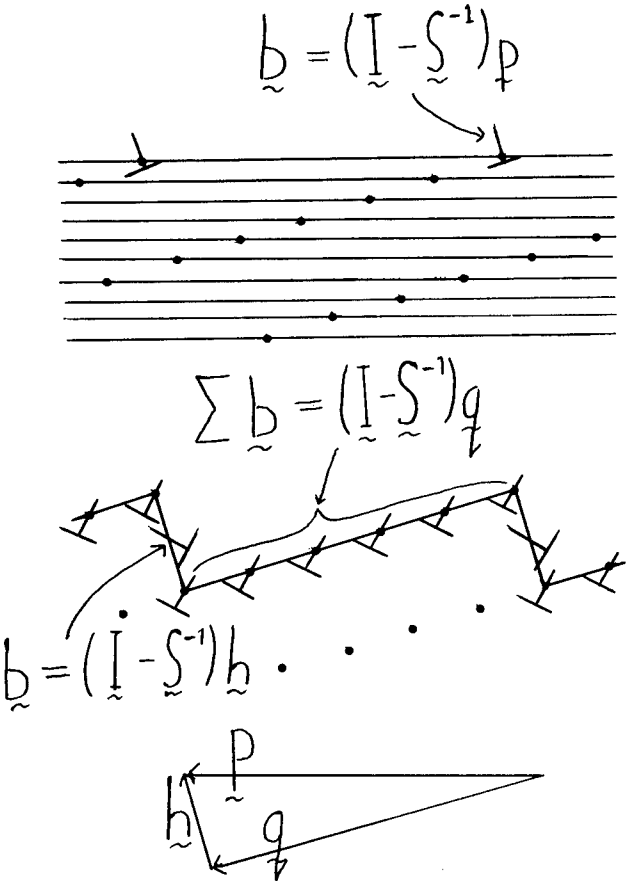


Fig. 12—Step model[43] of a high-index or irrational interface with a quasicontinuous distribution of coherency dislocations.

IX. SOME BONES OF CONTENTION

In this section, an attempt will be made to clarify some of the varied terminology and to resolve some of the more controversial issues that are found in recent literature. The previous statements may be quoted, but no additions will be suggested.

A. *Martensitic and Epitaxial Interface Models*

Coherent planar interfaces may usefully be divided first into those that conserve atom numbers (glissile or martensitic) and those that do not (sessile or epitaxial). All martensitic changes come into the first category (St. (5)) and may be either f-c or p-c, while changes involving some atomic mobility may have interfaces in either category (St. (11)). Fully coherent interfaces require **S** to be an IPS; the main experimental examples, apart from K_1 interfaces in deformation twins, are the rational interfaces between any two structures (fcc, hcp, dhcp, 9R, *etc.*) that differ only in the stacking sequences and interplanar spacings of almost identical hexagonal close-packed layers. Both martensitic and diffusion-controlled changes then utilize the same mechanism (transformation dislocations = growth ledges). Fully coherent irrational interfaces are rare, but they are found in some martensitic transformations in Ti alloys.[17] The p-c glissile interfaces predicted by the PTMC are usually irrational because of the IPS requirement and the assumption that the LID is a simple shear. Statement (37) shows that intrinsic steps (structural ledges) may be able (geometrically) to provide either martensitic or diffusional growth with PTMC irrational interfaces, but there is some experimental evidence[76] suggesting that such steps are not mobile during diffusional growth. In any event, a highly stepped interface will very probably contain a number of extrinsic steps because of fluctuations. The assumption that the interface and lattice orientation relations are rational usually requires epitaxial-type misfit dislocations, and hence, atomic and vacancy fluxes as the interface is displaced. More complex p-c interfaces may also be either conservative or non-conservative. Curved interfaces are modeled as stepped planar interfaces; they have a long-range stress field, as does any planar interface, the stress-free shape deformation across which is not an IPS.

Different interpretations of glissile and sessile may be found in the literature, but no such ambiguity extends to epitaxial. If the "terraces" of rational interfaces are p-c and epitaxial, a new layer, forming as a growth ledge migrates, must contain either more or fewer atoms in a particular interface direction than did the parent phase layer from which it formed. In contrast, atom numbers are conserved as the step moves over the terrace of a glissile p-c interface. Thus, the interface dislocations effectively either climb or glide through the new layer added by the step.

Much confusion, partly linguistic and partly scientific, has arisen in discussions of the motion of (postulated) epitaxial interfaces. As such an interface is displaced,

the misfit in any direction must be corrected by addition or removal of an appropriate fraction of the atoms, and however achieved, this is correctly described as a nonconservative process.[35,64] Frequent references to conservative transfer of misfit dislocations from an existing terrace to the new terrace produced by an approaching growth ledge, and an attempted justification of these claims by a later redefinition of conservative,[65] actually refer to nonconservative mechanisms based on prismatic punching or cross slip rather than on lattice diffusion. Statements[64,68] that the misfit dislocations must "climb" if the interface is to migrate have aroused much opposition,[63,65–67] but even with prismatic punching it is difficult to avoid some climb. In conservative mechanism (1) of Howe *et al.*,[91] this takes place when the expelled loop either shrinks on to the next terrace or disappears in the matrix; in mechanism (2), it occurs in the final reorientation of the screw element on the upper terrace; and mechanism (3) cannot retain the epitaxial structure without climb. Prismatic punching requires very high stresses[83,84] and seems improbable in comparison with simple climb (*e.g.*, addition or removal, at growth ledges, of vacancies as well as atoms) so that the misfit dislocations move from one terrace to the next as the growth ledge passes it.

An important result that now follows is that an epitaxial interface cannot be a correct model for bainitic or proeutectoidal plates of ferrite forming from austenite at temperatures where all atoms except interstitials are immobile. This implies, in turn, a rejection of the Rigsbee-Aaronson model[70] for ferrite plates formed with a Nishiyama orientation at temperatures as low as 400 °C. This criticism arises from the edge misfit dislocations, which ensure that displacement of this interface would be nonconservative, rather than the structural ledges on which much emphasis has been placed. Actually, Rigsbee and Aaronson's interesting transmission electron microscopy (TEM) observations of the single dislocation arrays in ferrite plates showed on analysis that in six out of seven cases, the Burgers vector of these dislocations was one of the two fcc ⟨110⟩-type vectors that correspond to bcc ⟨111⟩ vectors, whereas the Burgers vector of the model interface shown in Figure 1 of their article[70] corresponds to a ⟨100⟩ bcc vector. The observed Burgers vectors are both consistent with the usual application of the PTMC to ferrous martenites, and the implication[64] that the interfaces they studied were martensitic, not epitaxial, is confirmed by examining the relation between the Burgers vector of the dislocation array and the macroscopic habit plane.

More generally, the kinetic restraints on epitaxial growth may mean that such interfaces do not appear in competition with possible martensitic-type interfaces, or with habit planes or lattice orientations that are near to, but not exactly rational, but in which the remaining habit plane misfit has been reduced to a value that can be tolerated elastically. This latter possibility would imply a stress-free shape change **A** given by Eq. [4] (or Hirth's Eq. [4]). A probable example is the matrix-γ $AlAg_2$ interface where the networks found by Laird and Aaronson[86] imply a slight deviation from the rational f-c relation in order to compensate for the slight mismatch between the interatomic distances in the close-packed {111} planes of the cubic matrix and the (00.1) planes of the hexagonal γ precipitate.

B. *Structural Ledges and Invariant Lines*

Structural ledges, one to a few atom planes high, were introduced by Hall *et al.*[69,71] and Rigsbee and Aaronson[70] who pointed out that a regularly spaced array of steps of atomic height may be able to reduce the misfit along the rational interface direction normal to the ledge. A roughly equispaced array of such steps rotates the interface about the step direction **c** so that the interface vector normal to **c** approaches an IL. Dahmen[72] suggested that it becomes an IL, and although this is generally not possible (subsequently shown), it may be assumed as an ideal condition. If **c** is also an IL, the new high-index, or irrational, interface is now an f-c IP of **S**, and the structural ledges are the intrinsic steps of Sts. (36) and (37). If **c** is not an IL, the misfit may be corrected by a single set of misfit dislocations to give an epitaxial p-c interface. This is now a Rigsbee-Aaronson type model, but it is instructive to examine more closely the effects of the structural ledges. By choosing orthonormal coordinates x_i with x_2 parallel to **c** and x_3 normal to the stepped close-packed plane, and if vectors parallel to x_1 and x_2 change lengths in the ratio $1 + \varepsilon_1 : 1$ and $1 + \varepsilon_2 : 1$, then **S** is

$$\mathbf{S} = \begin{pmatrix} 1+\varepsilon_1 & 0 & s \\ 0 & 1+\varepsilon_2 & 0 \\ 0 & 0 & 1+\delta \end{pmatrix} \qquad [9]$$

The misfit dislocations are edges with Burgers vectors and spacings $\mathbf{b}_2$, and $\nu_2 = b_2/\varepsilon_2$. If there are ledges of height h and spacing ν_1 in the x_1 direction, the net Burgers vector of each ledge is given by Eq. [5] as the following:

$$\mathbf{b}_h = h[s\ 0\ \delta] \qquad [10]$$

$$\mathbf{b}_1 = \nu_1[\varepsilon_1\ 0\ 0] \qquad [11]$$

and the resultant Burgers vector along the rotated, high-index interface is the sum of $\mathbf{b}_h$ and $\mathbf{b}_1$. This is zero if

$$\frac{h}{\nu_1} = \frac{-\varepsilon_1}{s}; \quad \text{and } \delta = 0 \qquad [12]$$

Thus, the discrepancy in interplanar spacing must be small if the structural ledges are to be efficient, but the shear on the rational plane (which controls the "pattern advance"[73]) must not be small. Equation [12] is just the condition, $\mathbf{S}[\nu_1, 0, h] = [\nu_1, 0, h]$ for the rotated vector normal to **c** to be an IL; $\delta = 0$ is included because the line must be unrotated as well as unchanged in length. The difference in height is ignored in most discussions of the structural ledge but is included in the recent work of Van der Merwe *et al.*[73] Hirth,[24] following Howe and Smith,[74] considers first the case when there is perfect fit along the original interface so that the ledge is an ordinary transformation dislocation, and he shows how

a new equilibrium (stress free) interface can be constructed either by a planar array of steps with their stress field cancelled by a relative rotation of the two lattices or by having superledges of small net Burgers vectors, as previously discussed.

Some of the IL theories have been related to this model by a slightly different assumption, namely that the interface is retained as two parallel rational planes, but the azimuthal orientation of the two phases is rotated away from one of the "rational" relations so as to bring an IL within the interface. This has been used to rationalize some of the measured relations in fcc-hcp-bcc structures.[75] However, an interface that is planar has no obvious energetic advantage in containing an IL, and it follows from the complete range of measured lattice parameters for ferrous alloys that there can be no IL for any orientation within the usually reported Kurdjumov-Sachs/Nishiyama range.[21] Applications of the IL concept to plate-like precipitates may often be treated as simplified two-dimensional versions of the PTMC which predicts that the two close-packed planes cannot be exactly parallel. Unlike the IP conditions, those for an IL are readily satisfied, but it is not obvious which of the cone of potential ILs should be chosen. The IL theory should be applicable mainly to needles or laths.

In some recent articles,[76,77,78,92] a distinction is made between structural ledges and "misfit compensating ledges." The latter are stated to have larger Burgers vectors of the order of one-half that of a misfit dislocation on the close-packed interface and also have higher step heights (but no correlation between step height and Burgers vector has been reported). They are assumed to compensate for the misfit in a similar manner to epitaxial misfit dislocations. Misfit compensating ledges have been discussed mainly in relation to experimental investigations of the precipitation of a hcp phase from a bcc matrix, *e.g.*, in Ti-Cr alloys, where the lattice deformation differs only slightly from an IPS. It is difficult to believe that they are somehow different from ordinary structural ledges.

Distinctions between ledges and interface dislocations are usually unjustified and ignore the fact that the two terms generally have the same meaning. Thus, the only ledges without dislocation character are the pure steps of St. (34), and the only nonstepped interface dislocations are those with *lattice* Burgers vectors in the interface. Structural ledges are not, as often claimed, an alternative to misfit dislocations; they are intrinsic structural elements of an interface that is different from the rational interface of the misfit dislocations, and they achieve better matching because of the pattern advance, which is simply the geometrical consequence of the shear component of a nonlattice Burgers vector. On a sufficiently high ledge, the normal component might be compensated by misfit dislocations, but (St. (32)) this would prevent conservative migration of the ledge.

C. *Growth Ledges, Superledges, and Transformation Dislocations*

Ledges of any height that are not an intrinsic part of the interface structure may be regarded as growth ledges or transformation dislocations. Examples are isolated steps on a coherent interface or locally excess or deficit steps in a faceted high-index interface. The distinction between structural and growth ledges or, more generally, between intrinsic and extrinsic defects, is not always very sharp, and one function may merge into the other. As noted previously, an example of this is given in Hirth's article.[24]

In St. (30), it was emphasized that f-c superledges with large Burgers vectors should be unstable, but large semicoherent superledges are possible in some cases. Superledges with a very small Burgers vector have been observed in recent work on fcc-hcp diffusional transformations in Al-Ag alloys[79,80] and on martensitic transformation in Co-Ni alloys.[93] Both observations are consistent with the foregoing theory but apparently conflict with earlier reports of a large-shape shear.[11,65,66,86,94] Thus, in different experiments, both diffusional and martensitic transformations utilize coherent growth ledges of minimum height and a common Burgers vector and multiheight p-c growth superledges that migrate without appreciable shape strain. Aaronson has suggested that while an enclosed particle may prefer growth without shear, once a plate has emerged at a free surface, subsequent thickening might be facilitated by the choice of a single preferred displacement direction. The question of whether a coherent shape strain can be continuously removed during or after the growth has been addressed only by Liu and Aaronson,[11] whose interesting results showed that the surface tilt in Al-Ag alloys is removed slowly, subsequent to growth.

D. *Growth Mechanisms*

The analogy with the theory of crystal growth from the vapor must be used with caution; there is no concept of coherency across a solid-vapor interface, and kinks in steps are the ultimate sites at which atoms condense (and from which they evaporate) simply because they are the surface sites of maximum binding. In a simple model, an atom condenses on to the terraces and diffuses over the surface until it either evaporates back into the vapor or finds a kink in a step where it remains. While growth may also involve kinks in steps on solid-solid interfaces, it seems improbable that diffusion of atoms along the interface will be important. In twinning and martensitic transformations such as fcc-hcp, the evidence suggests that the step can glide as a unit with such a high velocity that separate kink motion seems unlikely. The kinks used in the crystal growth theories are defined for elementary steps; a kink on a superledge is presumably itself a short elementary step, which must be kinked again in order to give a condensation site of atomic dimensions.

Ledge growth in a diffusional transformation requires (a) lattice diffusion to effect the composition change, (b) some mechanism for the formation of new ledges, and (c) maintenance of the interface structure. In a steady state, these serial processes must all conform to the overall rate that is fixed mainly by the parameters of the "rate-controlling" process (*i.e.*, that which consumes the major part of the driving force). For an epitaxially p-c interface, the foregoing process (c) implies that the misfit dislocations climb. Any one process, (a), (b), or (c)

(and there may be others), could be rate controlling; what is not permissible is to select process (a) as rate controlling but resist the implication that (b) and (c) must then be faster processes than (a) at a fixed driving force.

E. *The Significance of a Shape Change Accompanying a Transformation*

A shape change with a significant shear component produces characteristic tilting of an originally flat surface and a change of direction of fiducial lines (scratches) inscribed on the surface before transformation. Such a change was for many years regarded as the hallmark of a martensitic transformation, but there is now good evidence that shape changes can occur in some diffusional transformations and indicate a lattice or shape correspondence that is maintained despite the diffusion. It was suggested[15] in 1962 that the shear component of any such shape change would be removed by a flux of atoms in and around the interface from highly compressed regions to highly expanded regions. The driving force for this flux is the elastic energy of a constrained particle, but it is not clear whether Bardeen-Herring sinks and sources are required or whether the critical nucleus size is so small that individual atoms can enter or leave the solid solution. The current belief is thus that a coherent diffusional transformation can lead to an IPS-type shape relief that may or may not subsequently disappear.

In some nonmartensitic transformations, the predictions of the PTMC all agree well with experimental results, and there can be a high degree of confidence (St. (13)) that the interface has a martensitic-type structure. Examples are the type II ordering of AuCu alloys[9] and the precipitation of $VH_{0.45}$ from a V-H solid solution.[87] In other cases, some of the predicted parameters agree fairly well with the experimental measurements, but others do not. This may indicate that the crystallography is not that predicted by the PTMC and that the interface migration is nonconservative (Sts. (11) and (22)), but it should be noted that the PTMC predictions are not readily reconciled with experimental data in some transformations that are quite clearly martensitic.

A division suggested between the formal and "derivative" requirements of the PTMC assigns the crystallography to the former and the interface structure to the latter category that also includes the absence of diffusion or changes in long-range order. The last two do indeed indicate a change of rate control, but with a different interface structure, the crystallography should not follow the formal requirements as already discussed. Much confusion derives from the belief that shear and diffusion are somehow incompatible. Apart from the large chemical effect on the velocity, there is no distinction at the atomic level between martensitic and nonmartensitic growth when the latter utilizes a martensitic-type, atom-conserving interface. Statements[65] such as "diffusional growth takes place by means of individual thermally activated diffusion jumps across the boundary level" and "growth by shear means . . . that individual atoms cross the interphase boundary by a glide-type process, much more nearly akin to plastic deformation . . ." do not have an exact meaning.

Measured shape changes are the most difficult parameters to fit to the predictions of the PTMC theory, and this is understandable in view of the possibility that the large stresses around even a thin plate-shaped particle may give rise to some plastic deformation that will accommodate (*i.e.*, conceal) the shape change. Simple examples in which the shape discontinuity was transferred to the matrix were discussed in Section VI; however, if the accommodating slip is less regular, the shape change is more diffuse and harder to measure.

Accommodating deformation along the face of a plate, rather than at its edge, produces an opposite tilt of a free surface and hence a "rooftop" appearance. This is often seen in "banded" products like Au-Cd or In-Tl where adjacent bands are product regions with almost the same habit plane but opposite shape changes. Rooftop structures may also be produced by slip, twinning, or kink band formation,[38] but some controversy has arisen from their appearance in some single-crystal products where the conventional descriptions are apparently invalid.[88,89] More complex relief effects often arise from the transformation of self accommodating groups of variants, especially in shape memory alloys.[90]

X. CONCLUSIONS

The statements are the main conclusions, but the following are some important assertions:

1. Coherent interfaces occur in wholly displacive (martensitic) and also in diffusional-displacive transformations.
2. Displacive transformations have crystallographies predicted by the PTMC that ensures that atom numbers are conserved during transformation; this also applies to the formation of bainites and ferrites at temperatures where only interstitial components are mobile.
3. Diffusional transformations may follow the same crystallography or may adopt interfaces that have some epitaxial character and require material transport as they migrate. As this may be a slow process, martensitic-type interfaces may be preferred, even with diffusion.
4. Ledges (steps) and interface dislocations are usually interchangeable terms for the same defects. Structural ledges are the naturally occurring steps or facets in a description of a high-index or irrational plane; growth ledges are equivalent in appropriate interfaces to twinning or transformation dislocations.
5. Coherent superledges have high energies and only minimum height ledges should form. Partly coherent superledges have small residual Burgers vectors and produce only small shape discontinuities.
6. Step growth in displacive transformations probably involves the continuous displacement of elementary steps without the need for an explicit kink mechanism.
7. Diffusional growth may be limited by the interdiffusion rates of various components, by matter transport (dislocation climb) if the interface is epitaxial, by step nucleation, or by interface processes.

ACKNOWLEDGMENTS

I wish to thank Professors J.P. Hirth and R.C. Pond and Dr. A.P. Sutton for useful discussion and correspondence. Professors Hirth and H.I. Aaronson kindly sent me advance copies of their articles, and Professor Aaronson also made what must surely be a record number of comments on the first draft of this article; I have to thank him sincerely for a lengthy and vigorous correspondence that ensued. Finally, I have to thank most warmly the Conference Organizers (especially Professors Aaronson and Wayman) for generous financial assistance and for their patience and persistence in inviting me (repeatedly) to participate in this enjoyable and friendly conference.

REFERENCES

1. Quoted by J. Horgan: *Sci. Am.*, 1992, (Dec.), pp. 8-11.
2. F.C. Larché and J.W. Cahn: *Acta Metall.*, 1973, vol. 21, pp. 1051-61.
3. R.O. Williams: *Metall. Trans. A*, 1980, vol. 11A, pp. 247-53.
4. J.W. Cahn and F. Larché: *Acta Metall.*, 1984, vol. 32, pp. 1915-23.
5. W.C. Johnson: *Metall. Trans. A*, 1987, vol. 18A, pp. 1093-97.
6. W.C. Johnson and P.W. Vorhees: *Metall. Trans. A*, 1987, vol. 18A, pp. 1213-28; *Acta Metall.*, 1990, vol. 38, pp. 1183-89.
7. W.C. Johnson and W.H. Müller: *Acta Metall.*, 1991, vol. 39, pp. 89-103.
8. J.W. Christian: in *Martensite—Fundamentals and Technology*, E.R. Petty, ed., Longman, London, 1970, pp. 11-41.
9. R. Smith and J.S. Bowles: *Acta Metall.*, 1960, vol. 8, pp. 405-15.
10. E. Hornbogen and H. Warlimont: *Acta Metall.*, 1967, vol. 15, pp. 943-51.
11. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, pp. 845-48.
12. K. Chattopadhyay and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, pp. 695-711.
13. J. Cornelis and C.M. Wayman: *Acta Metall.*, 1974, vol. 22, pp. 301-11.
14. M.H. Wu and C.M. Wayman: *Proc. Int. Conf. on Martensitic Transformations*, Jpn. Inst. Met., 1987, pp. 619-24; M.H. Wu, J. Perkins, and C.M. Wayman: *Acta Metall.*, 1989, vol. 37, pp. 1821-37.
15. J.W. Christian: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY 1962, pp. 371-86.
16. A.P. Sutton and J.W. Christian: *ICOMAT-82, J. Phys*, 1982, vol. 43, Suppl. 12, C4. 97-C4. 102.
17. K.A. Bywater and J.W. Christian: *Phil. Mag.*, 1972, vol. 25, pp. 1249-73.
18. A. Serra and D.J. Bacon: *Phil. Mag.*, 1986, vol. 54A, pp. 793-804.
19. V. Vitek: *Scripta Metall.*, 1970, vol. 4, pp. 725-32.
20. J.W. Christian: *Metall. Trans. A*, 1982, vol. 13A, pp. 509-38.
21. J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 799-803.
22. M.S. Wechsler, D.S. Liebermann, and T.A. Read: *Trans. AIME*, 1953, vol. 197, pp. 1503-15.
23. J.S. Bowles and J.K. MacKenzie: *Acta Metall.*, 1954, vol. 2, pp. 129-37, 138-47, 224-34.
24. J.P. Hirth: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1885-94.
25. A.G. Khachaturyan: *Sov. Phys. Solid State*, 1969, vol. 8, pp. 2163-67.
26. A.L. Roitburd: *Solid State Phys.*, 1978, vol. 33, pp. 317-90.
27. J.D. Eshelby: *Proc. R. Soc. A*, 1957, vol. 241, pp. 376-96; 1959, vol. 252, pp. 561-69.
28. J.D. Eshelby: *Prog. Solid Mech.*, 1961, vol. 2, pp. 89-140.
29. H.I. Aaronson: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, pp. 387-546.
30. L.J. Walpole: *Proc. R. Soc. A*, 1967, vol. 300, pp. 270-89.
31. J.K. Lee, D.M. Barnett, and H.I. Aaronson: *Metall. Trans. A*, 1977, vol. 8A, pp. 963-70.
32. A.G. Khachaturyan: *Theory of Structural Transformations in Solids*, Wiley, New York, NY, 1983.
33. A.G. Khachaturyan and G.A. Shatalov: *Sov. Phys. JETP*, 1969, vol. 29, p. 557.
34. J.W. Christian: *Int. Symp. on New Aspects of Martensitic Transformation*, Jpn. Inst. Met., 1976, pp. 21-33.
35. J.W. Christian and J.V. Edmonds: *Phase Transformations in Ferrous Alloys*, TMS-AIME, Warrendale, PA, 1984, pp. 293-325.
36. J.W. Matthews: in *Dislocations in Solids*, F.R.N. Nabarro, ed., North-Holland, Amsterdam, 1980, vol. 2, pp. 461-545.
37. C.M. Sargent and G.R. Purdy: *Phil. Mag.*, 1975, vol. 32, pp. 27-35.
38. J.W. Christian: *The Theory of Transformations in Metals and Alloys*, 1st ed., Ch. XX, Pergamon Press, Oxford, 1965; 2nd ed., Part 1, 1975.
39. J.W. Christian and A.G. Crocker: in *Dislocations in Solids*, F.R.N. Nabarro, ed., North-Holland, Amsterdam, 1980, vol. 3, pp. 165-249.
40. B.A. Bilby, R. Bullough, and E. Smith: *Proc. R. Soc. A*, 1955, vol. 231, pp. 263-73.
41. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer-Verlag, Berlin, 1970.
42. R. Bullough and B.A. Bilby: *Proc. Phys. Soc.*, 1956, vol. 69B, pp. 1276-86.
43. J.W. Christian and K.M. Knowles: *Proc. Int. Conf. on Solid-Solid Phase Transformations*, TMS-AIME, Warrendale, PA, 1982, pp. 1185-1207.
44. G.B. Olson and M. Cohen: *Acta Metall.*, 1979, vol. 27, pp. 1907-18.
45. G.B. Olson: *Acta Metall.*, 1981, vol. 29, 1475-84.
46. R.W. Balluffi and G.B. Olson: *Metall. Trans. A*, 1985, vol. 16A, pp. 529-41.
47. R. Bonnet and F. Durand: *Phil. Mag.*, 1975, vol. 32, pp. 997-1006.
48. R. Bonnet: *Acta Metall.*, 1981, vol. 29, pp. 437-45.
49. R. Bonnet: *Phil. Mag.*, 1981, vol. A44, pp. 625-42.
50. M. Dupeux: *Phil. Mag. Lett.*, 1987, vol. 55, pp. 7-13.
51. B.A. Bilby: *Rep. Conf. on Defects in Crystn. Solids*, The Physics Society, London, 1955, pp. 123-33.
52. J.F. Nye: *Acta Metall.*, 1953, vol. 1, pp. 153-62.
53. K.V. Vladimiriskij: *Zh. Eksp. Teor. Fiz.*, 1947, vol. 17, p. 530.
54. F.C. Frank and J.H. Van de Merwe: *Proc. R. Soc. A*, 1949, vol. 198, pp. 205-16.
55. B.A. Bilby: *Phil. Mag.*, 1953, vol. 44, pp. 782-85.
56. P.D. Bristowe and A.G. Crocker: *Phil. Mag.*, 1975, vol. 31, pp. 503-17.
57. M. Yamaguchi and V. Vitek: *Phil. Mag.*, 1976, vol. 34, pp. 1-11.
58. R.C. Pond and D.S. Vlachauas: *Proc. R. Soc. A*, 1983, vol. 386, p. 95-143.
59. R.C. Pond: *Dislocations and Properties of Real Materials*, Inst. Met. London, 1985, pp. 71-93.
60. R.C. Pond: in *Dislocations in Solids*, F.R.N. Nabarro, ed., North-Holland, Amsterdam, 198, vol. 8, pp. 1-93.
61. R.C. Pond and J.P. Hirth: *Solid State Phys.*, 1994, vol. 47, in press.
62. A.W. Sleeswyk: *Acta Metall.*, 1962, vol. 10, pp. 803-12.
63. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1988, vol. 22, pp. 567-72, 575-76.
64. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1988, vol. 22, pp. 573-74, 577-80.
65. H.I. Aaronson, T. Furuhara, J.M. Rigsbee, W.T. Reynolds, Jr., and M.J. Howe: *Metall. Trans. A*, 1990, vol. 21A, pp. 2369-2409.
66. H.I. Aaronson, *Metall. Trans. A*, 1993, vol. 24A, pp. 241-76.
67. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1989, vol. 23, pp. 279-84.
68. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1989, vol. 23, pp. 285-90.
69. M.G. Hall, H.I. Aaronson, and K.R. Kinsman: *Surface Sci.*, 1972, vol. 25, pp. 257-74.
70. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 351-76.

71. M.G. Hall, J.M. Rigsbee, and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, pp. 1419-31.
72. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, pp. 1029-34.
73. J.H. Van der Merwe, G.J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, pp. 1165-75.
74. J.M. Howe and D.A. Smith: *Acta Metall. Mater.*, 1991, vol. 40, pp. 2343-50.
75. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-67.
76. T. Furuhara, J.M. Howe, and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2873-86.
77. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2887-99.
78. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2857-72.
79. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 639-48, 649-58.
80. J.M. Howe, U. Dahmen, and R. Gronsky: *Phil. Mag.* A, 1987, vol. 56, pp. 31-61.
81. H.I. Aaronson, C. Laird, and K.R. Kinsman: *Phase Transformations*, ASM, Metals Park, OH, 1970, pp. 313-96.
82. J.P. Hirth and R.W. Balluffi: *Acta Metall.*, 1973, vol. 21, pp. 929-42.
83. G.C. Weatherley: *Phil. Mag.*, 1968, vol. 17, pp. 791-99.
84. L.M. Brown and G. Woolhouse: *Phil. Mag.*, 1970, vol. 21, pp. 329-45.
85. R. Beanland: *Phil. Mag.*, 1993, vol. 67, pp. 585-603.
86. C. Laird and H.I. Aaronson: *Acta Metall.*, 1977, vol. 15, pp. 73-103; 1969, vol. 17, pp. 505-519.
87. J.S. Bowles, B.C. Muddle, and C.M. Wayman: *Acta Metall.*, 1977, vol. 25, pp. 513-20.
88. M.G. Hall: *Metall. Trans. A*, 1994, vol. 25A, in press.
89. H.I. Aaronson and M.G. Hall: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1797-1819.
90. T. Saburi, S. Nenno, and C.M. Wayman: *Proc. Int. Conf. on Martensitic Transformations, ICOMAT 1979*, Cambridge, MA, 1979, pp. 619-32.
91. J.M. Howe, T. Abinandanan, C.S. Chang, T. Ruruhara, N. Prabhu, and H.I. Aaronson: *Scripta Metall.*, 1987, vol. 21, pp. 1639-43.
92. Y. Mou: Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, PA, 1992.
93. C. Hitzenberger and H.P. Karnthaler: *Phil. Mag.*, 1991, vol. 64A, pp. 151-163.
94. P. Gaunt and J.W. Christian: *Acta Metall.*, 1959, vol. 7, pp. 529-533.

Application of the Theory of Martensite Crystallography to Displacive Phase Transformations in Substitutional Nonferrous Alloys

B.C. MUDDLE, J.F. NIE, and G.R. HUGO

It has been demonstrated that the theory of martensite crystallography is capable of accounting successfully for the form and crystallography of a range of plate- or lath-shaped transformation products, even when the formation of the product phase involves significant substitutional diffusion. These transformations include the precipitation of metastable hexagonal γ' (Ag_2Al) plates in disordered face-centered cubic (fcc) solid-solution Al-Ag alloys, the formation of ordered AuCu II plates from disordered fcc solid solution in equiatomic Au-Cu alloys, and the formation of metastable $9R$ α_1 plates in ordered ($B2$) Cu-Zn and Ag-Cd alloys. The application of the theory to these transformations is reviewed critically and the features common to them identified. It is confirmed that, in all three transformations, the product phase produces relief at a free surface consistent with an invariant plane-strain shape change and that the transformations are thus properly described as displacive. The agreement between experimental observations and theoretical predictions of the transformation crystallography is in all cases excellent. It is proposed that successful application of the theory implies a growth mechanism in which the coherent or semicoherent, planar interface between parent and product phases maintains its structural identity during migration and that growth proceeds atom by atom in a manner consistent with the maintenance of a correspondence of lattice sites. In the case of the coherent, planar interfaces associated with γ' precipitate plates in Al-Ag alloys, there is direct experimental evidence that this is accomplished by the motion of transformation dislocations across the coherent broad faces of the precipitate plates; the transformation dislocations define steps that are two atom layers in height normal to the habit plane and have a Burgers vector at least approximately equivalent to an $(a/6)\langle 112\rangle$ Shockley partial dislocation in the parent fcc structure. However, for AuCu II plates, where the product phase is twinned on a fine scale, and for α_1 plates, for which the lattice invariant strain leads to a substructure of finely spaced stacking faults, the structures of the semicoherent interphase boundaries and thus the details of the transformation mechanism remain less clearly defined.

I. INTRODUCTION

IN the approximately 40 years since the original formulations[1,2] of the crystallographic theory of martensitic transformations first appeared, the theory[3] has been applied successfully to classically diffusionless martensitic transformations in a wide range of metallic[4,5,6] and ceramic systems.[7] The theory has not yet proved to be of universal applicability to such transformations. However, the large number of transformations to which it has been shown to be applicable, and the accuracy with which it has proved capable of accounting for the crystallographic features of these transformations, have led to general acceptance of the theory as providing a reliable basis for predicting transformation crystallography. During that same period, the theory has also been applied to a number of transformations for which the transformation product exhibits morphology and crystallography typical of a martensitic product, but which involve significant long-range diffusion or atomic interchange over distances exceeding unit cell dimensions. These transformations include the formation of Widmanstätten[8] and bainitic ferrite[9–12] in steels, and the precipitation of a range of hydride,[13,14] nitride,[15] and oxide[16,17] products, where it is conveniently assumed that the displacive component of the transformation involves the substitutional constituents alone and that the compositional change accompanying transformation is accomplished by essentially independent interstitial diffusion.

However, there are also members of this class of transformations which involve a composition change or ordering reaction that inevitably requires vacancy diffusion and for which it is thus more difficult to reconcile the displacive and diffusional components of transformation. It is the purpose of the present article to review the applicability of the formal crystallographic theory to selected transformations in this latter category and to examine critically the implications of successful application of the theory for the mechanism of transformation.

B.C. MUDDLE, Reader in Materials Engineering, is with the Department of Materials Engineering, Monash University, Clayton, Victoria, 3168 Australia. J.F. NIE, formerly with the Department of Materials Engineering, Monash University, is Research Fellow with the Department of Mining and Metallurgical Engineering, University of Queensland, St. Lucia, Queensland, 4072 Australia. G.R. HUGO, formerly with the Department of Materials Engineering, Monash University, is Research Scientist with Aeronautical Research Laboratory, Defence Science and Technology Organisation, Fishermens Bend, Victoria, 3207 Australia.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

The three transformations examined are those that have attracted by far the most attention and controversy: the face-centered cubic (fcc) to hexagonal close-packed (hcp) transformation associated with the precipitation of metastable γ' phase in Al-Ag alloys,[18–23] the formation of ordered orthorhombic CuAu II plates from disordered fcc solid solution in equiatomic Cu-Au,[24–30] and the $B2$ to $9R$ transformation associated with the formation of metastable α_1 plates in Cu-Zn(-X)[31–39] and Ag-Cd alloys.[40,41]

II. CRYSTALLOGRAPHIC THEORY

The initial step in the application of the crystallographic theory to any transformation involves the identification of the correspondence between parent and product lattices that is characteristic of martensitic transformations and which relates uniquely vectors, planes, and unit cells of the parent with corresponding vectors, planes, and cells in the product phase. It is the maintenance of this lattice correspondence that allows the change in structure to be described by a homogeneous lattice deformation, **B**. When combined with an appropriate rigid body rotation, **R**, this homogeneous strain defines a total lattice strain, $\mathbf{S_T} = \mathbf{R} \cdot \mathbf{B}$, that generates the product lattice in its observed orientation relationship with the parent phase. For a given transformation, it is commonly assumed[42] that the operative lattice correspondence will be that which involves the minimum principal deformations (η_i) for the strain **B**, and this correspondence may often be identified simply by inspection of the crystal structures of parent and product phases and the observed orientation relationship between them. The homogeneous strain **B** is completely defined by the assumed lattice correspondence and the lattice parameters of parent and product lattices.

The total lattice strain $\mathbf{S_T}$ is rarely such that it leaves a plane of the parent structure undistorted and unrotated in the product, and, for most transformations, it is incompatible with a plate-shaped transformation product with an invariant habit plane and a macroscopic shape deformation that is an invariant plane strain. This discrepancy is resolved by considering the shape strain $\mathbf{S_R}$ to be the product of the total lattice strain $\mathbf{S_T}$ and a second strain **L**, which is lattice invariant and which occurs inhomogeneously throughout the transformed volume so as to maintain an interface plane that is macroscopically undistorted, *i.e.*, $\mathbf{S_R} = \mathbf{S_T} \cdot \mathbf{L}$. In the simplest form of the theory, the lattice invariant strain **L** is assumed to be the product of slip or a partial twinning shear in parent or product lattices. Alternatively, the total lattice strain $\mathbf{S_T}$ may be represented as the product of consecutive invariant plane strains, *i.e.*, $\mathbf{S_T} = \mathbf{S_R} \cdot \mathbf{L}^{-1}$, and is thus an invariant line strain, with the invariant line parallel to the line of intersection of the habit plane and the plane of lattice invariant shear (LIS). Given as input the lattice correspondence and lattice parameters of initial and final lattices, and the plane and direction of LIS, the theory permits calculation of the magnitude, plane, and direction of the shape strain $\mathbf{S_R}$, the magnitude of the LIS, and the predicted orientation relationship between lattices.

Solutions to the theory can exist only if the homogeneous lattice deformation **B** is such that at least one of the principal deformations is less than unity and one greater than unity. If, in addition, the remaining principal deformation is exactly unity, then $\mathbf{S_T}$ is itself an invariant plane strain and there is no requirement for an LIS. In such a transformation, the two lattices related by **B** may be fully coherent across a planar interface. The classic example of such a transformation is the fcc (f) to hcp (h) transition occurring in pure Co,[43] certain Co-based alloys,[44,45,46] and a range of highly alloyed steels of relatively low stacking fault energy.[6,47] In this case, the interatomic spacings within parallel close-packed planes are very nearly identical in the fcc and hcp structures, and the change in structure may be simply accomplished by a shear, equivalent to one-half the twinning shear, between alternate pairs of $\{111\}_f$ planes in any one of the three $\langle 112\rangle_f$ twinning directions in a given $\{111\}_f$ plane.[48] The magnitude of this shear is $(8^{-1/2})$* and, if the shear occurs on a single shear system, then the shear component of the shape deformation is appreciable (~0.35). If, however, shear occurs in all three $\langle 112\rangle_f$ directions in a given $\{111\}_f$ habit plane, then the transformation may be partially or wholly self-accommodating and the shear component of the shape strain may vary from a maximum of $8^{-1/2}$ to zero. In the latter extreme, the shape deformation $\mathbf{S_R}$ may be considered to be the average of equal volumes of the three possible lattice deformations such that $\mathbf{S_R} = \mathbf{I}$, where **I** is the identity matrix.[49]

*This value is based on the assumption that the two structures are perfectly coherent in the habit plane and that the hcp structure has the ideal axial ratio of $(8/3)^{1/2}$.

The glissile motion of a rational, coherent interface as an entity is energetically unfavorable, and it is commonly assumed that transformations for which the interface is fully coherent proceed *via* the formation and migration of transformation dislocations across the otherwise planar interface (Figure 1). The transformation dislocations constitute atomic scale steps on the interface and, following Christian,[48,49,50] have an effective Burgers vector of

$$\mathbf{b_T} = hm\mathbf{d} = h\tau\mathbf{t} + h\xi\mathbf{n} \qquad [1]$$

where h is the step height, m is the magnitude of the shape strain, **d** is a unit vector defining the direction of the shape strain, **t** is a unit vector in the direction of the shear component of the strain, **n** is a unit vector normal to the habit plane, and τ and ξ are the magnitudes of the shear and normal components of displacement, respectively. In general, the Burgers vector is not a lattice vector in either parent or product lattices. In the specific case of the fcc to hcp transformation, the minimum step height corresponds to the spacing of two close-packed planes normal to the coherent interface plane, and a transformation dislocation that will accomplish the required structural adjustment in the packing of the close-packed planes has the following Burgers vector:[49]

$$\mathbf{b_T} = \frac{a}{6}\langle 112\rangle + \frac{2\xi a}{3}\langle 111\rangle \qquad [2]$$

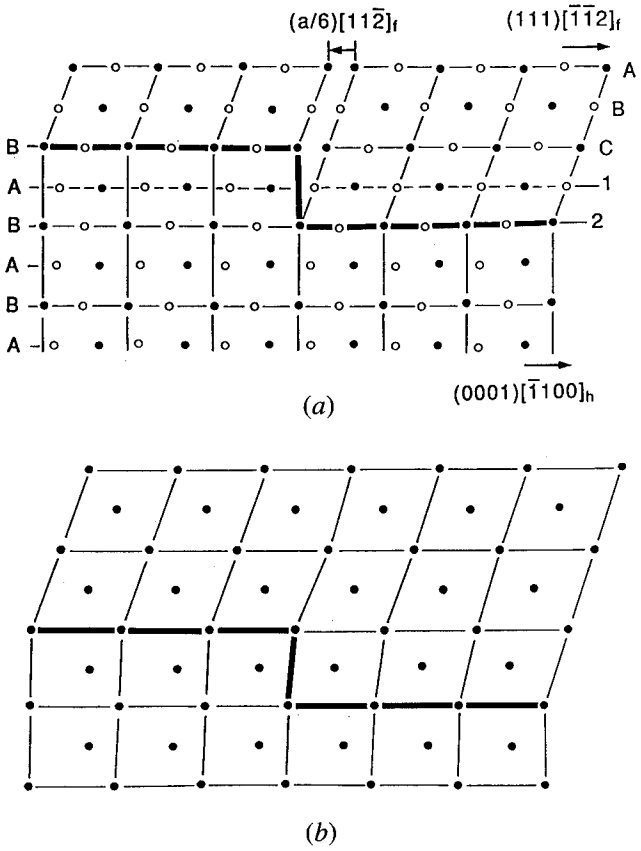


Fig. 1—(*a*) The $(111)_f \| (0001)_h$ interface between fcc and hcp structures containing a step that is two atom layers in height normal to the interface. Open and filled circles represent atoms on successive planes of the type $(2\bar{2}0)_f$ or $(11\bar{2}0)_h$. Note that planes 1 and 2 are common to both structures. (*b*) Schematic representation of the resulting transformation dislocation with Burgers vector $(a/6)[11\bar{2}]$.

This transformation dislocation reduces to a simple Shockley partial dislocation of the fcc structure if there is no change in the spacing of the close-packed planes normal to the habit plane (*i.e.*, $\xi = 0$). In the schematic representation of the fcc to hcp transition shown in Figure 1, the transformation dislocation has the specific Burgers vector $(a/6)[11\bar{2}]$.

In those transformations for which the total lattice strain $\mathbf{S}_T$ does not itself allow a plane between parent and product lattices to remain invariant, the crystallographic theory allows for the total strain to be combined with a lattice invariant strain $\mathbf{L}$ to define a habit plane which is macroscopically invariant and across which there is sufficient structural continuity to permit rapid glissile motion of the interface. Two simple models of a semicoherent martensitic interface, compatible with the formal requirements of the theory, have emerged.[5,48] In the first, the LIS is assumed to involve slip in either parent or product lattices, and the misfit in the interphase boundary is accommodated discontinuously by a single array of parallel interface dislocations with a common Burgers vector and a spacing determined by the magnitude of the LIS required to preserve a planar interface. The dislocation lines lie parallel to the intersection of the LIS plane and the habit plane, and are thus parallel to the invariant line $\mathbf{l}$ of $\mathbf{S}_T$. The interface is glissile and may move readily as an entity normal to itself, provided that the individual dislocations may move conservatively with the interface on slip planes in parent and product lattices that meet edge-on in the interface. This will be the case provided that the Burgers vector of the dislocations is either not parallel to the interface or, in the case of pure screw dislocations, is parallel to the invariant line $\mathbf{l}$.

In the second of the models of a semicoherent martensitic interface, the mismatch at the interface is corrected periodically by what is normally fine-scale twinning of the product phase, and the LIS corresponds to a partial twinning shear of the product lattice. In this case, the accommodation within the habit plane extends over the widths of individual twins, and the theory determines the relative widths of the twin-related crystals required to achieve a habit plane that is macroscopically invariant. If the fractions of the two twin orientations are f_t and $(1 - f_t)$, then there are two solutions for the equivalent lattice invariant shear with magnitudes:

$$g_1 = f_t s \quad \text{and} \quad g_2 = (1 - f_t)s \qquad [3]$$

where s is the magnitude of the twinning shear on the appropriate system in the product lattice. In the simplest form of this model, the segments of interface between each of the twin orientations and the parent phase are assumed to be in a state of forced elastic coherence and the interface is dislocation-free.[50] The elastic strain field associated with the interface has a periodicity equivalent to the average combined width of adjacent twin pairs, which is to be compared with the spacing of the interface dislocations in the dislocation model discussed previously. The interface is readily glissile, and growth of the product phase is assumed to involve motion of the interface as an entity normal to itself.

These models of coherent and semicoherent martensitic interfaces are such as to permit a correspondence of lattice sites to be preserved between parent and product lattices, and, provided that this correspondence is not destroyed by simultaneous diffusion, they imply a shape change over the transformed volume that may be described by an invariant plane strain. The assumption of a lattice correspondence is intrinsic to the application of formal crystallographic theory, and successful application of the theory, together with the observation of an invariant plane-strain shape change, would seem to imply the maintenance of a lattice correspondence. It is to be emphasized that, being phenomenological in nature, the theory requires a correspondence of equivalent lattice sites and not necessarily the maintenance of a correspondence of atoms occupying equivalent positions. It has been commonly assumed, however, that any significant atomic interchange over distances typically larger than the dimensions of the unit cell would destroy the correspondence of lattice sites and any associated shape change.[51] This argument is usually justified on thermodynamic grounds; the additional strain energy that would necessarily accompany a change of shape is considered likely to provide a driving force for the elimination of the correspondence under conditions permitting rapid atomic interchange.[51] A potential exception to this proposal has been identified in those cases, such as the formation of Widmanstätten or bainitic ferrite, where a change in composition is achieved through diffusion of

a highly mobile interstitial constituent, while the displacive component of the transformation is sustained by a constituent for which the range of atom movements remain less than the scale of the unit cell. However, it has long been considered unlikely that a shape change would be sustained and that the crystallographic theory would be proved applicable for those transformations that involve significant long-range substitutional (*i.e.*, vacancy) diffusion. The transformations reviewed here appear to constitute notable exceptions to this expectation and thus continue to attract significant attention.

III. PRECIPITATION OF γ′ PHASE IN Al-Ag ALLOYS

A. *Experimental Observations*

In Al-Ag alloys containing typically 15 to 20 wt pct Ag, solution-treated (550 °C, 0.5 hour), quenched, and aged (350 °C to 400 °C, 0.5 hour), the supersaturation of Ag in fcc solid solution is relieved initially by the precipitation of a metastable γ' phase and eventually by formation of the equilibrium γ (Ag_2Al) precipitate phase. The transition from γ' to γ phase occurs directly within individual precipitate particles and is described[19] as being associated with a loss of coherency of the γ' particles. The coherent γ' and semicoherent γ phases are both hcp, differ by only small changes in lattice parameters,[19,22] and are thus difficult to distinguish by most conventional diffraction techniques. However, most of the very extensive and detailed experimental work that has been carried out on the system[18–23] has been concerned with the initial stages of precipitation. The present discussion will thus focus on the formation of the γ' phase.

The crystal structure that has been proposed[23] for the γ' phase is, as shown in Figure 2, hcp (a = 0.2858 nm, c = 0.4607 nm, c/a = 1.612), with alternating close-packed A and B layers of pure Ag and Al_2Ag, respectively, to give an average precipitate composition of Ag_2Al. Although yet to be confirmed, this structure is qualitatively consistent with detailed contrast analysis of high-resolution images of γ' plates, and electron diffraction evidence confirms that the phase exhibits long-range order on alternate basal planes during at least the initial stages of growth. The γ' phase precipitates as thin, hexagonal-shaped plates with large aspect ratios (10:1 to 100:1) on $\{111\}_\alpha$ planes of the α-Al matrix (a_o = 0.4059 nm). The orientation relationship between parent and product structures is such that $(0001)_\gamma \| (111)_\alpha$ and $[11\bar{2}0]_\gamma \| [1\bar{1}0]_\alpha$. The two lattices are almost perfectly coherent within the $(0001)_\gamma \| (111)_\alpha$ habit plane, and there is a difference of only ~1.7 pct[23] in the spacing of the close-packed planes normal to the habit plane.

The transformation associated with the formation of γ' plates is thus structurally similar to the martensitic fcc to hcp transformation in Co and Co-based alloys.[43–46,48–50] However, there can be no question that there is a substantial composition change involving long-range substitutional diffusion of Ag associated with the formation of the γ' plates. Plates of the γ' phase have, for example, been extracted from the α-Al matrix and analyzed using energy-dispersive X-ray spectroscopy, and the composition has been confirmed to be consistent with the compound Ag_2Al to within ~4 at. pct.[23] Despite this composition change, there is ample evidence that there may be a shape change associated with the formation of γ' and γ plates, and that those plates formed at a free surface produce a surface relief consistent with a shape deformation that is an invariant plane strain.[19,21] Measurements of the surface tilt for single plates have been reported to be consistent with a simple shear equivalent to one-half of the twinning shear (0.354) in the hcp lattice.[21]

B. *Application of the Theory*

The structural change from fcc to hcp associated with the precipitation of the γ' phase is readily interpreted in terms of the formal crystallographic theory.[48,49,50] Based on the lattice parameters discussed previously, the misfit in interatomic spacings along parallel close-packed directions within the common close-packed planes is <0.5 pct. If it is assumed that the structures are perfectly coherent in the habit plane at the transformation temperature, and that there is no change in the spacing of the close-packed planes normal to the habit plane, then the structural change may be considered to be accomplished by a simple shear of magnitude $8^{-1/2}$ on the $(111)_f$ plane in the $[11\bar{2}]_f$ direction. In the formalism of the theory, the total strain $\mathbf{S_T}$ is an invariant plane strain equivalent in this case to the shape deformation $\mathbf{S_R}$:

$$\mathbf{S_T} = \mathbf{S_R} = \mathbf{I} + m\mathbf{dp}' \qquad [4]$$

where $\mathbf{I}$ is the identity matrix, m is the strain magnitude, $\mathbf{d}$ is a unit vector in the direction of displacement, and $\mathbf{p}'$ is a unit vector normal to the shear plane. Using a three-axis system to define the basis of the hcp lattice, the lattice correspondence between hcp and fcc lattices compatible with the observed orientation relationship is, as shown in Figure 3, of the form

$$ {}_f\mathbf{C}_h = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ \bar{1} & 0 & 1 \\ 1 & \bar{1} & 2 \end{bmatrix} \quad \text{and} \quad {}_h\mathbf{C}_f = \frac{1}{2}\begin{bmatrix} 1 & \bar{3} & 1 \\ 3 & \bar{1} & \bar{1} \\ 1 & 1 & 1 \end{bmatrix} \qquad [5]$$

The total strain may be expressed[48] as

$$\mathbf{S_T} = \frac{1}{12}\begin{bmatrix} 13 & 1 & 1 \\ 1 & 13 & 1 \\ \bar{2} & \bar{2} & 10 \end{bmatrix} \qquad [6]$$

For the present transformation, the lattice parameters imply a reduction of ~1.7 pct in the spacing of the close-packed planes parallel to the habit plane, and the displacement vector $m\mathbf{d}$ thus has a component normal to the habit plane. From Eq. [2],

$$m\mathbf{d} = \mathbf{b_T} = \frac{a}{6}[11\bar{2}] - \frac{2 \times 0.017a}{3}[111] \qquad [7]$$

and the total lattice strain

$$\mathbf{S_T} = \begin{bmatrix} 1.0364 & 0.0364 & 0.0364 \\ 0.0364 & 1.0364 & 0.0364 \\ -0.0808 & -0.0808 & 0.9192 \end{bmatrix} \qquad [8]$$

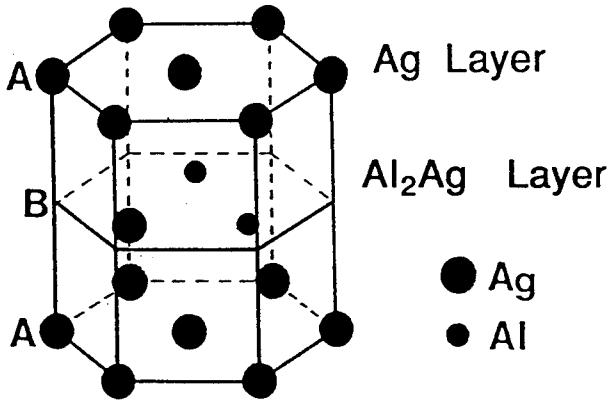


Fig. 2—Schematic representation of the crystal structure proposed for the γ phase (Ag_2Al) in Al-Ag alloys.[23]

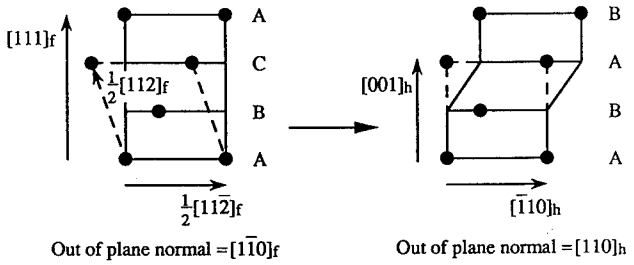


Fig. 3—Schematic representation of the lattice correspondence between parent and product lattices for the fcc to hcp transformation.

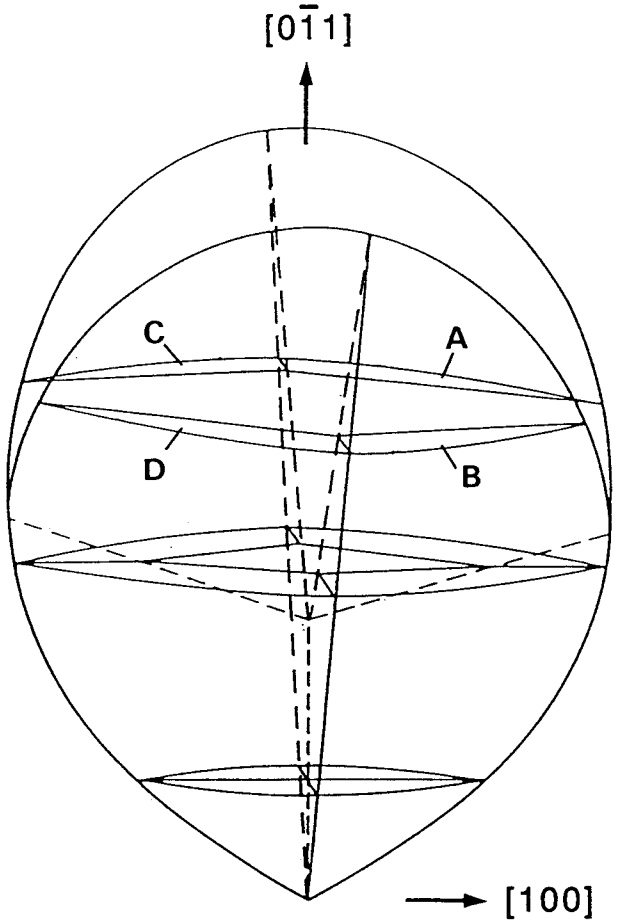


Fig. 4—Schematic representation of the morphology of AuCu II plates forming a pyramidal group of four.[29]

is an invariant plane strain rather than a simple shear. This total lattice strain $\mathbf{S}_T$ defines phenomenologically both the change in structure and the shape deformation.

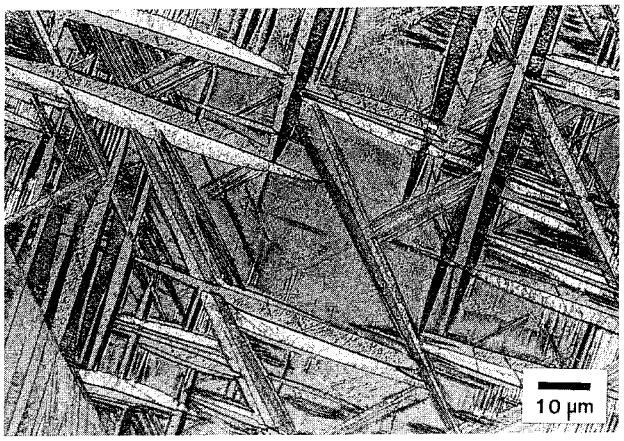
IV. FORMATION OF ORDERED AuCu II PLATES IN EQUIATOMIC Au-Cu

A. *Experimental Observations*

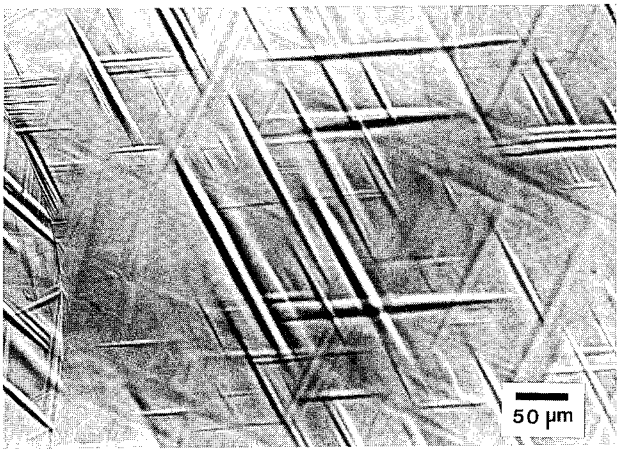
Equiatomic Au-Cu has a disordered fcc structure (a_c = 0.3872 nm) at temperatures above 410 °C and is ordered below this temperature. Below 380 °C, the product of ordering (AuCu I) is face-centered tetragonal (fct) $L1_o$ (a_t = 0.395 nm, c_t = 0.368 nm, c/a = 0.932),[25] with alternate layers of Au and Cu atoms parallel to the (001) basal plane. In the interval between 380 °C and 410 °C, a one-dimensional long-period superlattice designated AuCu II is formed. The orthorhombic unit cell of the AuCu II corresponds to an assembly of ten $L1_o$ unit cells stacked parallel to [010], with antiphase domain boundaries every five unit cells characterized by a displacement vector of the form $^1/_2$[101](010). Each $L1_o$ cell suffers a small additional distortion parallel to [010], and it has become common to describe the AuCu II structure in terms of the resulting orthorhombic (o) pseudocell with a_o = 0.3979 nm, b_o = 0.3963 nm, and c_o = 0.3678 nm.[25]

The ordered AuCu II phase forms characteristically as plates in pairs or pyramidal groups of four which, as shown schematically in Figure 4, share a common apex and grow cooperatively by motion of the inner apex parallel to the [0$\bar{1}$1] axis of the pyramid.[29] The forms that the plates are observed to take in section depends on the location and orientation of the plane of section with respect to the pyramid axis. The range of forms consistent with the pyramid morphology proposed in Figure 4 are shown in the reflected-light micrograph of Figure 5(a). At a free surface, the plates produce a well-defined surface relief (Figure 5(b)), which has been shown[25] to be consistent with an invariant plane-strain shape deformation. The inner and outer surfaces of the AuCu II plates are not parallel, leading to some uncertainty in the determination of the true habit plane. However, it is well established that the habit plane is irrational and close to {011}. Within a pyramidal group of four plates, the habit plane poles cluster close to a single (0$\bar{1}$1) plane, as shown schematically in Figure 4. The results of habit plane measurements from one such group, determined from trace measurements in two surfaces,[25] are reproduced in Figure 6, where it is clear that they are equivalent variants of a single irrational plane to within ±1 deg.

The orientation relationship between ordered AuCu II and disordered cubic (c) matrix is such that the principal axes of the ordered phase are approximately parallel to those of the matrix.[25,52] This is illustrated using electron diffraction in Figure 7, where the AuCu II plate in



(*a*)



(*b*)

Fig. 5—Reflected light micrographs of (*a*) polished and etched section showing AuCu II plates in an equiatomic Au-Cu alloy aged 0.75 h at 390 °C;[52] and (*b*) surface relief on a prepolished surface associated with growth of AuCu II plates produced under similar conditions.[52]

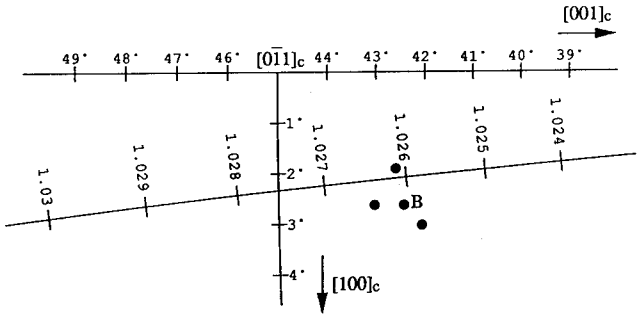


Fig. 6—Stereographic projection comparing the positions of measured habit plane normals for a pyramidal group of four AuCu II plates[29] with the locus of predicted habit planes plotted as a function of lattice parameter ratio a_o/a_c, where a_c refers to the fcc matrix and a_o refers to the orthorhombic product, for fixed values of the ratios $b_o/a_o = 0.9960$ and $c_o/a_o = 0.9244$. The measured habit planes have been converted by symmetry operations to that variant B consistent with the assumed variant of the LIS, *i.e.*, $(101)[\bar{1}01]_o$. Measured lattice parameters[25] give $a_o/a_c = 1.0276$.

Figure 7(a) is oriented with a $\langle 001\rangle_o$ direction parallel to the incident electron beam.[52] The accompanying selected-area electron diffraction (SAED) pattern (Figure 7(b)) includes reflections from the cubic matrix and from two $[001]_o$ variants of the ordered orthorhombic structure, twin-related about a $\{101\}_o$ plane. For the indexing adopted in Figure 7(c), the measured deviation of the $(001)_o$ plane from $(001)_c$ is on the order of 2 deg, and that between (100) planes is also approximately 2 deg. More accurate measurements of the angles between principal axes determined in previous work[25] using X-ray diffraction are reproduced in Figure 8 for a plate with habit plane variant equivalent to variant B in Figures 4 and 6.

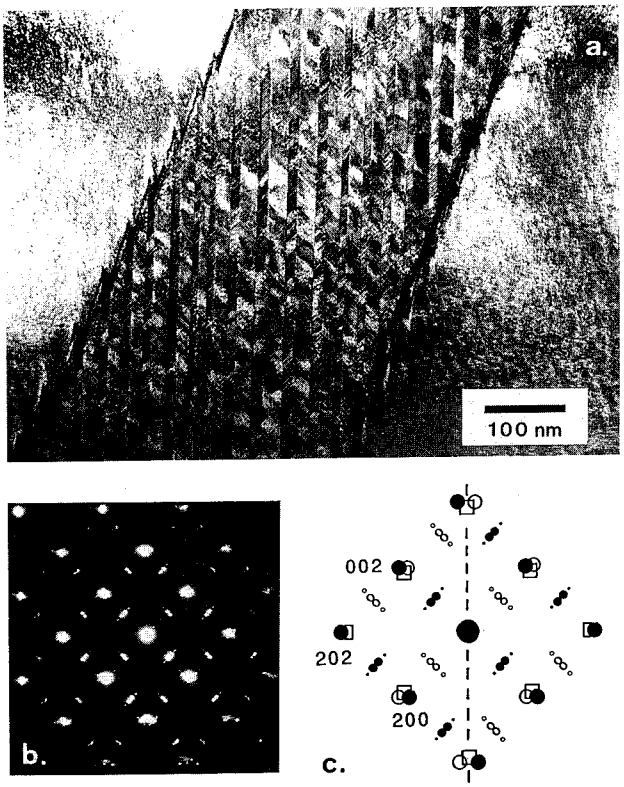


Fig. 7—(*a*) Transmission electron micrograph of single AuCu II plate recorded with the electron beam parallel to a $\langle 100\rangle$ fcc matrix direction; (*b*) corresponding SAED pattern; and (*c*) schematic indexed solution to (b) showing two $[010]_o$ variants of the orthorhombic product phase (open and filled circles) twin-related about $(101)_o$. Those reflections represented by open squares define an $[010]_c$ zone of the matrix phase.[52]

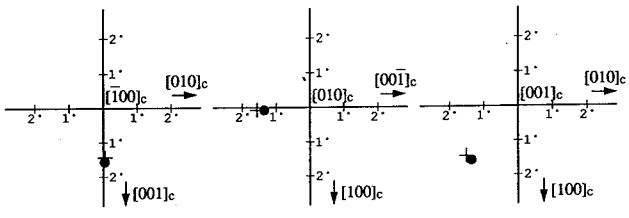


Fig. 8—Local sections of stereographic projections comparing measured[25] (crosses) and calculated (closed circles) positions of the principal axes of the orthorhombic unit cell of AuCu II plotted with respect to the principal axes of the parent fcc lattice. The variant of the orientation relationship shown corresponds to the variant B of the habit plane in Fig. 6.

As indicated in Figure 7, the AuCu II plates have a substructure of fine-scale twins on a $\{101\}_o$ twin plane. The pair of AuCu II plates shown in Figure 9 were imaged with the electron beam parallel to $\langle 111\rangle_c$ matrix and to corresponding $\langle 111\rangle_o$ axes in each of the plates. In this orientation, the twin planes in both plates and the junction plane between plates are parallel to the electron beam, and the habit planes of both plates are also approximately edge-on. Electron microdiffraction patterns recorded from the plates have been used to confirm[52] that each comprises twin-related variants of the ordered structure and that the variant that is in the lesser fraction is common to each. If the plates are interpreted, as indicated, as an A,B pair (*c.f.* Figure 4) and the electron beam is assumed to be parallel to $[\bar{1}11]_c$, then the twin plane for plate B is $(101)_o$ and the trace normal to the habit plane is observed to deviate by approximately 3 deg from $(0\bar{1}1)_o$ toward $(101)_o$. These electron microscope observations are thus consistent with earlier results,[25] which established that the twin plane corresponding to the habit plane variant B was $(101)_o$, derived from the corresponding $(101)_c$ plane of the parent phase.

When the $\{101\}_o$ twin plane is oriented parallel to the electron beam, as in Figure 9, it is possible to make accurate measurements of the relative thicknesses, x_1 and x_2, of the two twin-related orientations in the AuCu II plates. The results of measurements[52] of the average twin widths for a series of six different plates imaged under these conditions are recorded in Table I. Each value of twin width represents an average of at least 15 individual measurements on a given plate. The twin width ratio, x_2/x_1, has an average value of 0.49 ± 0.05.

B. *Application of the Theory*

The crystallographic theory has been applied[25] to the formation of AuCu II plates assuming that the correspondence between parent and product lattices is identity and that the LIS comprises a partial twinning shear on the system $\{101\}_o\langle\bar{1}01\rangle_o$. The predictions of the theory, for the specific variant of the correspondence

$$ {}_oC_c = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad [9] $$

and an LIS on the system $(101)[\bar{1}01]$, are presented in Table II. These predictions vary with a variation in the input lattice parameters, and a segment of the locus of predicted habit planes as a function of lattice parameter ratio for the low shear solution (2A) is shown in Figure 6. The curve passes close to $(0\bar{1}1)$ and, for the measured lattice parameter ratio of 1.0276, the predicted habit plane is within 3 deg of the measured habit plane variant B, for which the twin plane variant has been confirmed experimentally to be $(101)_o$ (Figure 9). In the original application of the theory, this difference between measured and predicted habit planes was interpreted[25] in terms of a dilatation parameter δ, which was considered a component of the shape strain and in essence relaxed the condition that the habit plane remains completely undistorted by allowing a small uniform dilatation of the habit plane. However, there is little or no direct evidence to support such a dilatation in this or other such transformations, and it appears likely that this difference is attributable to the difficulty of identifying the true habit plane for the tapered plates and to experimental errors in measurement.

For the previous low shear solution, the predicted shape strain direction is approximately parallel to $[011]_f$ and the magnitude of the shape deformation is ~0.047. Although there have been no direct measurements of the shape strain to date, the predicted shape strain has been used to calculate the surface tilts expected for CuAu II plates in certain orientations, and both the sense and magnitudes of these tilts have been found to be in excellent agreement with experimental observations.[25]

Aside from providing accurate predictions of the components of the shape strain consistent with the observed crystallographic variants, it is also to be emphasized that, by considering appropriate variants of the lattice correspondence and the LIS, the theory permits four variants of the low shear solution, for which the habit plane poles cluster about $(0\bar{1}1)_f$, in the manner observed experimentally. The theory is thus capable of accounting for the crystallography of each pyramidal group of four plates, for which the shape strain directions cluster about $[011]_f$. Each group of four may be considered to comprise two pairs in which the shear components of the shape strain are largely self-accommodating to minimize the net shape deformation.

As indicated in Table II, the principal axes of the product phase are predicted to be close to parallel to those of the cubic parent phase and, as shown in Figure 8, the agreement between the predicted and measured[25] orientation relationship for variant B is excellent. The twinned substructure of the AuCu II plates implies an LIS whose magnitude is determined by the relative proportions of the twin-related orientations within the plates. The magnitude g of the LIS has been

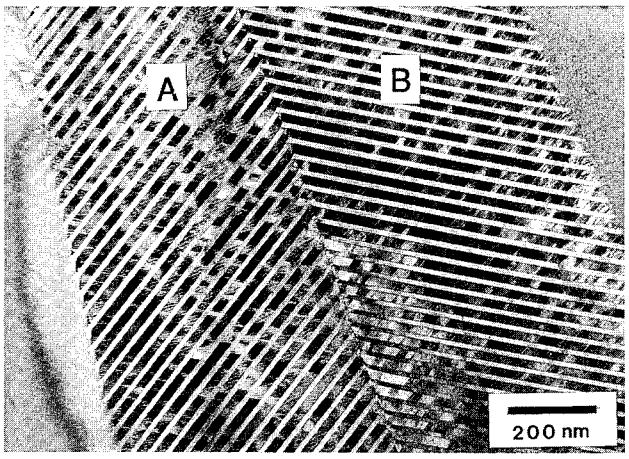


Fig. 9—Transmission electron micrograph showing a pair of AuCu II plates recorded with the electron beam parallel to a $\langle 111\rangle_c$ axis and to corresponding $\langle 111\rangle_o$ axes in each of the plates. The twin planes in both plates and the junction plane between plates are parallel to the electron beam, and the habit planes of both plates are also approximately edge-on.

Table I. Measurements of the Relative Twin Widths and the Magnitude (g) of Lattice Invariant Shear in AuCu II Plates[52]

Plate	x_1 (nm)	x_2 (nm)	x_2/x_1	g
1	12.71	6.32	0.497	0.052
2	21.32	11.32	0.531	0.055
3	30.87	13.64	0.442	0.048
4	21.83	10.87	0.498	0.052
5	21.13	11.37	0.538	0.055
6	24.25	10.73	0.442	0.048
Average:			0.49 ± 0.05	0.052 ± 0.005

Table II. Predictions of the Crystallographic Theory for the Disordered fcc to Ordered Orthorhombic Transformation Associated with the Formation of AuCu II Plates

Lattice Parameters: $a_c = 0.3872$ nm
$a_0 = 0.3979$ nm, $b_0 = 0.3963$ nm, $c_0 = 0.3678$ nm
Lattice Invariant Shear System: $(101)_0[\bar{1}01]_0$
Lattice Correspondence: Identity

Solution	g*	Habit Plane Normal Direction Cosines	Shape Strain Direction Cosines	m**	Orientation Relationship: Direction Cosines of Product Axes a_0	b_0	c_0
1A	0.1024	0.7044, −0.7087, 0.04104	−0.7209, −0.6918, −0.04200	0.04736	0.9984	0.02364	0.05121
					−0.02364	0.9997	−0.002587
					−0.05126	0.001377	0.9987
1B	0.1024	0.7044, 0.7087, 0.04104	−0.7209, 0.6918, −0.04200	0.04736	0.9984	−0.02364	0.05121
					0.02354	0.9997	0.002587
					−0.05126	−0.001377	0.9987
2A	0.0546	0.04104, −0.7087, 0.7044	−0.04200, −0.6918, −0.7209	0.04736	0.9996	0.001377	0.02730
					−0.000731	0.9997	−0.02367
					−0.02732	0.02364	0.9994
2B	0.0546	0.04104, 0.7087, 0.7044	−0.04200, 0.6918, −0.7209	0.04736	0.9996	−0.001377	0.02730
					−0.000731	0.9997	0.02367
					−0.02732	−0.02364	0.9994

*g is the magnitude of the lattice invariant strain.
**m is the magnitude of the shape strain.

calculated[52] from the measured twin widths x_1 and x_2 reported in Table I using the following expression:

$$g = \left(\frac{x_2}{x_1 + x_2}\right)s = f_2 s \qquad [10]$$

where s is the magnitude of the twinning shear and f_2 represents the volume fraction of orientation 2 in the orthorhombic product. For these lattice parameters, the magnitude of the twinning shear,[5] $s = [(a/c + c/a)^2 - 4]^{1/2}$ is 0.1575 and the resulting magnitudes of g in Table I vary between 0.048 and 0.055 for the six plates examined.[52] The average value for g over the six plates is 0.052 ± 0.005 and is thus in excellent agreement with the predicted value of 0.055 shown in Table II.

V. FORMATION OF α_1 PLATES IN Cu-Zn(-X) AND Ag-Cd ALLOYS

A. *Experimental Observations*

When body-centered cubic (bcc) β-phase alloys, based on the systems Cu-Al, Cu-Sn, Cu-Zn, and Ag-Cd, are quenched from temperatures at which the β phase is stable, there is invariably an ordering transition to a CsCl-type ($B2$) structure (β'). The β' phase is metastable and, on cooling below room temperature, undergoes martensitic transformation to one or more transformation products. In Cu-Zn alloys, two forms of martensite have been reported, one which inherits the ordered CuAu I structure (fct) from the matrix β' phase and has a substructure of fine-scale twinning, and a second needlelike martensite that has a faulted substructure and the $9R$ crystal structure, which may be regarded as a variant of the conventional close-packed fcc (f) structure ($3R$) with an $(a/6)\langle 112\rangle$ $\{111\}$ displacement on every third close-packed plane. However, if the β' phase in Cu-Zn alloys containing typically 38 to 44 wt pct Zn is reheated to temperatures in the two-phase ($\alpha + \beta'$) field, isothermal decomposition occurs and, at the lower end of the transformation temperature range, a metastable plate-shaped product designated α_1 is formed. The α_1 plates produce a surface relief consistent with an invariant plane-strain shape deformation[31,32,36] and have a morphology, structure, and crystallography closely similar to that of the needlelike $9R$ martensite formed athermally at low temperatures.[53,54]

Figure 10(a) shows the results of habit plane measurements for α_1 plates formed isothermally in two Cu-Zn alloys[36] and a Ag-Cd alloy,[40] plotted with respect to

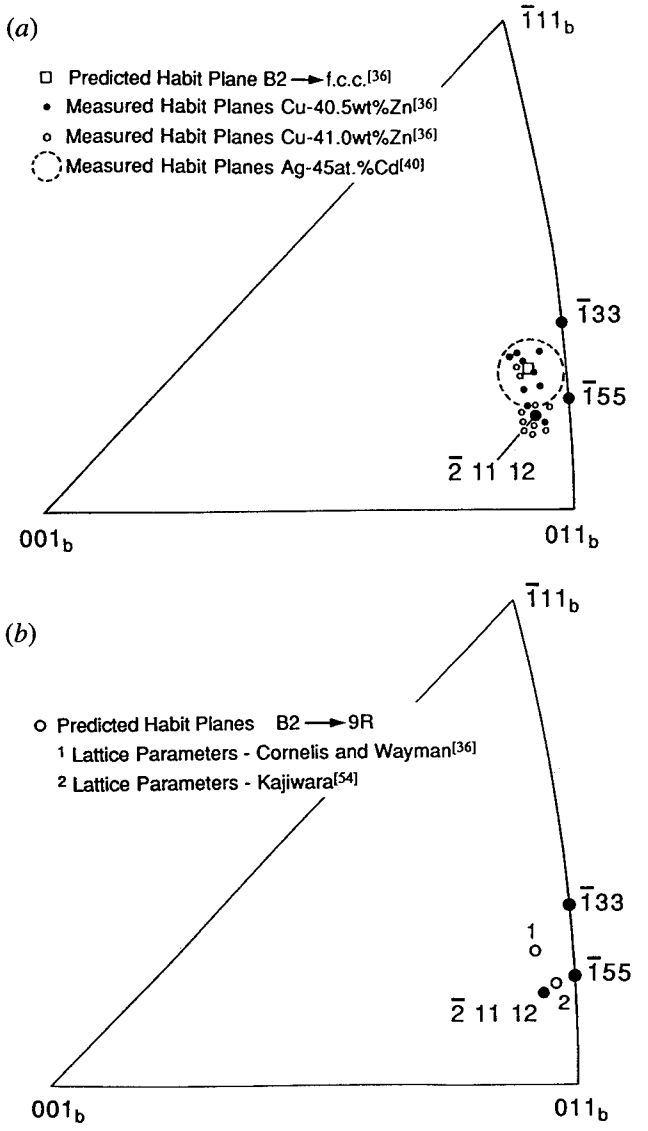


Fig. 10—Stereographic projections showing (*a*) measured habit plane normals for α_1 plates formed in two Cu-Zn alloys[36] and a Ag-Cd alloy[40] and (*b*) theoretical predictions of the habit plane for the *B*2 to 9*R* transformation.

the parent bcc (b) lattice in stereographic projection. The circle of radius 3.5 deg represents the scatter of two-thirds of individual measurements recorded for a Ag-45 at. pct Cd alloy,[40] while the individual measurements are reproduced for the Cu-Zn alloys. The habit plane is irrational and some 3.5 deg from $(\bar{1}44)_b$ or $(\bar{1}55)_b$; in the case of the Cu-41 wt pct Zn alloy, it is described[36] as being close to $(\bar{2},11,12)_b$. The large scatter in individual results reflects in part the limited accuracy of measurements made on what are very fine-scale plates, but it has also been suggested[36] that there is a significant variation in habit plane between the two Cu-Zn alloys due to a variation in lattice parameters. In both alloys, it is common for the α_1 plates to form in chevron pairs, with the habit plane poles of the pair crystallographically equivalent and symmetrically distributed with respect to the nearest $\{011\}_b$ pole. Given the fine scale of the plates, complete measurements of the shape strain have not been made, but the surface tilts associated with plates formed at free surfaces have been reported to range up to 11.75 deg in the case of Cu-Zn alloys[36] and 12.0 ± 1.0 deg for the Ag-Cd alloy.[40]

Figure 11(a) shows an electron micrograph of an α_1 plate formed during the early stages of transformation in a Ag-45 at. pct Cd alloy.[41] Given the limited scale of such plates, the orientation relationship between the β' and 9*R* α_1 structures has commonly been determined approximately using conventional SAED patterns such as that of Figures 11(b) and (c), in which a $\langle 110\rangle_{9R}$ zone axis is approximately parallel to a $\langle 111\rangle_b$ zone of the parent phase and the close-packed planes of the α_1 phase, $(009)_{9R}$, are parallel to the electron beam. The relationship that may be derived[36,38,40,41] from this and other zone axis patterns is irrational and is usually described approximately by the relationships between planes shown in Table III. The variant of the orientation relationship recorded in Table III corresponds to the variant of the habit plane shown in Figure 10.

The α_1 plates have a substructure of parallel stacking faults distributed regularly along the length of each plate (Figure 11). For the variant of the habit plane shown in Figure 10 and the orientation relationship reported earlier, it has been established that the fault plane is $(111)_f$, derived from a $\{101\}_b$ plane of mirror symmetry in the parent bcc phase. With extended isothermal aging, the faulted substructure observed in freshly formed plates is annealed from the plates, and this recovery is accompanied by a transition in structure from the metastable 9*R* to the equilibrium 3*R* packing of close-packed planes (α).[36,40,41] Local growth of the plates proceeds behind incoherent, faceted interfaces, with interface facets bearing no apparent relationship to the original habit plane.

B. *Application of the Theory*

The formal theory has been applied to the formation of α_1 plates in Cu-Zn[36] and Ag-Cd[40] alloys assuming a product structure that is fcc (3*R*) or fct ($c/a \sim 0.98$). The correspondence between bcc and fct lattices, identified as compatible with the observed orientation relationship, is equivalent to the reverse of the familiar Bain correspondence observed in association with the fcc to bct transformation in ferrous martensites.[4,42] When a strict correspondence of atom positions is preserved during transformation, this correspondence requires that the product phase remain ordered and acquire the fct CuAu I structure ($L1_0$). If the stacking faults within the plates are taken as evidence of the LIS, then the fault plane consistent with the variant of the correspondence is $(111)_f$ and, of the three $\langle 112\rangle_f$ directions in this plane, only a shear in the $[11\bar{2}]_f$ direction preserves the ordered arrangement in the product structure.[36] It has thus been assumed that the LIS comprises a shear on the system $(111)_f[11\bar{2}]_f$, which corresponds to a $\{110\}_b\langle 110\rangle_b$ shear system in the bcc matrix phase.

Based on these assumptions, the predictions of the theory for the components of the shape strain and the orientation relationship have been shown to be in good general agreement with experimental observations. The magnitude of the LIS predicted is in the range of 0.234[36] to 0.238[40] (corresponding to shear angles of

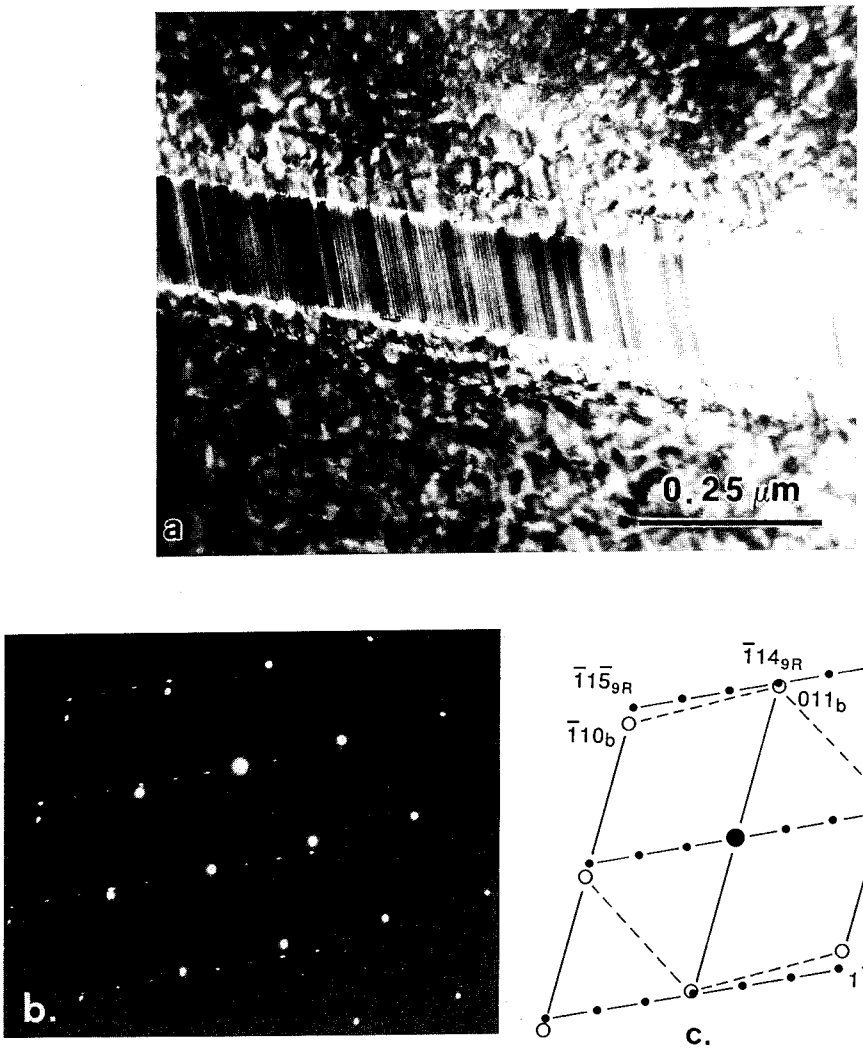


Fig. 11—(*a*) Transmission electron micrograph of α_1 plate in Ag-45 at. pct Cd alloy aged 30 s at 240 °C; (*b*) corresponding SAED pattern indicating approximately parallel $[\bar{1}\bar{1}1]_b$ and $[\bar{1}\bar{1}0]_{9R}$ zone axes from matrix and precipitate phases, respectively; and (*c*) schematic indexing of (b).

13.18 to 13.56 deg) and is almost exactly equivalent to the shear of $(a/6)[11\bar{2}]_f$ on every third $(111)_f$ plane that is required to generate the 9*R* stacking from a regular 3*R* fcc structure (*i.e.*, a shear of 13.26 deg). However, given the periodicity of these displacements, it is not expected that the modulation in stacking characteristic of the 9*R* structure will be resolvable using conventional electron imaging techniques. The stacking faults that are readily resolved within the α_1 plates are much more widely spaced and less regularly distributed than those that would be required to produce the 9*R* structure from conventional fcc, and this approach to application of the theory does not yield an obvious explanation of these faults. To account for faulting on the scale observed requires a more rigorous application of the theory, assuming that the product structure is 9*R* rather than fcc(t). Indeed, there is evidence to suggest that the 9*R* structure might properly be treated as a structure intermediate between the *B*2 parent and the equilibrium 3*R* phase, and that it is inappropriate to interpret the transformation as *B*2 to fcc 3*R*.

As indicated schematically in Figure 12, the 9*R* structure may be described by a unit cell that is orthorhombic if the structure is ideally close-packed and monoclinic (with a monoclinic β angle close to 90 deg) if there is a significant deviation from close packing normal to the close-packed planes.[54] The correspondence between parent bcc and product monoclinic (m) 9*R* lattices is such that the vectors $[10\bar{1}]_b$, $[010]_b$, and $[504]_b$ become the base vectors $[100]_m$, $[010]_m$, and $[001]_m$, respectively, in the product lattice. The LIS system corresponding to $(101)[10\bar{1}]_{bcc}$ and to $(111)[11\bar{2}]_{fct}$ is $(001)[100]_{9R}$. Reliable measurements of the lattice parameters of the isothermal 9*R* product are not readily available, but a range of parameters has been reported for the 9*R* martensite in Cu-Zn and Cu-Al systems.[36,54] Table IV presents a set of solutions to the crystallographic theory based on the above lattice correspondence and LIS, and a monoclinic 9*R* structure with a representative set of lattice parameters.

For the lattice parameters used in the compilation of Table IV, the predicted habit plane for low shear solution 1A, shown in Figure 10(b), is irrational. It lies within approximately 1.5 deg of $(\bar{1}55)_b$ and 1 deg of $(\bar{2},11,12)_b$ and is thus in good agreement with experimental observations of the habit plane for the Cu-41Zn

Table III. Comparison of Measured and Predicted Orientation Relationships for the *B2* to *9R* Transformation in Cu-Zn and Ag-Cd Alloys

	Measured Angles (deg)			Predicted Angles* (deg)
	Cu-Zn[36]	Ag-Cd[40]	Ag-Cd[41]	
$[110]_{9R} \wedge [\bar{1}\bar{1}1]_b$	~0.0	0.7 ± 0.2	<1.0	0.54
$(009)_{9R} \wedge (101)_b$	5.5	4.3 ± 0.4	5.0 ± 1.0	4.10
$(\bar{1}14)_{9R} \wedge (011)_b$	0.0 ± 0.5	—	<0.5	0.24
$(\bar{1}1\bar{5})_{9R} \wedge (\bar{1}10)_b$	5.5 ± 0.5	—	4.0 ± 1.0	3.90

*Calculated using the lattice parameters shown in Table IV.

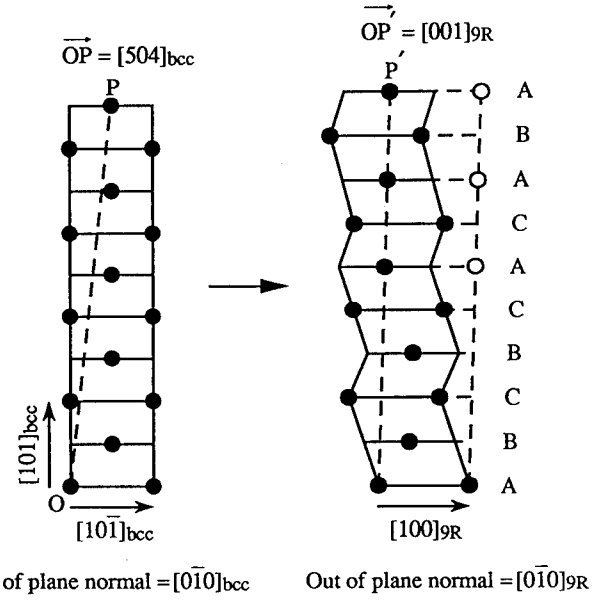


Fig. 12—Schematic representation of the structural transformation from *B2* to *9R*, indicating the lattice correspondence that is sustained.

alloy reproduced in Figure 10(a). The habit plane predicted for the bcc to *9R* transformation using the parameters of Cornelis and Wayman[36,53] (Figure 10(b)) is similar to that reported by those authors for the bcc to fct transformation and is in good agreement with experimental observations for the Cu-40.5Zn alloy in Figure 10(a). It is clear that the position of the predicted habit plane normal is sensitive to small variations in lattice parameters and that the theory is capable of accounting for the variation in the observed habit planes in Cu-Zn and Ag-Cd alloys, allowing for reasonable variations in the lattice parameters of the parent and product phases with varying alloy compositions.

For the habit plane variant shown in Figure 10 (low shear solution 1A), the predicted shape strain direction is also irrational, and approximately 8 deg from $[01\bar{1}]_b$ and within 2 deg of being parallel to the habit plane. The predicted shape strain displacement (m) is approximately 0.203. If an α_1 plate were to form at a free surface in such a way that the shear component of displacement was parallel to the surface normal, then the shape strain displacement would lead to a maximum surface tilt of ~11.48 deg. The predicted shape strain is thus capable of accounting well for observed surface tilts, which range up to a maximum of 11.75 deg and

Table IV. Predictions of the Crystallographic Theory for the *B2* (b) to Monoclinic (m) *9R* Transformation Associated with the Formation of α_1 Plates

Lattice Parameters: **Parent bcc: a_b = 0.2930 nm**
Product monoclinic *9R*: a_m = 0.4439 nm, b_m = 0.2638 nm, c_m = 1.921 nm, β_m = 89 deg
Lattice Invariant Shear: $(101)[10\bar{1}]_{bcc} \rightarrow (111)[11\bar{2}]_{fct}$ **and** $(001)[100]_{9R}$
Lattice Correspondence: $[10\bar{1}]_b \rightarrow [100]_{9R}$ $\quad [100]_b \rightarrow 1/9\ [401]_{9R}$
$[010]_b \rightarrow [010]_{9R}$ $\quad [010]_b \rightarrow [010]_{9R}$
$[504]_b \rightarrow [001]_{9R}$ $\quad [001]_b \rightarrow 1/9[\bar{5}01]_{9R}$

Solution	g	Habit Plane Normal Direction Cosines	Shape Strain Direction Cosines	m	Orientation Relationship: Direction Cosines of Product Axes		
					a_m	b_m	c_m
1A	0.00571	0.1325, −0.6821, −0.7192	0.1193, 0.7527, −0.6475	0.2031	0.67368	−0.01836	0.75044
					0.08596	0.99485	−0.05215
					−0.73400	0.09965	0.65888
1B	0.00571	0.1325, 0.6821, −0.7192	0.1193, −0.7527, −0.6475	0.2031	0.67368	0.01836	0.75044
					−0.08596	0.99485	0.05215
					−0.73400	−0.09965	0.65888
2A	0.18294	0.7192, −0.6821, −0.1325	0.6475, 0.7527, −0.1193	0.2031	0.73400	−0.09965	0.68450
					0.08596	0.99485	0.05515
					−0.67368	0.01836	0.72692
2B	0.18294	0.7192, 0.6821, −0.1325	0.6475, −0.7527, −0.1193	0.2031	0.73400	0.09965	0.68450
					−0.08596	0.99485	−0.05515
					−0.67368	−0.01836	0.72692

12.0 ± 1.0 deg in Cu-Zn[36] and Ag-Cd[40] alloys, respectively. Measured and predicted orientation relationships are also in excellent agreement, as shown in Table III. The angles between the pairs of planes and directions identified conveniently from SAED patterns agree to within experimental error on the measurements, and it is possible, on closer analysis, to confirm that the sense of the rotations predicted agrees with those observed.

An interesting feature of the solutions to the theory for the *B*2 to 9*R* transformation presented in Table IV is that they show that, for the lattice parameters employed, the total lattice strain may approximate to an invariant plane strain and that the magnitude of the LIS required is small ($g \sim 0.006$). Indeed, as indicated in Figure 13, the predicted magnitude of LIS may be reduced to zero with small changes in lattice parameters. Figure 13 shows the predicted variation in g with systematic variation in the parameters of the product monoclinic 9*R* phase about the commonly reported values, for a constant value of $a_b = 0.2930$ nm. For the low shear solutions, the magnitude of LIS is seen to be independent of b_m, but to vary significantly with a_m, c_m, and β_m and to approach zero at values of each of these parameters that are close to those used in the calculations for Table IV. For $g = 0$, the product phase would be expected to have a fault-free 9*R* structure and to be coherent with the matrix β' phase across the irrational habit plane. However, for lattice parameters leading to a finite value of g, the α_1 plates would be expected to have a substructure of coarsely spaced stacking faults, in qualitative agreement with common observations for α_1 plates in Cu-Zn and Ag-Cd alloys. The magnitude of g and thus the spacing of the stacking faults arising as the result of the LIS would be expected to vary with variation in lattice parameters. For the parameters employed by Cornelis and Wayman,[36,53] for example, the predicted magnitude of g rises to ~0.012 and the corresponding α_1 plates would be expected to contain a higher density of internal faulting arising from the LIS.

VI. DISCUSSION

As is the case for the martensitic fcc to hcp transformation, the structural change associated with the γ' precipitate plates in Al-Ag alloys may be considered to be accomplished by the passage of transformation dislocations across the broad faces of the plates.[48,49] Defining steps that are two close-packed planes in height normal to the habit plane, these transformation dislocations have a Burgers vector that is of the form $(a/6)\langle 11\bar{2}\rangle$ in the case of coherent transformation to an ideal hcp structure with zero volume change, but that must acquire a small additional component of magnitude ξ normal to the habit plane, *i.e.*, $(2a\xi/3)\langle 111\rangle$, for the γ' phase. For such transformations, formal application of the crystallographic theory, assuming the total lattice strain $\mathbf{S}_T$ to be an invariant plane strain on the $(111)_f$ habit plane with a displacement $m\mathbf{d}$ equivalent to the Burgers vector of the transformation dislocations, provides a complete description of the transformation crystallography.

For both the martensitic fcc to hcp transformation and the formation of γ' plates, there is strong experimental evidence to support the proposal that the fcc to hcp transition involves the lateral migration of steps across the otherwise coherent broad faces of the product phase, with the steps having the configuration of the transformation dislocations required by the theory. High-resolution electron microscope images,[22,23] recorded parallel to close-packed (*i.e.*, $\langle 110\rangle_f \| \langle 11\bar{2}0\rangle_h$) directions in the habit plane, reveal that the interface commonly contains steps normal to the habit plane that are invariably an even multiple of close-packed layers in height and that may be interpreted as an accumulation of $(a/6)\langle 11\bar{2}\rangle$ Shockley partial transformation dislocations on every second close-packed $(111)_f$ plane. Single transformation dislocations are observed, but it is more common for the observed steps to comprise multiple dislocations. Since the layer of atoms (1) immediately

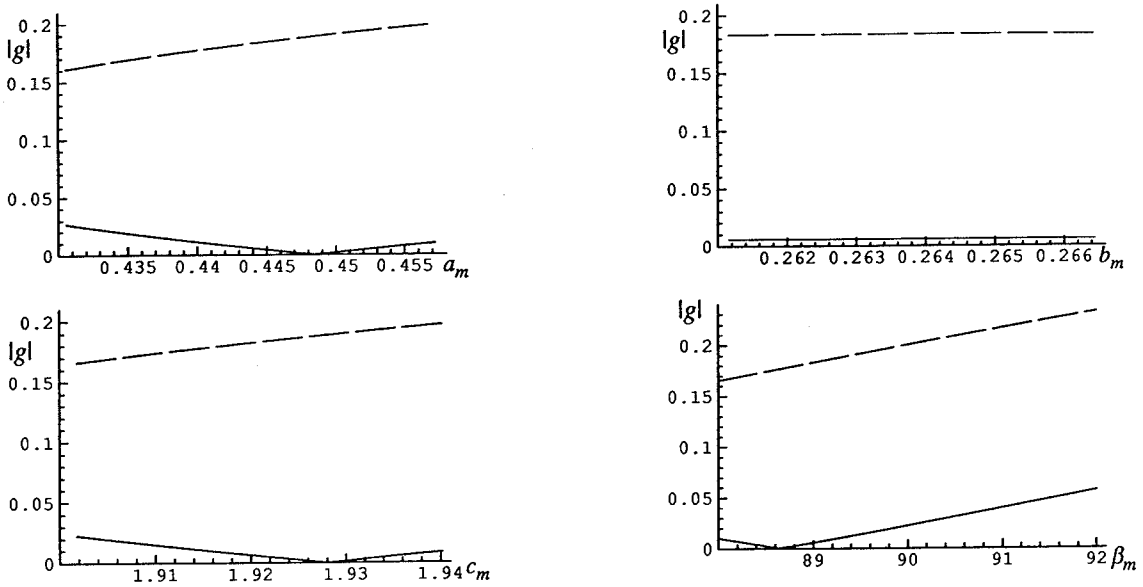


Fig. 13—Predicted variation in the magnitude (g) of the LIS with variation in each of the product lattice parameters a_m, b_m, c_m, and β_m for the *B*2 to 9*R* transformation. The lattice parameter of the parent phase a_b is assumed constant.

ahead of the interface in Figure 1(a) may be regarded as common to the fcc and hcp lattices, it has been suggested[23] that there may be some adjustment in composition in this layer ahead of the passage of a transformation dislocation. However, the completion of the structural transition requires the migration of a transformation dislocation over the next adjacent layer in the fcc phase and an adjustment in the composition that allows this plane of the precipitate to contain wholly Ag atoms. It is inferred that passage of the transformation dislocation involves atom by atom transfer between parent and product phases, with Ag atoms attaching themselves preferentially at kink sites along the Shockley partial transformation dislocations. The rate of migration may be controlled either by long-range transport of Ag atoms ahead of the migrating transformation dislocations or by the substitutional diffusion of Ag atoms across the kinks. Although there is a substantial compositional change associated with the transformation, the diffusion-controlled migration of the transformation dislocations in the manner described permits a coherent interface to be sustained and to migrate normal to the habit plane to sustain thickening of the plate and the maintenance of a correspondence of lattice sites.

In a given $\{111\}_f$ habit plane there are three equivalent $\langle 11\bar{2}\rangle_f$ twinning directions, and the passage of a transformation dislocation with a Burgers vector given by Eq. [2] in any one of these directions will generate the observed hcp structure. If the transformation involves migration of transformation dislocations in a single $\langle 11\bar{2}\rangle_f$ direction on alternate $(111)_f$ planes, then an appreciable shape change will accumulate over the transformed volume. When the component of displacement normal to the habit plane is zero, this shape change will be a simple shear, but if the normal component of displacement is finite, then the shape change will have the more general form of an invariant plane strain. If, alternatively, transformation dislocations are generated in all three $\langle 11\bar{2}\rangle_f$ directions and the transformed volume contains equal volume fractions of all three variants of the hcp structure, then the variants will be self-accommodating and the shear component of the net shape deformation over the transformed volume will be zero. Of course, it is possible to imagine situations between these two extremes,[49] but for transformation product forming within the bulk constraints of the parent phase it is to be expected that all three transformation variants will operate within a given product plate to minimize the shape deformation, which is consistent with direct experimental observations for bulk γ' plates in Al-Ag.[23] If, however, the transformation were to occur under the influence of an applied stress with a significant shear component, then an interaction that relaxed the potential of the stress field might be expected to favor the adoption of one preferred shear component of transformation and the development of a significant shape change. Similarly, for those plates nucleating at or in the immediate vicinity of a free surface, the relaxed matrix constraints may lead to preference for a single set of transformation dislocations and the accumulation of a finite shape deformation that approaches one-half the twinning shear in magnitude. Such is the explanation offered[55] for the observation of invariant plane-strain surface relief associated with at least some γ' precipitate plates at free surfaces in Al-Ag alloys.

The formation of the ordered orthorhombic phase AuCu II from disordered fcc parent phase also generates invariant plane-strain surface relief and conforms with predictions of the crystallographic theory.[25] Based on the assumptions that the lattice correspondence is identity and the LIS is a partial twinning shear on the system $(101)[\bar{1}01]$, the theory is capable of accounting consistently and accurately for the observed habit plane, surface relief, and orientation relationship. With appropriate choice of variants of the lattice correspondence, it is also capable of generating crystallographic variants of the habit plane normal and displacement direction that each cluster about common $\langle 110\rangle_f$ directions, and thus account for the growth of AuCu II plates in partially self-accommodating pairs or pyramidal groups. Moreover, recent measurements[52] have confirmed that the observed fine-scale twinning within AuCu II plates is on the appropriate twinning system required by the theory and that the relative fraction of the twin-related variants within a given plate is accurately consistent with the magnitude of the LIS required to ensure that the irrational habit plane is macroscopically invariant. It is thus strongly suggested that the twinned substructure of the plates is a direct manifestation of the LIS.

The AuCu II plates have a form and substructure similar to those of martensite plates in which the LIS involves twinning of the product structure, and the faceted interface between twinned product and single-crystal matrix that is evident in Figures 7 and 9 is typical of one of the two common models proposed for a semicoherent martensite interface.[48,49,50] In this model, the interface facets define areas of coherence between parent and product lattices, separated periodically in the product phase by twin boundaries. For a diffusionless, martensitic transformation, such an interface is considered to be glissile and capable of migration as an entity normal to the habit plane through homogeneous transformation of the coherent regions combined with extension of the twin boundaries to preserve an invariant habit plane. However, for an ordering transformation involving nearest-neighbor exchange of atoms, such as the formation of AuCu II plates, it is currently not clear that migration of such a semicoherent interface in the direction of its normal will sustain the coordinated atom displacements necessary to the maintenance of a correspondence of lattice sites.

To resolve this apparent difficulty, it has been proposed[29] that both lengthwise growth and thickening of AuCu II plates involve the generation and glide of transformation dislocations parallel to the habit plane. For a pyramidal group of four plates such as that illustrated schematically in Figure 4, it is suggested that the dislocations responsible for growth on the inner surfaces of the pyramid would be generated cooperatively at the common inner apex. The Burgers vectors $m\mathbf{d}$ of the dislocations on the four plates would be equal in magnitude and approximately parallel to the same $[110]_f$ direction. It is anticipated that the individual dislocation segments would form a continuous loop around the inner surface

of the pyramid and that cooperative glide across the broad inner surfaces of the plates would allow the four plates to grow at the same rate, with the inner apex moving parallel to the $[0\bar{1}1]_f$ axis of the pyramid. The inner surfaces of the pyramid would likely contain steps corresponding to single or multiple transformation dislocations and, if growth were largely confined to the inner surfaces of the pyramid, the plates would become tapered in section, with the inner surfaces and their junction lines deviating significantly from the corresponding habit planes and habit plane intersections.

The form and crystallography of the proposed transformation dislocations have not been addressed in detail, and the compatibility of the transformation dislocations, which generate the observed shape deformation, with the internal twinning, which must occur within the lattice created by the shape strain and, on this interpretation, constitutes evidence of the complementary strain of the original Bowles-Mackenzie formulation of the crystallographic theory,[2] requires further examination. However, it has been demonstrated that the model of the growth mechanism based on such transformation dislocations is consistent with the morphology of AuCu II plates and with measurements of both the lengthening and thickening kinetics of plate growth. The migration of transformation dislocations within the habit plane also provides a plausible mechanism by which a correspondence of lattice sites may be sustained in the face of the atomic interchange necessary to accomplish the ordering reaction. The atomic displacements are restricted to coordinated atom transfer at single dislocation steps, and it is suggested that the dislocations will migrate at a rate determined by the rate of ordering of atoms at the steps. It is in the nature of these steps and the mechanism of atom transfer across the steps that the proposed mechanism differs significantly from earlier models,[27,28] in which separate mechanisms were proposed for lengthening and thickening of plates and in which plate thickening was proposed to involve the migration of disordered ledges across the habit plane interface. It remains difficult to reconcile the maintenance of a lattice correspondence and the resulting generation of an invariant plane-strain shape change with a mechanism in which random atom transfer occurs across disordered ledges in the interface. Unlike the γ' precipitate plates in Al-Ag alloys, there have to date been no detailed studies of the interface structure in AuCu II plates, and there thus remains little direct evidence on which to assess the proposed transformation mechanisms.

While there is no composition change associated with the growth of ordered AuCu II plates, the question of whether there is a change in composition accompanying the initial formation of α_1 plates in Cu-Zn(-X) and Ag-Cd alloys is central to the present discussion and the answer remains controversial.[41,56–62] The bulk of current evidence strongly suggests that there is a significant difference in composition between the α_1 plates and β' matrix at the earliest stages of transformation for which measurements can presently be made. There is also evidence from some studies[37,38,41] that the order present in the matrix β' phase is not necessarily preserved in the formation of the 9*R* plates, confirming the existence of significant atomic interchange accompanying initial growth. In the Ag-Cd system it has been confirmed[41] that the α_1 plates have a disordered 9*R* structure, in conjunction with a composition approaching that of the equilibrium α phase, well before any evidence of the 9*R* to 3*R* structural transition that accompanies prolonged isothermal aging. It is thus difficult to sustain an argument that initial growth is diffusionless with a subsequent adjustment in composition, for it would require this adjustment in composition to preserve the 9*R* structure, only to have it subsequently transform to equilibrium 3*R*. Even if, as has been suggested,[41,60] the α_1 plates lengthen in a displacive (diffusionless) manner and then thicken at a rate controlled by substitutional volume diffusion, there is no evidence to suggest that the plates are composite in structure and the lateral growth must thus occur in such a way that the 9*R* structure, lattice correspondence, and shape change are sustained during diffusional growth.

Formation of the α_1 plates produces a shape change compatible with an invariant plane strain, and the crystallographic theory again provides an excellent basis for accounting for the observed crystallographic features of the transformation. Although initial applications[36,40] of the theory, which assumed for convenience the product structure to be fcc(t), produced predictions in general agreement with experimental observations, the present analysis suggests that this interpretation may be in part misleading. In the early analyses, the magnitude of the LIS on the observed slip system required to preserve an invariant habit plane was found to be approximately equivalent to the shear of $(a/6)[11\bar{2}]$ on every third (111) plane that would generate 9*R* stacking of the close-packed planes of a 3*R* fcc structure. This approach thus has the potential to create the impression that the 9*R* structure of the α_1 plates is a consequence of the LIS component of the shape strain, rather than the preferred product structure in the initial stages of decomposition of supersaturated β' phase. It also fails to provide a sound basis for accounting for the substructure of stacking faults in α_1 plates, for those faults detectable are much more widely spaced than every third {111} plane.

If a comparison is made of the principal distortions (η_i) of the homogeneous lattice deformations **B** for the *B*2 to 9*R* and *B*2 to fcc(t) transformations, then, as shown in Table V, those associated with the *B*2 to 9*R* transition are much closer to unity than those for the transformation to the fcc(t) structure. This implies that the lattice distortion associated with formation of the 9*R* structure is significantly less than that required to form the fcc structure and suggests that it is more appropriate to regard the 9*R* structure as a preferred metastable phase, intermediate between the parent *B*2 β' and equilibrium product α. In support of this interpretation, it is to be noted that the magnitude (g) of the LIS required for the *B*2 to 9*R* transition approaches zero for small variations in lattice parameters from observed values. For $g = 0$, the α_1 plates would be expected to be fully coherent with the parent phase and grow fault-free. This may account for the recent, and yet to be confirmed, observation of coherent α_1 plates free of faulted substructure in the very early stages of transformation in a Cu-Zn alloy.[39] It has been suggested[39] that such plates represent the initial growth stage for α_1 plates and that

Table V. Comparison of the Principal Distortions η_i of the Homogeneous Lattice Strain B for the *B*2 to fcc(t) and *B*2 to 9*R* Transformations in Cu-Based Alloys

Transformation	η_1	η_2	η_3
$B2 \rightarrow$ fcc[36]	0.891894	0.891894	1.26133
$*B2 \rightarrow 9R$[54]	0.900341	0.997322	1.10645
$**B2 \rightarrow 9R$[36]	0.891891	0.994588	1.13109

*Calculated using the lattice parameters of Kajiwara[54] (Table IV).
**Calculated using 9*R* lattice parameters equivalent to those of Cornelis and Wayman.[36]

the plates subsequently acquire the faulted substructure in a second stage of growth, during which they lose coherency with the matrix. In formal terms, the crystallographic theory allows for the magnitude of the LIS to become finite, with small local variations in the lattice parameters at the transformation interface, perhaps as a result of local changes in composition. The theory thus allows for the development of a faulted substructure, although the mechanism by which faults apparently propagate through the existing volume of plate established during the initial coherent growth stage perhaps needs further attention. Irrespective of the outcome of such analysis, it is to be emphasized that, for the *B*2 to 9*R* transition and the available lattice parameters, the theory requires a small finite value of g and thus a substructure of relatively widely spaced stacking faults that is qualitatively consistent with observations.

Following a recent study of the structure of the broad faces of α_1 plates in a Cu-Zn alloy, it was reported that the plates develop growth ledges during the second stage of growth and suggested that growth proceeds *via* diffusion-controlled ledge wise thickening.[39,63] The observation of an invariant plane-strain shape change and the well-established applicability of the crystallographic theory imply that this growth should proceed in such a way that a correspondence of lattice sites is preserved. It thus seems likely that the growth ledges responsible for thickening should have the form of transformation dislocations gliding across the broad faces of the plates at a rate that is diffusion controlled. It is possible that a correspondence of sites may be maintained if there is a coordinated atom by atom transfer from parent to product phase at individual dislocation steps in the interface, but again unlikely that this will be so if the risers of the growth ledges are disordered or incoherent. The form and crystallography of any such ledges or transformation dislocations on the broad faces of α_1 plates remain unresolved in detail, and it would seem that prospects for progress in resolving details of the mechanism of the transformation rest with further detailed high-resolution studies of the interface and, as in the case of AuCu II plates, analysis of the compatibility of potential transformation dislocations with those responsible for the lattice invariant shear.

VII. SUMMARY

1. The transformation products associated with

(a) precipitation of hcp Ag_2Al (γ') plates in Al-Ag alloys;

(b) formation of ordered orthorhombic AuCu II plates in equiatomic Au-Cu; and

(c) formation of 9*R* α_1 plates in ordered *B*2 Cu-Zn(-X) and Ag-Cd alloys

produce relief at a free surface consistent with an invariant plane-strain change. The transformations are thus properly described as displacive.

2. The observed shape change implies that a lattice correspondence is maintained between parent and product phases. The existence of such a lattice correspondence is fundamental to justifying the application of the theory of martensite crystallography to this class of transformations.
3. The theory of martensite crystallography may be applied successfully to account for the crystallographic features of each of these transformations. The agreement between experimental observations and theoretical predictions is exceptionally good in all cases.
4. Although it has been common to assume that atomic interchange or a compositional change involving long-range substitutional diffusion is incompatible with the maintenance of a lattice correspondence, the formal crystallographic theory does not necessarily require a correspondence of atom positions and thus does not preclude interchange of atom positions or a change in composition. Successful application of the theory does, however, imply a transformation mechanism that is compatible with the maintenance of a correspondence of lattice sites.
5. It seems likely that the mechanism of transformations in this group involves the migration of transformation dislocations in the planar interface between parent and product, and coordinated atom by atom transfer across the interface associated with motion of the corresponding dislocation steps in the interface. The rate of migration of the transformation dislocations is controlled by the rate at which the required atoms assemble ahead of the migrating dislocation steps. The migration of the transformation dislocations produces the displacements that generate the surface relief.

REFERENCES

1. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. AIME*, 1953, vol. 194, p. 1503.
2. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 129, 138, and 224.
3. A.F. Acton, M. Bevis, A.G. Crocker, and N.D.H. Ross: *Proc. R. Soc. (London) A*, 1970, vol. 320, p. 101.
4. C.M. Wayman: *Introduction to the Crystallography of Martensitic Transformations*, Macmillan, New York, NY, 1964.

5. J.W. Christian: *The Theory of Transformations in Metals and Alloys*, Pergamon Press, Oxford, 1975.
6. Z. Nishiyama: *Martensitic Transformation*, Academic Press, New York, NY, 1978.
7. B.C. Muddle and G.R. Hugo: *Proc. 7th Int. Conf. Martensitic Transformations (ICOMAT '92)*, Monterey, CA, 1992, in press.
8. J.D. Watson and P.G. McDougall: *Acta Metall.*, 1973, vol. 21, p. 961.
9. J.S. Bowles and N.F. Kennon: *J. Aust. Inst. Met.*, 1960, vol. 5, p. 106.
10. G.R. Srinivasan and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, p. 621.
11. S. Hoekstra: *Acta Metall.*, 1980, vol. 28, p. 507.
12. B.P.J. Sandvik: *Metall. Trans. A*, 1982, vol. 13A, p. 777.
13. J.S. Bowles, B.C. Muddle, and C.M. Wayman: *Acta Metall.*, 1977, vol. 25, p. 513.
14. M.P. Cassidy, B.C. Muddle, T.E. Scott, C.M. Wayman, and J.S. Bowles: *Acta Metall.*, 1977, vol. 25, p. 829.
15. U. Dahmen, P. Ferguson, and K.H. Westmacott: *Acta Metall.*, 1987, vol. 35, p. 1037.
16. J. Van Landuyt and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, p. 803.
17. C.M. Wayman and J. Van Landuyt: *Acta Metall.*, 1968, vol. 16, p. 815.
18. J.A. Hren and G. Thomas: *Trans. Met. Soc. AIME*, 1963, vol. 227, p. 308.
19. C. Laird and H.I. Aaronson: *Acta Metall.*, 1967, vol. 15, p. 73.
20. C. Laird and H.I. Aaronson: *Acta Metall.*, 1969, vol. 17, p. 505.
21. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, p. 845.
22. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 639 and 649.
23. J.M. Howe, U. Dahmen, and R. Gronsky: *Phil. Mag.*, 1987, vol. 56, p. 31.
24. J.S. Bowles and A.S. Malin: *J. Aust. Inst. Met.*, 1960, vol. 5, p. 131.
25. R. Smith and J.S. Bowles: *Acta Metall.*, 1960, vol. 8, p. 405.
26. A.M. Hunt and D.W. Pashley: *J. Aust. Inst. Met.*, 1963, vol. 8, p. 61.
27. A.J. Pedraza and J. Kittl: *Acta Metall.*, 1976, vol. 24, p. 835.
28. H.I. Aaronson and K.R. Kinsman: *Acta Metall.*, 1977, vol. 25, p. 367.
29. J.S. Bowles and C.M. Wayman: *Acta Metall.*, 1979, vol. 27, p. 833.
30. G. Van Tendeloo, S. Amelinckx, S.J. Jeng, and C.M. Wayman: *J. Mater. Sci.*, 1986, vol. 21, p. 4395.
31. R.D. Garwood: *J. Inst. Met.*, 1954–55, vol. 83, p. 64.
32. R.D. Garwood: *Iron Steel Inst. Spec. Rep.*, No. 93, 1965, p. 90.
33. P.E.J. Flewitt and J.M. Towner: *Acta Metall.*, 1966, vol. 14, p. 1013.
34. E. Hornbogen and H. Warlimont: *Acta Metall.*, 1967, vol. 15, p. 943.
35. P.E.J. Flewitt and J.M. Towner: *J. Inst. Met.*, 1967, vol. 95, p. 273.
36. I. Cornelis and C.M. Wayman: *Acta Metall.*, 1974, vol. 22, p. 301.
37. M.-H. Wu and C.M. Wayman: *Proc. Int. Conf. Martensitic Transformations*, Japan Institute of Metals, Sendai, Japan, 1986, p. 619.
38. M.-H. Wu, J. Perkins, and C.M. Wayman: *Acta Metall.*, 1989, vol. 37, p. 1821.
39. K. Chattopadhyay and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, p. 695.
40. M.M. Kostic and E.B. Hawbolt: *Metall. Trans. A*, 1979, vol. 10A, p. 165.
41. M.-H. Wu, B.C. Muddle, and C.M. Wayman: *Acta Metall.*, 1988, vol. 36, p. 2095.
42. J.S. Bowles and C.M. Wayman: *Metall. Trans.*, 1972, vol. 3, p. 1113.
43. J.W. Christian: *Proc. R. Soc. (London) A*, 1951, vol. 206, p. 51.
44. P. Gaunt and J.W. Christian: *Acta Metall.*, 1959, vol. 7, p. 529.
45. S. Mahajan, M.L. Green and D. Brasen: *Metall. Trans. A*, 1977, vol. 8A, p. 283.
46. C. Hayzelden, K. Chattopadhyay, J.C. Barry, and B. Cantor: *Phil. Mag.*, 1991, vol. A63, p. 461.
47. J.W. Brooks, M.H. Loretto, and R.E. Smallman: *Acta Metall.*, 1979, vol. 27, p. 1839.
48. J.W. Christian: in *Interfaces Conf.*, R.C. Gifkins, ed., Butterworth's, London, 1969, p. 159.
49. J.W. Christian: in *Martensite*, G.B. Olson and W.S. Owen, eds., ASM International Materials Park, OH, 1992, p. 103.
50. J.W. Christian and K.M. Knowles: in *Proc. Int. Conf. Solid-Solid Phase Transformations*, H.I. Aaronson, D.E. Laughlin, R.F. Sekerka, and C.M. Wayman, eds., TMS-AIME, Warrendale, PA, 1982, p. 1185.
51. J.W. Christian: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 371.
52. J.F. Nie and B.C. Muddle, Monash University, Clayton, Victoria, Australia, unpublished research, 1993.
53. I. Cornelis and C.M. Wayman: *Acta Metall.*, 1974, vol. 22, p. 291.
54. S. Kajiwara: *Trans. Jpn. Inst. Met.*, 1976, vol. 17, p. 435; 1976, vol. 17, p. 448.
55. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, p. 1029.
56. I. Cornelis and C.M. Wayman: *Scripta Metall.*, 1973, vol. 7, p. 579.
57. G.W. Lorimer, G. Cliff, H.I. Aaronson, and K.R. Kinsman: *Scripta Metall.*, 1975, vol. 9, p. 271.
58. M.M. Kostic and E.B. Hawbolt: *Scripta Metall.*, 1975, vol. 9, p. 1173.
59. G.W. Lorimer, G. Cliff, H.I. Aaronson, and K.R. Kinsman: *Scripta Metall.*, 1975, vol. 9, p. 1175.
60. P. Doig and P.E.J. Flewitt: *Met. Sci.*, 1983, vol. 17, p. 601.
61. B.C. Muddle and H.L. Fraser: in *Proc. Int. Conf. Solid-Solid Phase Transformations*, H.I. Aaronson, D.E. Laughlin, R.F. Sekerka, and C.M. Wayman, eds., TMS-AIME, Warrendale, PA, 1983, p. 987.
62. Y. Nakata, T. Tadaki, and K. Shimizu: *Mater. Trans. JIM*, 1989, vol. 30, p. 107.
63. H.I. Aaronson, T. Furuhara, J.M. Rigsbee, W.T. Reynolds, Jr., and J.M. Howe: *Metall. Trans. A*, 1990, vol. 21A, p. 2369.

A Comparison between Three Simple Crystallographic Principles of Precipitate Morphology

U. DAHMEN

The connection between the optimum shape and orientation relationship of precipitates in a solid is examined. Three simple criteria for precipitate morphology are compared and illustrated schematically: the principle that precipitate dimensions tend to be inverse to the magnitude of the transformation strain; the postulate that precipitates are bounded by unrotated planes (eigenplanes); and the proposal that interfaces are parallel to the planes of three independent dislocation loop arrays necessary to accommodate the transformation strain completely. These principles are illustrated for different orientation relationships, and it is shown that special features are displayed by invariant-line precipitates. The implications of these criteria for experimental studies of precipitate morphologies are discussed and their predictions compared with results from a recent study of lath-shaped precipitates in Cu-Cr alloys.

I. INTRODUCTION

THE morphology of precipitates in a solid matrix is of fundamental importance in materials science. Underlying most investigations of microstructures and their development is the assumption that the shape, orientation relationship, and interface structure of a product forming in a parent phase are interlinked. However, the nature of this link is not well understood, and morphological prediction is of limited success. It is generally implied that the link lies in the mechanism of strain accommodation, either during nucleation or in the process of coherency loss or growth.[1,2] Yet, for most cases, there is still no clear and simple relationship between the crystallographically observable features of a solid inclusion and its mechanism of formation.

A notable exception is the phenomenological theory of martensitic transformations (*e.g.*, Reference 3), which has been remarkably successful in many of its predictions. However, its application is limited to inclusions that are dominated by a single planar, glissile interface, and to cases where diffusion is negligible, so that an atomic correspondence is maintained during transformation. Most inclusions do not fulfill these conditions, since diffusion prohibits an atom-by-atom correspondence, and interfaces are neither constrained to be glissile nor to be planar. In fact, inclusions are found in a great variety of shapes and orientations, depending on crystal structure and transformation conditions.

Since no atomic (as opposed to lattice) correspondence can be maintained during bainitic or diffusional transformations, one might expect random shapes and orientations, determined only by the parameters governing diffusion. Instead, most diffusional transformation products adhere to well-defined orientation relationships and exhibit a narrow range of shapes. In fact, a large number of common precipitation systems form plate-shaped inclusions similar to those of martensitic transformation products[4] which offer themselves readily to application of the concepts of martensite crystallography, regardless of their formation mechanism. Such attempts have often met with some degree of success,[5,6] showing that a lattice correspondence can be maintained during precipitation reactions.

Other evidence for a lattice correspondence in a diffusional transformation was shown by Ryder and Pitsch,[7] who observed a limited range of orientation relationships during precipitation in a Co-Fe alloy. More recently, detailed studies have shown that many precipitation systems exhibit a single preferred orientation relationship and corresponding shape. The orientation relationship is often found to be such that the transformation strain is of the invariant-line type, and the corresponding shape is a needle or lath along the invariant line. Experimental observations indicate that for surface precipitates, the invariant line lies at the intersection of the cone of unextended lines with the surface,[8] while bulk precipitates lie along invariant lines at the intersection of the cone with a low-index crystallographic plane.[8,9,10] These observations are clear evidence that an optimum orientation relationship and shape are maintained during many diffusional transformations.

It is the intent of the present contribution to outline some crystallographic principles linking precipitate shapes and orientation relationships, and to examine the relationship between the geometrically observable features and the underlying principles.

II. RESULTS AND DISCUSSION

A. *Shape and Orientation Relationship*

The optimum shape of an inclusion depends on the transformation strain, which in turn depends on the lattice orientation relationship. To illustrate this interrelationship, consider the simple transformation shown in Figure 1. When a precipitate lattice is derived from the matrix lattice by an expansion a along the x direction and a contraction b along the y direction, the transformation $\mathbf{B} = \left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$ can be illustrated schematically, as

U. DAHMEN, Senior Scientist and Director, is with the National Center for Electron Microscopy, Lawrence Berkeley Laboratory, Berkeley, CA 94720.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

shown in Figure 1(a). The unit circle is transformed into the strain ellipse,[11] here shown for the specific case of $a = 1.12$ and $b = 0.91$. Directions at the intersection of the unit circle with the ellipse are unextended but rotated by the transformation. If one of these unextended lines is rotated back to its original orientation (dashed line in Figure 1), it becomes an invariant line; *i.e.*, it will be unextended and unrotated and therefore a direction in which the two lattices match without strain. This is illustrated in Figure 1(b), where a small lattice rotation of $\alpha = 5.76$ deg has changed the orientation relationship such that the transformation $\mathbf{A} = \mathbf{RB} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ is an invariant-line strain.

Both the direction of the invariant line and the lattice rotation to produce it depend on the lattice parameters of the matrix and precipitate phase and therefore are usually not rational. The angle θ of the invariant line with the x-axis is given by[10]

$$\tan\theta = \sqrt{\frac{a^2 - 1}{1 - b^2}} \qquad [1]$$

and the angle α of lattice rotation necessary to form an invariant line is found from

$$\cos\alpha = \left(\frac{1 + ab}{a + b}\right) \qquad [2]$$

The effect of the rotation on the transformation strain is seen more clearly from the displacement field diagrams. Figure 1(c) illustrates the displacement field of the transformation for $\alpha = 0$ deg. The displacement vectors take on all directions in the plane. Wherever the displacement is tangential, the corresponding radius vector is an unextended line that is simply rotated. By adding an equal and opposite rotation of $\alpha = 5.76$ deg, we change the displacement field dramatically to that shown in Figure 1(d). The displacements have been eliminated along one direction (the invariant line) and increased in other directions. In addition, all displacements are now parallel (or antiparallel) to each other.

It is apparent from these schematics that the two displacement fields would demand very different shapes, even though the transformations differ only by a small relative rotation. Intuitively, it would be expected for a precipitate to choose a shape of minimum strain and hence be elongated for the invariant-line strain, while a more equiaxed inclusion would be expected for the unrotated case. Several different criteria have been given for the optimum shape of an inclusion, and the three that are most commonly implied will be reviewed in Sections C through E. Since the corresponding shapes are strongly affected by the orientation relationship, the effect of a lattice rotation α will be illustrated for all of

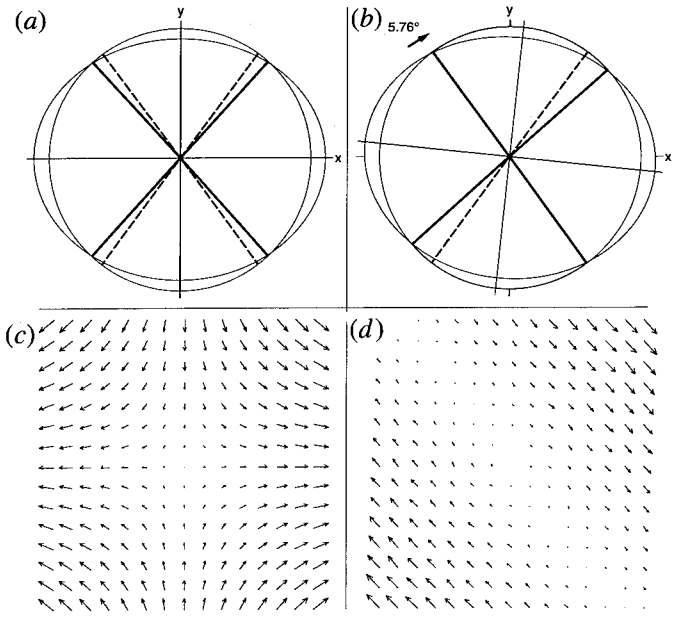


Fig. 1—Schematic illustration of transformation strains with and without an added lattice rotation, using complementary representations of strain ellipse and displacement field. In (*a*), a unit circle is deformed by expansion/contraction into an ellipse. The dashed lines are unextended but rotated to the positions shown by solid lines. In (*b*), a 5.76 deg rotation has turned one of the unextended lines into an invariant line. In (*c*) and (*d*), the same transformations are illustrated as vector fields, showing the magnitude and direction of displacements. In (*d*), the displacements vanish along the invariant line, and all displacement vectors have become parallel to each other.

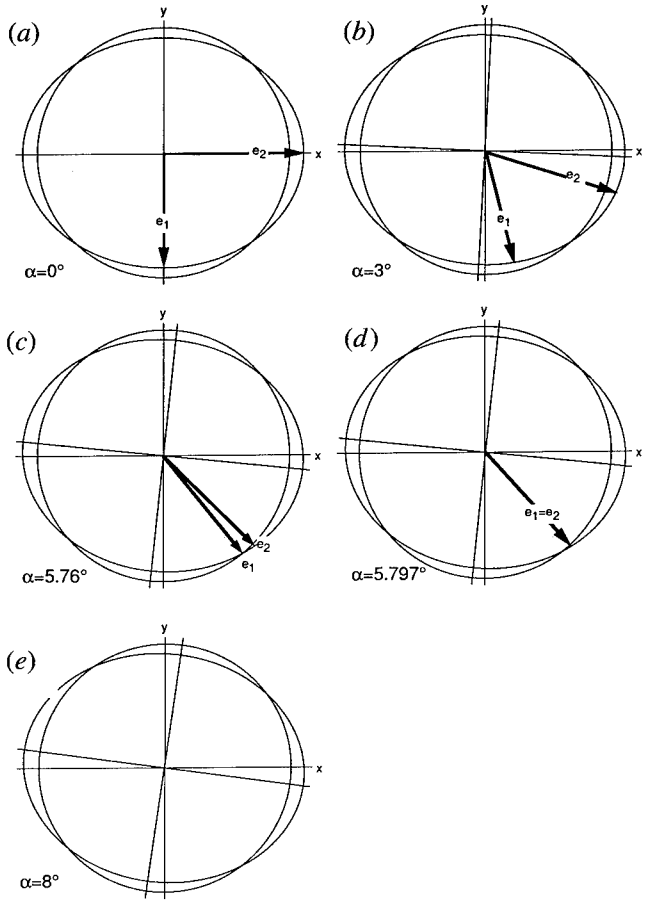


Fig. 2—Behavior of eigenvectors for the example transformation as a function of orientation relationship (rotation angle α). With increasing rotation, the eigenvectors approach each other in angle as the eigenvalues approach each other in magnitude. No real eigenvalues exist beyond $\alpha = 5.797$ deg.

them. Figures 2 through 6 show different features for the same sequence of orientation relationships described by rotation angles α = 0, 3, 5.76, 5.797, and 8 deg. This parallel presentation simplifies a direct comparison between different figures illustrating the three criteria.

B. *Effect of Orientation Relationship on Eigenvectors of the Transformation*

Because the eigenvectors and eigenvalues of the matrix $\mathbf{A} = \mathbf{RB}$ have a strong bearing on the optimum shape, it is useful to illustrate briefly their behavior as a function of orientation relationship. Figure 2 shows this behavior for the five different rotation angles α, using the strain ellipse for representation. For α = 0 deg, the eigenvectors are the principal axes, enclosing an angle of 90 deg. With increasing lattice rotation, the eigenvectors approach each other in direction as the eigenvalues approach each other in magnitude. For α = 3 deg, the eigenvalues are 1.10 and 0.93, while the eigenvectors enclose an angle of 59 deg. For α = 5.76 deg, the eigenvalues are 1.0 (invariant line) and ab = 1.02, and the eigenvectors enclose an angle of 6 deg. For α = 5.797 deg, the eigenvalues are equal, with $\sqrt{ab}$ = 1.01, and the eigenvectors are parallel to each other, enclosing an angle of $\theta = \tan^{-1}\sqrt{a/b}$ = 47.9 deg with the x-axis. Any further lattice rotation leads to the loss of real eigenvalues, and no directions remain unrotated. These different cases are illustrated in Figures 2(a) through (e). Notice how a small lattice rotation α can have a large effect on the orientation of the eigenvectors. For example, between Figures 2(c) and (d), the lattice rotation differs by only 0.03 deg, causing the eigenvectors to rotate toward each other by about 3 deg each. This magnification is related to the moiré effect and is of practical importance for the crystallographic analysis. To distinguish the cases in Figures 2(c) and (d) experimentally by their orientation relationships requires an accuracy on the order of 0.01 deg, whereas to distinguish the two cases by their eigenvectors, an accuracy of 1 deg is sufficient.

Figure 3 illustrates the displacement fields for the same transformation and orientation relationships as Figure 2. Unextended lines are radius vectors pointing to tangential displacements, whereas unrotated lines (eigenvectors) are radius vectors pointing to radial displacements. The eigenvectors shown schematically in Figure 2 can be verified by inspection in Figure 3. It can also be seen from Figure 3(e) that with an 8 deg rotation, the transformation has no eigenvectors because all directions suffer some tangential displacement (a net rotation). For the special case of an invariant-line strain (Figure 3(c)), the displacement vectors are aligned in a single direction parallel to the eigenvector $\mathbf{e}_2$, while the invariant line is parallel to $\mathbf{e}_1$. The angle between these two directions reflects the volume change of the transformation. Because the volume change for this example is small (det $B = ab$ = 1.02), the angle between $\mathbf{e}_1$ and

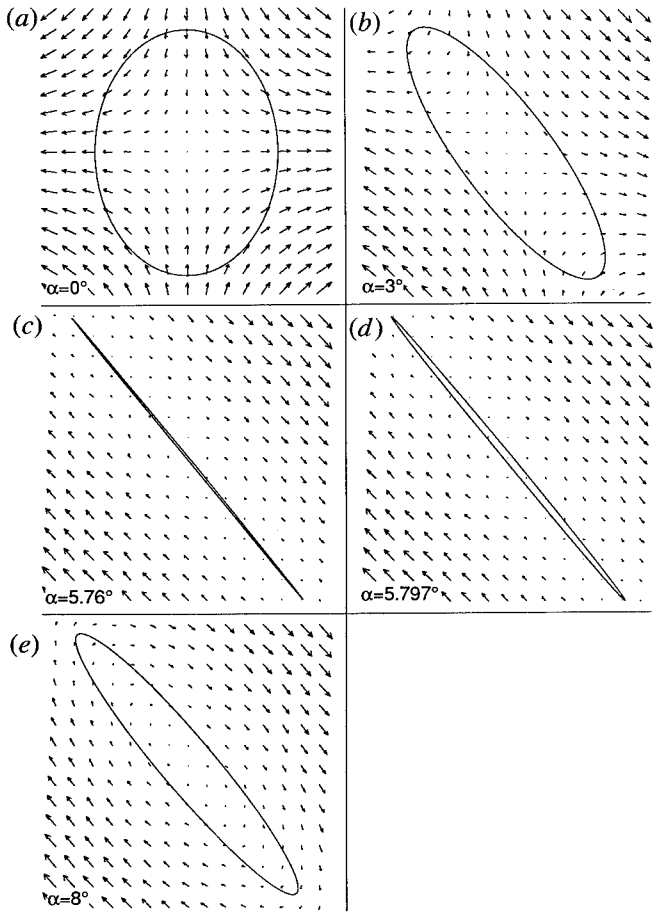


Fig. 3—Behavior of displacement fields for the same cases illustrated in Fig. 2. The inscribed elliptical shapes are surfaces of constant magnitude of the displacement vector; *i.e.*, their dimensions are inverse to the size of the displacement. As the orientation relationship approaches 5.76 deg, the aspect ratio increases, and the major axis approaches the invariant-line direction.

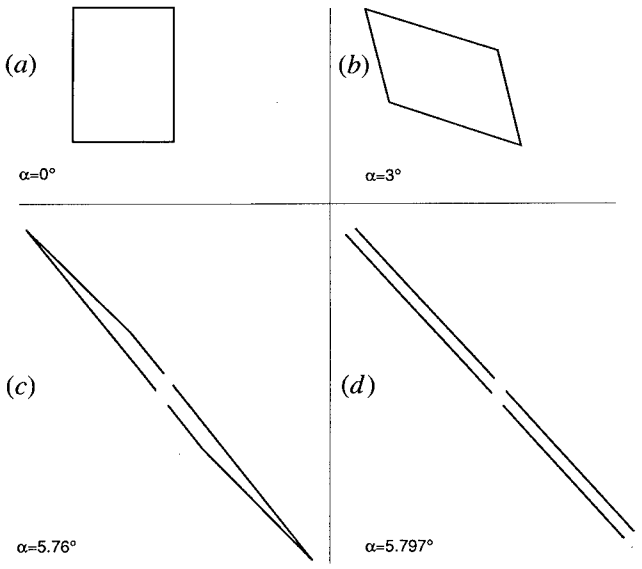


Fig. 4—Precipitate shapes bounded by eigenvectors of the transformation with aspect ratios given by the eigenstrains. The eigenvectors are those in Fig. 2. At 5.76 deg rotation, there are two eigenvectors and hence two corresponding facets, but the aspect ratio is infinite. At 5.797 deg rotation, there is only one degenerate eigenvector, and the shape is indeterminate. At 8 deg rotation, no eigenshape exists, because there are no real eigenvalues.

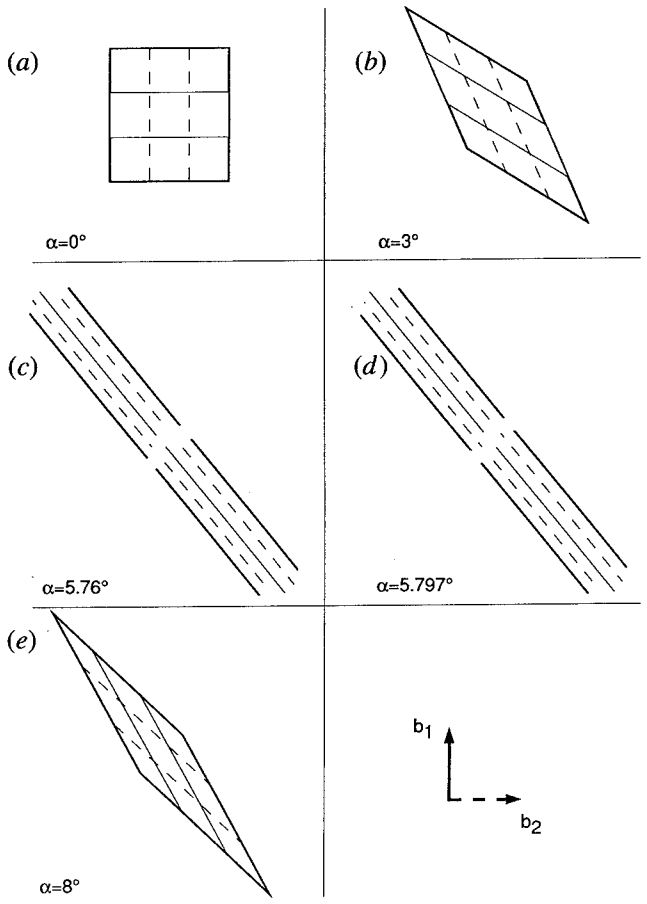


Fig. 5—Precipitate shapes bounded by dislocation loops with Burgers vectors $\mathbf{b}_1$, $\mathbf{b}_2$ parallel to the coordinate axes, illustrated for the same cases as shown in Figs. 1 through 4. As the lattice rotation approaches 5.76 deg, the loop arrays rotate toward each other and become parallel to the invariant line (*c*), and the aspect ratio increases, as seen for the other two criteria.

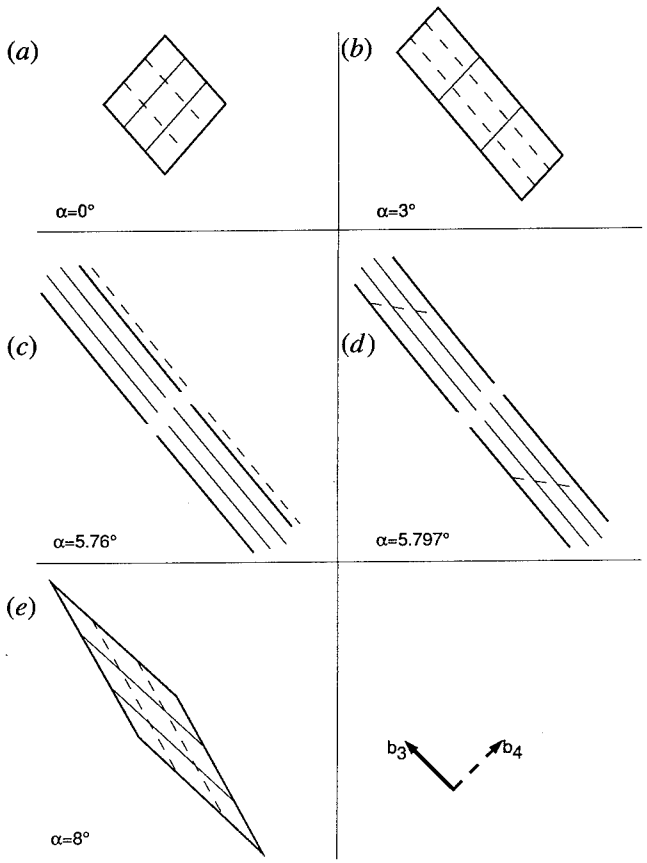


Fig. 6—Precipitate shapes bounded by dislocation loops with Burgers vectors $\mathbf{b}_3$, $\mathbf{b}_4$ diagonal to the coordinate axes, illustrated for the same cases as shown in Figs. 1 through 5. As the lattice rotation approaches 5.76 deg, the behavior is similar to that shown in Fig. 5, but the dislocation spacing is different. At a given orientation relationship, aspect ratios and shape alignments are similar in Figs. 3 through 6.

$\mathbf{e}_2$ is small, and the transformation in Figure 3(c) is close to a simple shear.

C. *Inverse Strain Shape*

The most elementary principle governing precipitate shapes is the postulate that their dimensions tend to be inverse to the magnitude of the transformation strain.[9,12] Thus, inclusions tend to be long in directions of small strain and short in directions of large strain. (Conversely stated, the transformation strain tends to be small in directions of large extent. We assume here a linear inverse relationship to help illustrate the principle, even though the actual form of the dependence will be more complex and has, in fact, not been established for a general inclusion.) Thus, a precipitate that has formed by a transformation **A**, for example, the Bain distortion

$$\mathbf{B} = \begin{pmatrix} a & & \\ & a & \\ & & c \end{pmatrix} \qquad [3]$$

where a and c are the principal deformations, would be expected to have a shape with aspect ratio $|1 - c|/|1 - a|$. According to this principle, the shape would be an ellipsoid, given by the condition that the magnitude of the displacement be constant:

$$|\mathbf{s}| = |(\mathbf{A} - \mathbf{I})\mathbf{r}| = \text{const} \qquad [4]$$

The corresponding elliptical shapes derived from Eq. [4] for the example under consideration are inscribed in Figure 3, where it can be seen that on each ellipse, the length of displacement vectors remains constant. Notice how the aspect ratio increases and the long axis rotates toward the invariant line as the orientation relationship deviates from 0 deg. Even the small 3 deg change in lattice orientation between Figures 3(a) and (b) changes the orientation of the inclusion by about 35 deg and increases its aspect ratio by 70 pct. At the invariant-line orientation in Figure 3(c), the aspect ratio tends to become infinite and the displacement vectors parallel to each other. At a lattice rotation of 8 deg, the optimum shape is again a finite ellipse with its major axis near the invariant-line direction. However, the displacement field has a net rotation component, and even the long axis of the inclusion undergoes a rotation.

D. *Eigenplane Shape*

An alternative suggestion for the optimum shape is based on the postulate that one plane and one direction

within that plane remain unrotated. Ryder and Pitsch used this criterion to predict orientation regions for grain boundary precipitates.[7] Luo and Weatherly subsequently extended this to include more than one plane, implying that inclusions tend to exhibit facets that are defined by unrotated directions (eigenvectors) of the transformation.[13] Thus, it may be postulated that inclusions tend to be faceted on unrotated planes (eigenplanes) of the transformation, where each eigenplane is defined by two eigenvectors (if, in fact, the transformation is one that has three real eigenvectors). The shape predicted on that criterion will be a parallelepiped with edges parallel to the eigenvectors and, hence, faces parallel to the eigenplanes. Its aspect ratio would tend to be inverse to the magnitude of the eigenstrains.

For the case of the example considered here, the corresponding shapes are illustrated in Figure 4. The rectangle shown in Figure 4(a) has edges along the principal axes. The aspect ratio inverse to the eigenstrains becomes $0.12/0.09 = 1.3$. The eigenplane shapes for the other orientation relationships can be derived similarly from the eigenvectors shown in Figure 2, with the aspect ratios given by the corresponding eigenstrains. This principle does not predict a shape for the case of 8 deg rotation because no eigenvectors exist.

E. *Dislocation Loop Shape*

A third proposal for the optimum shape of inclusions is embodied in Bilby's theory of surface dislocations[14] and Bollmann's *O*-lattice theory,[15] which tries to find a network of dislocations that accommodates a transformation strain at the interface between two crystalline regions. An optimum interface is one that lies parallel to planes of the *O*-lattice cell. An inclusion that undergoes a transformation $\mathbf{A}$ can be accommodated by a network of three dislocations with noncoplanar Burgers vectors whose combined effect is the deformation $\mathbf{A}^{-1}$ which exactly cancels the transformation $\mathbf{A}$. Because dislocations cannot end in a crystal, they must form closed loops around the inclusion. Forwood and Clarebrough have shown that in the *O*-lattice formalism, these dislocations form regularly spaced planar loops.[16] Each array of planar loops can be described as an invariant plane strain[12] given by the outer product $\mathbf{b} * \mathbf{n}$, where $\mathbf{b}$ is the Burgers vector and $\mathbf{n}$ the loop plane normal. The deformation resulting from the three sets of dislocation loops enveloping the inclusion is thus

$$\mathbf{A}^{-1} = \mathbf{I} - \mathbf{b}_1 * \mathbf{n}_1 - \mathbf{b}_2 * \mathbf{n}_2 - \mathbf{b}_3 * \mathbf{n}_3 \qquad [5]$$

For a given set of noncoplanar Burgers vectors $\mathbf{b}_i$, the corresponding loop plane normals $\mathbf{n}_i$ are found as[17]

$$\mathbf{n}_1 = (\mathbf{I} - \mathbf{A}^{-1})^T \mathbf{b}_1^* \qquad [6]$$

where

$$\mathbf{b}_1^* = \frac{\mathbf{b}_2 \times \mathbf{b}_3}{(\mathbf{b}_2 \times \mathbf{b}_3) \cdot \mathbf{b}_1} \qquad [7]$$

The loop plane normals $\mathbf{n}_2$ and $\mathbf{n}_3$ are obtained by permutation of the indices. The spacing between loops in the ith array is[18]

$$d_i = \frac{1}{|\mathbf{n}_i|} \qquad [8]$$

If the optimum shape is one bounded by dislocation loop planes (equivalent to *O*-lattice planes), then a parallelepiped will be predicted, or, for the two-dimensional case, a parallelogram. It is interesting to note that the loop plane $\mathbf{n}_1$ in Eq. [6] is independent of the Burgers vector $\mathbf{b}_1$ to ensure that the three sets of dislocation loops are independent. The aspect ratio of each shape is assumed to be proportional to the ratio of the loop spacings. Figures 5 and 6 illustrate this for the cases considered in this section, assuming two alternative sets of dislocation Burgers vectors. (For the two-dimensional case, $\mathbf{b}_1^*$ in Eq. [7] is given by $\mathbf{b}_1^* = \mathbf{b}_2^\perp/(\mathbf{b}_2^\perp \cdot \mathbf{b}_1)$, with $\mathbf{b}_2^\perp \cdot \mathbf{b}_2 = 0$.) Figure 5 shows the predicted shapes for the different orientation relationships considered in this section, assuming dislocations with Burgers vectors $\mathbf{b}_1$, $\mathbf{b}_2$ along the principal axes. As the lattice rotation α increases, the aspect ratio increases, and the loop planes approach each other. At the invariant-line orientation, the two sets of loops are exactly parallel to each other and to the invariant line. However, the loop plane spacings d_1 and d_2 are slightly different. Notice that dislocation loops are predicted, even though the displacement is not parallel to a lattice translation anywhere. However, the combination of the two sets of dislocations precisely cancels the displacement field. The aspect ratio is infinite since the two sets of dislocation loops do not intersect.

At 5.79 deg rotation, both sets of dislocation loops deviate by about 0.2 deg from the invariant-line direction, thus enclosing an angle of 0.4 deg. At 8 deg lattice rotation, this angle has increased to 17 deg, producing a shape and orientation very similar to that for $\alpha = 3$ deg but with the two sets of dislocations interchanged.

Figure 6 shows the same case for a different pair of Burgers vectors $\mathbf{b}_3$, $\mathbf{b}_4$ of the accommodating dislocations. The effect of changing Burgers vectors on the shape is most noticeable for small lattice rotations and disappears for the invariant-line orientation. However, even though at 5.76 deg rotation the loop planes are again both parallel to the invariant line, their spacings are now vastly different because $\mathbf{b}_3$ is nearly parallel to the direction of the strain (Figure 3(c)) and $\mathbf{b}_4$ nearly perpendicular. Most dislocation loops will therefore be of the $\mathbf{b}_3$ type, shown as solid lines. The dislocation spacings for the two arrays are in the ratio of about 1:300.

In comparing the overall shapes, *i.e.*, the aspect ratio and alignment of the long axis, it is found that for the same orientation relationship, the predicted shapes are quite similar for all three criteria. This can be seen, for example, by comparing the shapes for the 3 deg angle of rotation in Figures 3 through 6. Although the facets are very different, the long precipitate axis always lies near the invariant line, and the aspect ratio is near 3:1. This is particularly apparent for the invariant-line orientation, where all three criteria and both sets of Burgers vectors predict an inclusion shape that is infinitely elongated along the irrational invariant-line direction.

F. *The Three-Dimensional Case*

Although, for simplicity, the three principles discussed in Sections C through E were illustrated with

two-dimensional schematics, they are equally valid for the three-dimensional case and can be transferred directly.

The first principle of a shape that tends to be inverse to the magnitude of the strain predicts an ellipsoid for the three-dimensional case, with major axes given by the principal strains.

The second principle of a shape bounded by eigenplanes applies only if three real eigenvalues exist, a condition that is met only by a limited range of orientation relationships. For the case illustrated above, this range is 5.797 deg. Any further lattice rotation eliminates the eigenvectors and leads to imaginary eigenvalues. For the Bain correspondence between face-centered cubic (fcc) and body-centered cubic (bcc) lattices, this range is about 11 deg.[7] In fact, the condition of Ryder and Pitsch[7] (one unrotated plane and an unrotated direction in that plane) requires that there be three real eigenvalues.

The third principle of dislocation loop planes bounding the optimum shape is also directly applicable to three dimensions, using the procedure given in Eqs. [5] through [7]. The inclusion will be enveloped by three arrays of planar dislocation loops whose planes will depend on the triplet of (noncoplanar) Burgers vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$. The shape defined by the loop planes is a parallelepiped.

Thus, all principles outlined in this section will predict three-dimensional precipitate shapes for any given orientation relationship. Some of the underlying differences were illustrated by showing predicted shapes for several different orientation relationships. However, as mentioned in Section I, many precipitation systems are actually observed to exhibit needle- or lath-shaped precipitates along invariant-line directions. This is therefore the most important orientation relationship to consider.

As seen in Figures 2 through 6, the aspect ratio tends to infinity as the orientation approaches the critical value, predicting a needle-shaped precipitate regardless of the criterion used. However, for a three-dimensional inclusion, the cross-sectional shapes, normal to the needle axis, will be different. The first principle predicts an elliptical cross section, while the second criterion demands a parallelogram given by the projection of the eigenvectors $\mathbf{e}_2$ and $\mathbf{e}_3$ along the direction $\mathbf{e}_1$ of the invariant line.

For the third principle, there will be three sets of dislocation loops, with the invariant line as the common zone axis. Thus, all dislocations are parallel to the needle axis. However, they generally have three different Burgers vectors, and they lie on different loop planes. According to the third principle, precipitate cross sections normal to the invariant-line direction can therefore have three distinct facets, one for each loop plane in this zone, leading to an unusual cross-sectional shape bounded by three, rather than only two, separate facets. The three independent arrays of dislocation loops which together exactly cancel the transformation **A** rotate toward a common zone as the invariant-line orientation is approached.

Equation [6] shows that at the critical orientation, all loop planes lie in the zone of the invariant line, regardless of their Burgers vectors. Only the individual array spacings will depend on the operating triplet of Burgers vectors. This behavior is seen in Figures 5 and 6, where the two dislocation arrays rotate toward each other and become parallel to the invariant line at the critical lattice rotation of 5.76 deg. In the three-dimensional case, this process may be thought of as an infinite elongation of the parallelepiped (defined by the three dislocation loop planes in Eq. [6]) along its body diagonal, until all three faces lie in the common zone of the body diagonal which has become parallel to the invariant line. For different dislocation Burgers vectors, the shape is identical, but the array spacings are different. Of the two Burgers vectors in Figure 6, one is almost exactly aligned with the single direction of the displacement (Figure 3). The dominant dislocation loop array has this Burgers vector and almost perfect shear loop character. The second array has a spacing that is almost 300 times as large, and has near-edge loop character.

It should be noticed that the three sets of dislocation loops exactly cancel the transformation strain, even if none of the displacements of the transformation are aligned with a Burgers vector. For example, to accommodate an invariant plane strain with irrational plane and direction of displacement, three arrays of loops parallel to the irrational invariant plane are added in the right proportion to produce the irrational direction of displacement from the sum of three rational Burgers vectors.

G. *Implications for Experimental Studies*

Experimental investigation of precipitate shapes, orientation relationships, and interface structure is particularly difficult for invariant-line precipitates. As illustrated in Section B, a very small change in orientation relationship can make a very large difference in the crystallography. For example, a rotation of 0.03 deg makes the difference between an invariant-line strain and a transformation with degenerate eigenvectors. This sensitivity is highest near the invariant-line condition. Since the orientation relationship cannot be determined with an experimental accuracy of 0.03 deg, it is not possible to rule out either case from diffraction measurements. On the other hand, the invariant-line direction varies rapidly with a very small change in orientation relationship. For large aspect-ratio precipitates, the corresponding change in needle axis can be measured with sufficient accuracy to distinguish between different cases.

Similarly, an experimental determination of the dislocation Burgers vectors is extremely difficult because the loops are often closely spaced and have small anisotropic strain fields. In addition, an image matching analysis applied to fcc/bcc interfaces in Cu-Fe alloys has shown that the Burgers vectors of interfacial dislocations are not lattice vectors, as commonly assumed, but vectors of the displacement shift complete lattice.[16]

Although, for the three-dimensional case, the loops all lie in the zone of the invariant line, different loop planes will be predicted for different sets of Burgers vectors. It is thus important to measure cross-sectional shapes carefully when making comparisons between observed and predicted morphologies. In Cu-Cr alloys, three planar

facets were found to characterize the cross-sectional shape of bcc Cr laths in an fcc Cu-rich matrix.[19] These facets corresponded to a triplet of Burgers vectors in three close-packed $\{111\}_{Cu}$ planes, in agreement with the predictions of the third principle. Two of these facets were also parallel to eigenplanes of the transformation, as demanded by the second shape criterion.

H. *Orientation Relationship*

Finally, it should be noted that for the three-dimensional case, the orientation relationship is not uniquely determined by the invariant-line criterion, even if its location on the cone of unextended lines is known. Experimentally, it is found that the invariant line lies at the intersection of a close-packed plane with the cone of unextended lines and that the corresponding close-packed planes remain parallel.[13,19,20,21] This is unlike the case of martensite crystallography, where usually no rational elements remain parallel. It has been suggested that this preference for a parallel set of low-index rational lattice planes serves to maximize coincidence in an interface.[12] Thus, the observed crystallography is usually one in which a set of corresponding low-index planes are parallel to each other and twisted to contain an invariant line.

This preference is the reason for the choice of transformation strains used in Figures 1 through 6. The principal strains are those characteristic for Cu-Cr in the Nishiyama–Wassermann orientation relationship. The rotation necessary to produce an invariant line with parallel close-packed planes is 5.76 deg, 0.5 deg beyond that necessary for the Kurdjumov–Sachs orientation relationship. This has recently been confirmed experimentally and correlated with careful observations on the precipitate morphology.[19]

III. SUMMARY

The link between orientation relationship and optimum shape of a solid inclusion is investigated and illustrated schematically. Three alternative criteria for the optimum shape of precipitates are presented and compared: the principle that precipitate dimensions tend to be inverse to the magnitude of the transformation strain; the postulate that precipitates are bounded by unrotated planes (eigenplanes); and the proposal that interfaces are parallel to the planes of three independent dislocation loop arrays necessary to accommodate the transformation strain completely. The results of this investigation are as follows.

(1) It is shown that, although the degree of faceting and type of facets differ, all three criteria predict broadly similar shapes. Near an invariant-line orientation relationship, the difference between the three criteria is minimized.

(2) Small changes in orientation relationship can lead to large changes in optimum shape and orientation of an inclusion.

(3) Many precipitates prefer an invariant-line orientation relationship and take the shape of a needle or a lath, aligned along the invariant line. The direction of an invariant line and the corresponding orientation relationship depend on the lattice parameters and are usually irrational. The orientation relationship usually maintains a pair of parallel low-index planes which contain an invariant line.

(4) The cross-sectional shape predicted for invariant-line precipitates is an ellipse for the first criterion, a parallelogram for the second criterion, and bounded by three planar facets for the third criterion.

ACKNOWLEDGMENT

I am grateful to K.H. Westmacott for helpful discussions and a critical reading of the manuscript and to H.I. Aaronson for a stimulating conference that rekindled my interest in this topic. This work is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the United States Department of Energy under Contract No. DE-ACO3-76SFOOO98.

REFERENCES

1. H.I. Aaronson: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, p. 387.
2. K.C. Russell, M.G. Hall, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1974, vol. 5, pp. 1503-05.
3. C.M. Wayman: *Introduction to the Crystallography of Martensitic Transformations*, MacMillan, New York, NY, 1964.
4. U. Dahmen and K.H. Westmacott: in *Proc. Int. Conf. on Phase Transformations '87*, pp. 511-14.
5. J.S. Bowles and W.J. McG. Tegart: *Acta Metall.*, 1955, vol. 3, pp. 590-97.
6. H.M. Otte and T.B. Massalski: *Acta Metall.*, 1958, vol. 6, pp. 494-99.
7. P.L. Ryder and W. Pitsch: *Acta Metall.*, 1966, vol. 14, pp. 1437-47.
8. A. Crosky, P.G. McDougall, and J.S. Bowles: *Acta Metall.*, 1980, vol. 28, pp. 1495-1504.
9. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-73.
10. U. Dahmen, P. Ferguson, and K.H. Westmacott: *Acta Metall.*, 1984, vol. 32, pp. 803-10.
11. J.W. Christian: *The Theory of Phase Transformations in Metals and Alloys*, Pergamon Press, Oxford, United Kingdom, 1975.
12. G.B. Olson and M. Cohen: in *Dislocations in Solids*, F.R.N. Nabarro, ed., North-Holland, New York, NY, 1986, vol. 7, p. 297.
13. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, pp. 1963-72.
14. B.A. Bilby: in *Rept. of the Conf. on Defects in Crystalline Solids*, The Physical Society, London, 1955, p. 124.
15. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer, Berlin, 1970.
16. C.T. Forwood and L.M. Clarebrough: *Phil. Mag.*, 1989, vol. 59, pp. 637-65.
17. K.M. Knowles: *Phil. Mag. A*, 1982, vol. 46, pp. 951-69.
18. C.T. Forwood and L.M. Clarebrough: *Electron Microscopy of Interfaces in Metals and Alloys*, Adam Hilger, New York, NY, 1991.
19. C.P. Luo, U. Dahmen, and K.H. Westmacott: *Acta Metall. Mater.*, in press.
20. M.G. Hall, H.I. Aaronson, and K.R. Kinsman: *Surf. Sci.*, 1972, vol. 31, pp. 257-74.
21. C.P. Luo and G.C. Weatherly: *Phil. Mag.*, 1988, vol. 58, pp. 445-62.

The Invariant Line and Precipitate Morphology in Fcc-Bcc Systems

G.C. WEATHERLY and W.-Z. ZHANG

Second-phase precipitates in many face-centered cubic–body-centered cubic (fcc–bcc) systems (*e.g.*, Ni-Cr, Cu-Cr, Fe-Cu, and α-γ stainless steels) have a lath-shaped morphology, the long axis of the lath being an invariant line of the transformation. The invariant line direction and major (habit) facet plane of the product phase can be predicted by O-lattice (O-line) models. For *N-W*- and *K-S*-oriented precipitates, the habit plane is shown to be an unrotated plane of the transformation. This contains a single set of dislocations lying parallel to the invariant line, with their Burgers vector in the habit plane. Structural ledge models for the habit-plane interface also are considered. For the range of lattice parameter ratios of interest in this study, structural ledge and O-line models can make almost identical predictions as to the optimum habit plane. A variety of elasticity calculations for the energy of fully constrained or fully relaxed precipitates is presented. These models are shown to have limited predictive capabilities. It is suggested that better atomic matching along or near to the invariant line direction might explain the preference for *K-S*-related precipitates in many systems.

I. INTRODUCTION

ALTHOUGH the importance of the invariant line in martensite transformations was first appreciated some forty years ago,[1,2] its role in diffusion-controlled phase transformations involving close-packed [hexagonal close-packed (hcp) or face-centered cubic (fcc)] and body-centered cubic (bcc) parent and product phases was only recognized thirty years later.[3,4,5] With hindsight, this is somewhat surprising, as it now is clear that both types of transformation can share a common lattice correspondence (the Bain strain, B) and a similar or identical rigid body rotation, R, with the invariant line $\mathbf{x}$ given by

$$\mathbf{x} = A\mathbf{x} \qquad [1]$$

where $A = RB$. However, the two types of transformation differ in two important respects. The correspondence of lattice sites in a martensitic transformation must involve a correspondence of atom positions; *i.e.*, each atom in the parent phase is predestined to move to a unique site in the product phase as the transformation front passes. In a diffusion-controlled transformation, this clearly will not be the case. The second difference lies in the role of the lattice invariant shear. This shear is an integral part of nearly all martensitic transformations and leads to an invariant plane strain shape change for the product phase. Although there may be a lattice invariant shear in some diffusional phase transformations,[5] there is no strong evidence for one in any of the systems of interest here, *viz.*, Cu-Cr,[6,7,8] Ni-Cr,[9] Fe-Cu,[10,11] and two-phase α-γ stainless steels.[12,13,14] In the analysis which follows, we will assume that the matrix A in Eq. [1] is fully described by B and one (or more) rigid body rotations R.

Each of the four systems, Cu-Cr, Ni-Cr, Fe-Cu, and α-γ stainless steel, has a lattice parameter ratio (a_f/a_b) close to 1.25, the value for pure Fe. They also share a number of common crystallographic and morphological features.

(a) The product phase has a lath-shaped morphology. The long direction of the lath is the invariant line ($\mathbf{x}$ in Eq. [1]), and the lath is bounded by two or more well-developed facet planes.
(b) The orientation relationship is close to either Nishiyama-Wasserman[15] or Kurdjumov-Sachs.[16] In some systems (*e.g.*, Ni-Cr,[9]) the precipitates have a unique orientation relationship, while in others (*e.g.*, Cu-Cr[6,7,8] and α-γ stainless steels,[12,13,14]), a range of orientation relationships has been reported.
(c) For *K-S*-related precipitates, the invariant line $\mathbf{x}$ is always close to the common close-packed direction in the two phases ($\langle 1\bar{1}1\rangle_b$ or $\langle 1\bar{1}0\rangle_f$). If the orientation relationship varies in a system, each particle grows along a unique direction which can be predicted from Eq. [1] once R has been determined from experimental data.
(d) All sets of misfit dislocations or ledges observed at the facet planes lie parallel to the invariant line.[8,13,17]

Given this wealth of experimental information, one would anticipate that rapid progress might have been made in understanding the factors which control the selection of $\mathbf{x}$ and the development of facets bounding the precipitate. This is not the case. Although several different hypotheses have been advanced, based on elastic strain energy minimization,[3,11,18,19,20] O-lattice models,[8,9,21,22] or structural ledge models,[6,23,24,25] none of them is able to account for all of the experimental observations. Each of these models is considered below. Our discussion is restricted to transformations in fcc-bcc systems with a ratio $a_f/a_b \sim 1.25$ (a_f and a_b are the lattice parameters of the fcc and bcc phases, respectively).

G.C. WEATHERLY, Professor, and W.-Z. ZHANG, Postdoctoral Fellow, are with the Department of Materials Science and Engineering, McMaster University, Hamilton, ON L8S 4M1, Canada.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

II. ELASTIC STRAIN ENERGY MODELS

The starting point for any discussion of the elastic strain energy of the transformation product is the Bain strain (B) given by

$$B = \begin{bmatrix} \eta_1 & 0 & 0 \\ 0 & \eta_1 & 0 \\ 0 & 0 & \eta_2 \end{bmatrix} \qquad [2]$$

with $\eta_1 = \sqrt{2}\,\eta_2$ and $\eta_2 = a_b/a_f$, referred to axes $[\bar{1}10]$, [110], and [001] in the fcc lattice. If $\eta_1 > 1$ and $\eta_2 < 1$, the minimum elastic strain energy of a fully constrained particle corresponds to a plate-shaped product in a state of uniaxial stress[5,26] with the habit plane inclined at an angle

$$\theta = \tan^{-1} \sqrt{-\frac{1+\nu}{\eta + 1 + \nu}} \qquad [3]$$

to [001], where ν is Poisson's ratio and $\eta = \eta_2 - \eta_1/1 + \eta_1$.

Equation [3] has been used successfully to predict the habit plane of plate-shaped precipitates in a number of different alloy systems[5,26] where the strains are relatively small, but this approach does not explain any of the observations in systems with $a_f/a_b \sim 1.25$.

Most authors have concentrated instead on the role of the invariant line and followed the approach first suggested by Dahmen.[3] It is tacitly or explicitly assumed that the rotation to an invariant line position leads to a reduction in the elastic strain energy of the product phase. However, as the following simple example shows, if the rotation is in the plane of a plate-shaped particle, one must be very careful in reaching such a conclusion. Consider a particle in the *N-W* orientation relationship. This position can be obtained by rotating 9.74 deg about $[110]_f$ from the Bain relationship given in Eq. [2]. After transformation to a new set of axes defined by $[1\bar{1}\bar{2}]_f$, $[110]_f$, and $[1\bar{1}1]_f$ in the fcc lattice, the transformation matrix becomes

$$B' = \begin{bmatrix} 1+\varepsilon_1 & 0 & 0 \\ 0 & 1+\varepsilon_2 & 0 \\ \gamma_{31} & 0 & 1+\varepsilon_3 \end{bmatrix} \qquad [4]$$

For the systems of interest here, the strains $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$; an invariant line can be found by rotating about the normal to the common close-packed plane, $(1\bar{1}1)_f$ and $(101)_b$. This is the procedure adopted by Dahmen[3] and followed by Fuji *et al.*,[11] Kato and Mishima,[19] and Kato.[20] If the product of the transformation is a thin plate lying on the close-packed plane (or is an epitaxial thin film deposited on a $(1\bar{1}1)_f$ substrate), the elasticity problem reduces to one of plane stress, as the components ε_3 or γ_{31} do not contribute to the strain energy in this case. If $n_1 = 1 + \varepsilon_1$ and $n_2 = 1 + \varepsilon_2$, the two-dimensional (2-D) invariant line transformation becomes

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix} \qquad [5]$$

where $\cos\theta = (1 + n_1 n_2)/(n_1 + n_2)$ (Dahmen[3]).

An arbitrary unit vector $\mathbf{l}$ lying in the interface can be expressed as

$$\mathbf{l} = \cos\gamma \mathbf{v}_m + \sin\gamma \mathbf{v}_I \qquad [6]$$

with $\mathbf{v}_I$, a unit vector parallel to the invariant line, given by Kato[20] as

$$\mathbf{v}_I = \left[\sqrt{\frac{n_2^2 - 1}{n_2^2 - n_1^2}}, \quad \sqrt{\frac{1 - n_1^2}{n_2^2 - n_1^2}} \right] \qquad [7]$$

while $\mathbf{v}_m$ is perpendicular to $\mathbf{v}_I$ (Figure 1).

The displacement of the unit vector $\mathbf{l}$ after the transformation is

$$\mathbf{b} = \mathbf{l} - A\mathbf{l} = \cos\gamma(I - A)\mathbf{v}_m \qquad [8]$$

The displacement of $\mathbf{l}$ is parallel to $(I - A)\mathbf{v}_m$, and the maximum displacement occurs when $\gamma = 0$; *i.e.*, $\mathbf{l}$ is parallel to $\mathbf{v}_m$. If the maximum value of $\mathbf{b}$ ($\mathbf{b}_{max}$) is expressed relative to a set of axes lying along the normal to and parallel to the invariant line respectively, then its components are

$$\mathbf{b}_{max} = [1 - n_1 n_2, \quad -\sqrt{(n_2^2 - 1)(1 - n_1^2)}] \qquad [9]$$

One can also show that $|\mathbf{b}_{max}| = |\varepsilon_1| + |\varepsilon_2|$. This implies that if the misfit were relieved by one or two sets of necessary dislocations, the density of dislocations required before and after rotation would be the same (assuming the magnitudes of the Burgers vectors are identical, and the misfit before rotation is relieved by two orthogonal sets of edge dislocations).

We can now compare the elastic strain energies before and after the rotation to produce an invariant line, using the plane stress relationship,

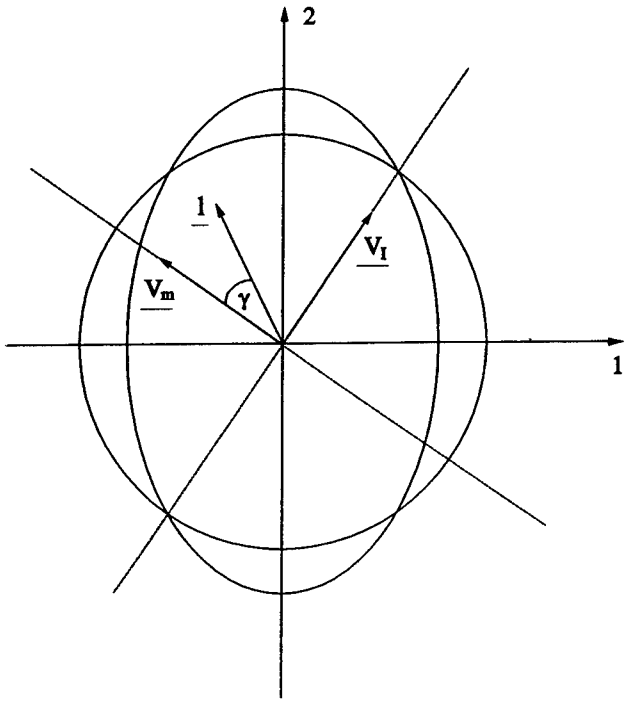


Fig. 1—The circle-to-ellipse transformation illustrating the orientations of axes (1,2) and (V_I,V_m) used for the 2-D invariant line model.

$E_{el} = 1/2\ (\sigma_1\varepsilon_1 + \sigma_2\varepsilon_2 + \tau_{12}\gamma_{12})$ for the elastic strain energy, with

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \tau_{12}\end{bmatrix} = \frac{2\mu}{1-\nu}\begin{bmatrix}1 & \nu & 0\\ \nu & 1 & 0\\ 0 & 0 & (1-\nu)/2\end{bmatrix}\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \gamma_{12}\end{bmatrix} \quad [10]$$

From Eqs. [4] and [10], we find that *before* the rotation

$$E_{el} = \frac{\mu}{1-\nu}[\varepsilon_1^2 + \varepsilon_2^2 + 2\nu\varepsilon_1\varepsilon_2] \quad [11]$$

while *after* the rotation, from Eqs. [9] and [10],

$$E_{el} = \frac{\mu}{1-\nu}\left[\varepsilon_1^2 + \varepsilon_2^2 + \frac{1+\nu}{2}\varepsilon_1\varepsilon_2(1 + 2\varepsilon_1 + 2\varepsilon_2 + \varepsilon_1\varepsilon_2) + \left(\frac{3\nu-1}{2}\right)\varepsilon_1\varepsilon_2\right] \quad [12]$$

Since a calculation of elastic strain energy can only depend on terms of $0[\varepsilon^2]$, Eqs. [11] and [12] predict essentially identical values of E_{el}; indeed, if $\nu = 1/3$ the two results are identical.

The analysis just presented clearly illustrates the shortcomings of plane stress (plate-shaped product), elastic strain energy arguments in explaining the selection of the invariant line in fcc ⇆ bcc transformations and in fcc/bcc epitaxial growth. While the orientation relationships and invariant lines predicted from this approach lie close to the observations in a variety of systems,[6–14] the strain energy is independent of whether the habit plane contains an invariant line or not, if the rotation is restricted to lie in the common close-packed plane. However, a further reduction in strain energy can be realized if a plate-shaped particle contains the invariant line, but the rotation of the habit plane is not restricted in the way prescribed above.[5]

There are several limitations to the model presented above. The terms γ_{31} and ε_3 in Eq. [4] should not be ignored. A plate-shaped morphology maximizes the surface energy-to-volume ratio of the particle. On the other hand, a lath- or needle-shaped morphology has a lower surface energy-to-volume ratio and can partly accommodate all the strain components out of the invariant line direction. Intuitively, a lath or needle morphology is the one which makes most effective use of the invariant line direction while trying to accommodate the large displacements in the other directions.

Before considering O-lattice models in Section III, it will be helpful to the arguments developed therein to point out that although Eqs. [11] and [12] lead to essentially identical values of E_{el}, the nature of the displacement field at the interface approaches a simple shear on rotating to the invariant line; *i.e.*, $n_1n_2 \simeq 1$. Before rotation (*i.e.*, at the *N-W* position), a network of three $a_f/2\ \langle 110\rangle$ dislocations, one pure edge, the other two mixed, would be required to accommodate the misfit strains at the common close-packed interface. After rotation to the invariant line position, a single set of near-screw dislocations lying parallel to the invariant line is needed. As shown in Section III, this single set might have a lower energy than the network.

III. O-LATTICE MODELS

Geometric Considerations

The O-lattice (O-line) model has been described a number of times.[9,21,22,27,28,29] In constructing O-lattice models for fcc-bcc transformations with an invariant line, the rules governing allowable displacement fields at any facet plane of the product phase must be considered. These are:

(a) The displacements at any facet must lie in the plane of the invariant normal; and
(b) The displacements at a particular facet plane are aligned in the direction $(I - A^{-1})\mathbf{v}_m$.

Note that this expression for the displacement is consistent with the usual convention adopted in Bollmann's O-lattice model. In Section II, the displacement was expressed as $(I - A)\mathbf{l}$ to be compatible with the definitions of strain used in elastic strain energy calculations. The displacement now need not lie in the facet plane, as was the case for the 2D model considered in Eq. [8]. These two results were first used in the development of the phenomenological theory of martensite crystallography.[1,2] If the displacements at the interface are parallel to a Burgers vector compatible with either the fcc or bcc lattice, the solution to the O-lattice problem then reduces to a set of parallel lines. We assume, following Bollmann,[27] that the interface adopted by nature would be one which conserves the optimum periodic structure. For a facet containing the invariant line, the lowest energy interface then would correspond to one where a single set of parallel lattice dislocations completely accommodates the misfit. Although more complex dislocation arrays could be envisaged at a facet,[34] these would have a larger residual strain and greater energy than an array composed of a single set of lattice dislocations. In the analysis which follows, we seek solutions to the O-line problem where the displacements are parallel to $\mathbf{b}_f$.

Descriptions equivalent to those given in Section II can be defined in a reciprocal lattice formulation.[28,29] This leads to a number of useful relationships. We can define a displacement in the reciprocal lattice $\Delta\mathbf{g}$ as

$$\Delta\mathbf{g} = \mathbf{g}_1 - \mathbf{g}_2 = [I - (A')^{-1}]\mathbf{g}_1 \quad [13]$$

provided that $\mathbf{g}_1$ and $\mathbf{g}_2$ are related by

$$\mathbf{g}_2 = [A']^{-1}\,\mathbf{g}_1 \quad [14]$$

where A' is the transpose of A defined in Eq. [1]. $\mathbf{g}_1$ and $\mathbf{g}_2$ are reciprocal vectors of lattices correlated through the transformation A. A list of low-index planes and directions related by the Bain correspondence is given in Table I. The facet plane containing the O-lines will be normal to a $\Delta\mathbf{g}$, provided the corresponding $\mathbf{g}_1$ (or $\mathbf{g}_2$) is contained in the zone axis of the Burgers vector, $\mathbf{b}_1$ (or $\mathbf{b}_2$), associated with the O-lines.[29] As a result, once the Burgers vector is known the facet plane containing the O-lines can be determined simply from $\Delta\mathbf{g}$.

Table I. Correlated Sets of Low-Index Planes and Directions in the Fcc and Bcc Lattices

Fcc	Bcc	Fcc	Bcc
[0 −1 −1]	[1 −1 −1]	(1 1 −1)	(0 1 −1)
[1 0 −1]	[1 1 −1]	(1 1 1)	(0 1 1)
[1 −1 0]	[2 0 0]	(2 0 0)	(1 1 0)
[0 −1 1]	[1 −1 1]	(1 −1 −1)	(1 0 −1)
[1 0 1]	[1 1 1]	(1 −1 1)	(1 0 1)
[−1 −1 0]	[0 −2 0]	(0 0 −2)	(0 0 −2)
		(0 −2 0)	(1 −1 0)

As there are an infinite number of choices for the invariant line (**x** in Eq. [1]), we have restricted our search to those solutions where at least one set of O-lines exists. Starting from the *N-W* orientation relationship, one can specify three rotation angles (α, θ, and φ) to define the final orientation relationship. The first rotation (α) lies in the common close-packed plane, $(1\bar{1}1)_f \parallel (101)_b$, and is limited to the range from 0 deg to −6 deg. If α = −5.26 deg, $[\bar{1}01]_f \parallel [\bar{1}\bar{1}1]_b$, and the rotation corresponds to an exact *K-S* relationship. The second and third rotation angles (θ, φ) are used to define a small misorientation between the close-packed planes. The rotation axis, defined by θ, lies in this plane but can vary from θ = 0 deg to θ = 180 deg. The rotation angle φ is restricted to the range from −1.5 deg to +1.5 deg. Because of the symmetry of the *N-W* relationship, this angular range covers any orientation relationship lying within ±1.5 deg between the *K-S* and *N-W* positions. The *N-W* and *K-S* orientation relationships are defined by (α, θ, φ) = (0 deg, 0 deg, 0 deg) and (−5.26 deg, 60 deg, 0 deg), respectively.

Figure 2 summarizes the results of searching for possible O-line solutions for Ni-Cr (a_f/a_b = 1.255) and Cu-Cr (a_f/a_b = 1.267) over the angular range noted. A single set of O-lines is found for only three $\langle 110\rangle_f$ Burgers vectors, *viz.*, $[110]_f$, $[011]_f$, and $[10\bar{1}]_f$, lying in $(1\bar{1}1)_f$. These are shown in Figure 2 for each combination of α and θ and are found to correspond to invariant

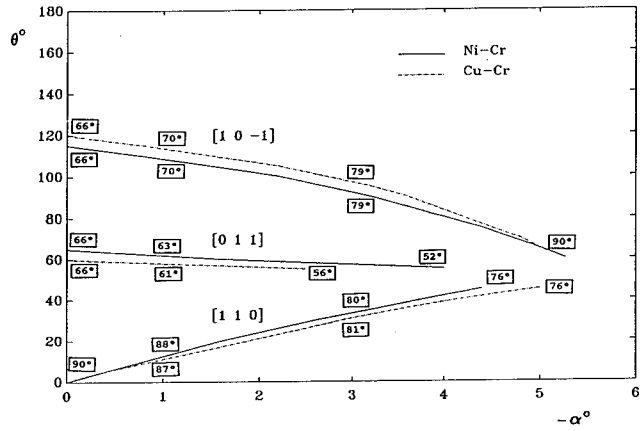


Fig. 2—O-line solutions for Ni-Cr and Cu-Cr in the angular range −6 deg ≤ α ≤ 0 deg and 0 deg ≤ θ ≤ 180 deg, for b_f = $[110]_f$, $[011]_f$, and $[10\bar{1}]_f$. The solid and dotted lines represent the angular range within which O-line solutions were found for the lattice parameter ratios of Ni-Cr and Cu-Cr. The numbers in the rectangular boxes are the angles between the direction of $\mathbf{b}_f$ and the normal to the habit plane. When α = θ = 0 deg [110] and α = −5.26 deg, θ = 60 deg $[10\bar{1}]$, $\mathbf{b}_f$ lies in the habit plane.

normals lying close to $(7\bar{7}6)_f$, $(4\bar{3}3)_f$, and $(3\bar{4}3)_f$, respectively, over the angular range defined in Figure 2. We find that there is *no* solution which simultaneously can be used to predict two *different* sets of O-lines associated with $\langle 110\rangle_f$ Burgers vectors, because $(1\bar{1}1)_f$ is not the plane of the invariant normal for the Cu-Cr and Ni-Cr systems. (The implications of this are considered in this section.) In general, the Burgers vectors of the dislocations do not lie in the plane of the O-lines, but there are two important exceptions to this observation. These two exceptions occur when the first rotation (α) leaves the Burgers vector in the parent lattice (fcc) parallel to its counterpart in the product lattice (bcc); *i.e.*, at the exact *N-W* or *K-S* orientation relationship. The second rotation, defined by (θ, φ), keeps this vector unrotated; *i.e.*, the rotation axis coincides with the Burgers vector. The corresponding facet plane will then be an unrotated plane of the transformation, containing the invariant line and the Burgers vector which defines a single set of dislocations that accommodates the misfit most effectively. We have also found that at these two positions, the unrotated direction of minimum misfit strain exists in the habit plane. Both these orientation relationships lie within ~0.5 deg of *K-S* or *N-W* and were selected to compare with experimental observations in Ni-Cr[9] and Cu-Cr.[8]

The results of the calculations for Ni-Cr and Cu-Cr are compared in Tables II and III. The predictions of the O-line model are almost identical for the two systems at either orientation relationship. The data listed in Tables II and III all can be verified experimentally; *viz.*, the orientation relationship, the invariant line (**x**), the facet planes, the Burgers vectors and spacings of dislocations at facets, and step heights at facet planes can be found by transmission electron microscopy (TEM). The step height should be related to integer multiples of the corresponding $|1/\Delta\mathbf{g}|$ values (Eq. [13]). Three of the correlated planes in Table I, $(111)_f$, $(1\bar{1}1)_f$, and $(0\bar{2}0)_f$ contain the lattice vector $\mathbf{b}_L = a_f/2[\bar{1}01]$, and any one of them can be used to predict the "major" observed facet plane, close to $(1\bar{2}1)_f$.[9] However, each of these three planes has associated with it a different $\Delta\mathbf{g}$. When only one set of O-lines is available, growth steps connecting equivalent habit-plane interfaces will always have a residual strain field associated with them. The "correct" choice for the correlated planes is made by assuming that the residual strain would be kept to a minimum. This corresponds to the condition where the correlated **g** ($\{111\}_f$ type) lies closest to the plane of the invariant normal. For the Cu-Cr and Ni-Cr systems, this leads to the $(111)_f$, $(011)_b$ pair as the optimum choice,

Table II. Predictions of O-Line Model for Near K-S-Oriented Precipitates in Cu-Cr and Ni-Cr Systems

Systems	Cu-Cr	Ni-Cr
a_f/a_b	1.267	1.255
Orientation relationship	$\alpha = -5.26$ deg, $\theta = 60$ deg, $\phi = 0.51$ deg	$\alpha = -5.26$ deg, $\theta = 60$ deg, $\phi = 0.44$ deg
Invariant line (**x**)	$[1.23, 0.13, -1]_f$	$[1.16, 0.09, -1]_f$
Major facet plane A	$(1, -1.70, 1)_f$	$(1, -1.87, 1)_f$
Correlated **gs**, $(1/\lvert\Delta\mathbf{g}\rvert)$	$(1\ 1\ 1)_f/(0\ 1\ 1)_b$, (10.5 Å)	$(1\ 1\ 1)_f/(0\ 1\ 1)_b$, (10.8 Å)
$\mathbf{b}_1$, $(\mathbf{b}_1 \wedge \mathbf{d})$	$[-1\ 0\ 1]_f/2$, (0 deg)	$[-1\ 0\ 1]_f/2$, (0 deg)
Dislocation spacing (l)	9.9 Å	9.7 Å
Minor facet plane B	$(1, 0.21, 1.26)_f$	$(1, 0.24, 1.18)_f$
Correlated **gs**, $(1/\lvert\Delta\mathbf{g}\rvert)$	$(1\ 1\ -1)_f/(0\ 1\ -1)_b$, (8.4 Å)	$(1\ 1\ -1)_f/(0\ 1\ -1)_b$ (8.5 Å)
$\mathbf{b}_1$, $(\mathbf{b}_1 \wedge \mathbf{d})$	$[0\ 1\ 1]_f/2$, (9.7 deg)	$[0\ 1\ 1]_f/2$, (7.1 deg)
Dislocation spacing (l)	13.6 Å	13.2 Å
Minor facet plane C	$(1, -0.41, 1.18)_f$	$(1, -0.42, 1.13)_f$
Correlated **gs**, $(1/\lvert\Delta\mathbf{g}\rvert)$	$(1\ -1\ -1)_f/(1\ 0\ -1)_b$, (5.8 Å)	$(1\ -1\ -1)_f/(1\ 0\ -1)_b$, (5.9 Å)
$\mathbf{b}_1$, $(\mathbf{b}_1 \wedge \mathbf{d})$	$[1\ 1\ 0]_f/2$, (9.7 deg)	$[1\ 1\ 0]_f/2$, (7.1 deg)
Dislocation spacing (l)	19.8 Å	19.1 Å

Table III. Predictions of O-Line Model for Near N-W-Oriented Precipitates in Cu-Cr and Ni-Cr Systems

Systems	Cu-Cr	Ni-Cr
a_f/a_b	1.267	1.255
Orientation relationship	$\alpha = 0$ deg, $\theta = 0$ deg, $\phi = 0.54$ deg	$\alpha = 0$ deg, $\theta = 0$ deg, $\phi = 0.35$ deg
Invariant line (**x**)	$[1, -1, -1.14]_f$	$[1, -1, -1.21]_f$
Major facet plane	$(1, -1, 1.75)_f$	$(1, -1, 1.65)_f$
Correlated **gs**, $(1/\lvert\Delta\mathbf{g}\rvert)$	$(1\ -1\ -1)_f/(1\ 0\ -1)_b$, (6.2 Å)	$(1\ -1\ -1)_f/(1\ 0\ -1)_b$, (6.3 Å)
$\mathbf{b}_1$, $(\mathbf{b}_1 \wedge \mathbf{d})$	$[1\ 1\ 0]_f/2$, (0 deg)	$[1\ 1\ 0]_f/2$, (0 deg)
Dislocation spacing (l)	24.6 Å	23.0 Å

with $|1/\Delta\mathbf{g}|$ = 10.5 Å or 10.8 Å. This should be the minimum height of growth or dissolution steps. This agrees with the experimental observations in Ni-Cr.[17] As noted earlier, the calculations summarized in Table II assume *a priori* that the single dislocation array ($b_L = a_f/2[\bar{1}01]$) at the major facet plane, A, completely accommodates the misfit there. One can use the O-lattice model to describe dislocation arrays at the other two potential side facet planes[9] (B and C in Table II), but these will not completely annul the misfit at these facets. A measure of this is provided by noting the angle between $\mathbf{b}_L$ and the net displacement **d** at the facet plane (Table II). This leaves a residual long-range strain at the side facet plane, which has been detected by TEM.[17]

The predictions of the O-lattice model are in reasonable agreement with the experimental measurements reported on the Ni-Cr system.[9,17] The close-packed planes are predicted to be misoriented by approximately 0.5 deg (Table II), while the experimental TEM observations suggest that these planes are parallel.[9] Figure 3 shows the cross-section of a Cr-rich lath in a Ni-45 wt pct Cr alloy, aged for 10 minutes at 1000 °C. The diffraction pattern superimposed on the image (without rotation) comes from the common zone axis of the $[\bar{1}01]_f$, $[\bar{1}\bar{1}1]_b$ *K-S* orientation relationship. Since the major facet plane, A, contains $[\bar{1}01]_f$, it lies parallel to the incident beam in this figure. The diffraction spots associated with the correlated pairs of planes $(111)_f$, $(011)_b$ and $(0\bar{2}0)_f$, $(1\bar{1}0)_b$ are marked. Both $\Delta\mathbf{g}$ vectors are perpendicular to the facet plane, in agreement with the predictions of Eq. [13]. This observation confirms the existence of a true invariant line, as described earlier in both experimental[9,17] and theoretical studies.[29] The minor facet plane predicted by the analysis, B in Table II, is not quite edge-on in Figure 3. The measured angle between the two facets is 63 deg, compared to a predicted value

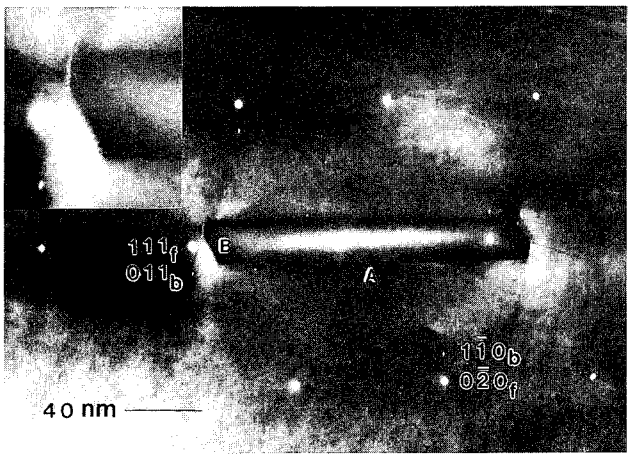


Fig. 3—Ni-45 wt pct Cr alloy, solution treated at 1200 °C, water quenched, and aged for 10 min at 1000 °C. A Cr-rich lath is seen in cross section with the superimposed $[\bar{1}01]_f$, $[\bar{1}\bar{1}1]_b$ diffraction pattern. The correlated diffraction spots $(111)_f$, $(101)_b$ and $(0\bar{2}0)_f$, $(1\bar{1}0)_b$ are marked, with $\Delta\mathbf{g}$ in each case normal to the major facet plane A. The insert (at higher magnification) shows a series of line defects, spacing ~1.2 nm, at the side facet planes. Note also the long-range strain field (black-white contrast) in the matrix associated with the side facet planes.

of 61.8 deg. A third facet also is seen in Figure 3. This is not predicted by the O-line analysis; *i.e.*, it is not facet C in Table II.

Two other features should be noted in Figure 3. A series of finely spaced defects are visible at facet B, lying parallel to the invariant line. The measured spacing of these defects, ~1.2 nm, compares to a predicted value of 1.3 nm. If we are correct in hypothesizing that this array will not completely compensate for the misfit at the side facet, a long-range residual strain field should be observed. This could account for the pronounced black-white contrast effects visible at either end of the faceted particle in Figure 3. On continued aging, this long-range strain field is relaxed by the accumulation of a second set of dislocations[17] (from the matrix) and at the same time, the side facet plane changes its orientation by about 10 deg.[9] The major facet plane shown in Figure 3 does not contain any visible steps or ledges. These are most clearly seen at coarser particles after prolonged aging.[17] Figure 4 shows individual growth or dissolution ledges imaged under diffraction conditions similar to those used in Figure 3. Although the diffraction contrast at the interface is complicated here by Fresnel effects, it is a simple matter to measure the ledge height directly from the image. One finds an average ledge height of 1 nm, in good agreement with the predicted minimum value of 1.1 nm (Table II).

There are several important differences between the predictions made for the near *K-S* (Table II) and *N-W* (Table III) orientations. At the *K-S* orientation relationship, the dislocations at the major facet plane are in near-screw orientation, while at the *N-W* orientation, they are in exact-edge orientation. The invariant line is close to the common close-packed direction, $[\bar{1}01]_f \sim \| [\bar{1}\bar{1}1]_b$, in the *K-S* case but is not very close to either of the principal axes, $[110]_f$ or $[\bar{1}12]_f$, in the *N-W* case. The dislocation spacing in the *N-W* habit plane is about 2.5 times larger than the *K-S* case (24.6 Å *vs* 9.9 Å for Cu-Cr). Finally, it is worth noting that the major facet planes predicted for either orientation of Cu-Cr both belong to a family which can be indexed as being near "{335}." This is fortuitous, but it highlights the importance of correctly establishing the orientation of facet planes relative to the orientation relationship for a particular crystallographic variant.

The results shown in Table III are similar to those published by Hall *et al.*[22] for *N-W*-related Cr laths in a Cu-Cr alloy. Although these authors did not consider O-line solutions explicitly, they were aware that their analysis reduced to the O-line model at certain orientation relationships. Hall and Aaronson[8] report interdislocation spacings (l) of 2.55 nm, with a single set of dislocations aligned along an irrational direction in a "near" $(3\bar{3}5)_f$ habit plane for Cr laths. Although the Burgers vectors of finely spaced arrays of dislocations are difficult to determine, the results found by Hall and Aaronson were also in good agreement with the predictions of the O-line model.

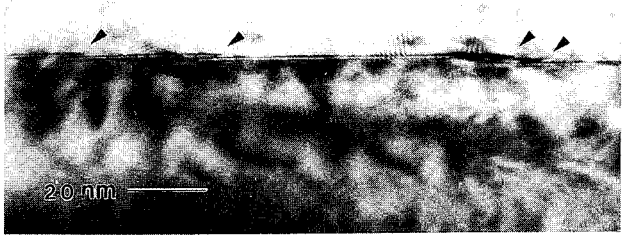


Fig. 4—The alloy from Figure 3, aged for 15 h at 1050 °C. A series of ledges, marked by arrows, are visible at the major facet plane.

IV. INTERFACE DISLOCATION STRAIN ENERGY

In view of the apparent success of O-line models in explaining the results in Cu-Cr[6,8] and Ni-Cr,[9,17] it is of some interest to explore whether strain energy arguments might help to rationalize the choice of the prominent facet planes found in these systems. If a single set of parallel dislocations relieves the misfit at the facet plane, the energy can be evaluated by a method suggested by Bonnet *et al.*.[31,32] Three possible configurations are shown in Figure 5. In the first (Figure 5(a)), the facet is coherent, and the displacement at the interface can be represented by a continuous distribution of infinitesimal dislocations, $d\mathbf{b}$, defined by $d\mathbf{b} = \mathbf{b}\, dx/l$, $-l/2 \le x \le l/2$. In the second (Figure 5(b)), the displacement of the facet can be described by a Somigliana dislocation. Here two discrete dislocations, Burgers vectors $= -\mathbf{b}/2$, located at either end of the facet, are added to the array shown in Figure 5(a). In the third (Figure 5(c)), repetition of the basic Somigliana dislocation unit side by

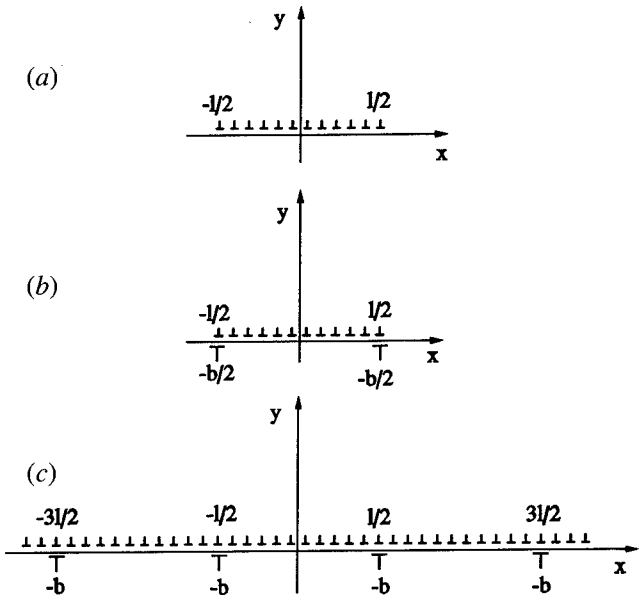


Fig. 5—Three different dislocation arrays used in the text to estimate the strain energy of faceted interfaces in invariant-line strain models. (*a*) The facet is coherent, and the misfit is represented by an array of infinitesimal dislocations. (*b*) The strain field of the facet is described by a Somigliana dislocation.[31] (*c*) The facet is represented by an array of discrete epitaxial dislocations, spacing l, superimposed on the infinitesimal array.

side along the interface leads to the familiar one-dimensional array of epitaxial dislocations, Burgers vector = $-\mathbf{b}$, spacing = l. The energies of the first two arrays* can be evaluated by considering the work done in making the cut at the facet interface in the stress field of the continuous (and discrete) dislocations.[31]

*If a lath-shaped particle were to be bounded on all faces by coherent boundaries or Somigliana dislocations, the total strain energy of the particle would be determined by summing the contribution of each facet with the interaction terms between facets.[31]

The analysis is described in more detail in the Appendix. For the three arrays shown in Figure 5, the corresponding strain energies are

$$E_1 = \frac{\mu b^2}{4\pi(1-\nu)}\left[\ln\left(\frac{R}{l}\right) + \frac{3}{2}\right] \qquad [15a]$$

$$E_2 = \frac{\mu b^2}{8\pi(1-\nu)}\left[\ln\frac{l}{r_0} - 1\right] \qquad [15b]$$

$$E_3 = \frac{\mu b^2}{4\pi(1-\nu)l}\left[\ln\frac{l}{2\pi r_0}\right] \qquad [15c]$$

where r_0, the cut-off radius at the singularity (dislocation core or end of the facet), is assumed to be $\ll l \cdot E_1$ and E_2 are the energies per unit length of a facet of width l, while E_3 is the energy/unit area of the array seen in Figure 5(c). For the coherent array, b in Eq. [15a] is interpreted as the product of l and the misfit strain. Neither the *K-S*- nor *N-W*-related systems satisfy the condition $r_0 \ll l$, as $l \sim 4b$ and $9b$, respectively. More complete expressions for E_1, E_2, and E_3 valid over a wide range of l/r_0 values (including $l/r_0 \gg 1$) can be found in the Appendix and in References 43 and 44. Equations [15a] through [15c] hold for edge dislocations. The results are readily extended to mixed or screw dislocations by replacing $\mu b^2/4\pi(1-\nu)$ with $\mu b^2/4\pi[\sin^2\theta/(1-\nu) + \cos^2\theta]$, where θ is the angle between the Burgers vector and the dislocation line.

Since b/l is the misfit strain at the facet (δ) and $r_0 \sim b$, Eq. [15c] reduces to the familiar form $E \sim -\mu b\delta/4\pi[\ln\delta + A]$ or $E \sim -\mu b\delta/4\pi(1-\nu)[\ln\delta + A]$ respectively for screw or edge misfit dislocations, where A is a constant, of order unity. For a fixed misfit, the screw dislocation array will have a lower strain energy by a factor $1/1-\nu$. This might explain why epitaxial thin films deposited on close-packed fcc or bcc substrates often "rotate" to the invariant line position.[20] As we showed earlier, this rotation does not lower the elastic strain energy of a coherent film, but in a relaxed film, one can demonstrate that the single set of near-screw dislocations required at the invariant line position will have a lower energy than the two (or more) sets of edge or mixed dislocations required before rotation.

Equation [15c] also may be used to compare the energies of the different facet planes. Similar expressions have been used in previous studies to rationalize the choice of habit plane in fcc-bcc systems.[22,30,33,34] However, it is obvious from an inspection of Table II and the results presented by Hall *et al.*[22] that the near $(1\bar{2}1)_f$ habit plane in Ni-Cr does *not* correspond to a minimum in strain energy, even allowing for the fact that the dislocations are in near-screw orientation. Indeed, the dislocations are so closely spaced at the habit plane, $l \sim 1$ nm, that the validity of any continuum strain energy argument becomes highly questionable. The answer might lie in the application of structural or misfit-compensating ledges to the problem, as discussed in the Section V. Before considering these models, however, two key points which emerge from the O-line analysis should be recognized. The three possible Burgers vectors of fcc dislocations compatible with the O-line model for the angular range (α, θ, φ) and a_f/a_b ratios considered in our analysis all lie in the common close-packed plane $(1\bar{1}1)_f/(101)_b$. Secondly, for the *K-S* or *N-W* orientations, the optimum set of dislocations have their Burgers vector lying in the habit plane, which must then be an unrotated plane of the transformation.

V. STRUCTURAL LEDGE MODELS

The structural ledge model was introduced by Hall *et al.*[6] and further developed by Rigsbee and Aaronson[23] to explain the experimental observations of irrational habit planes associated with bcc precipitates in fcc matrices. The close-packed planes $(1\bar{1}1)_f$ and $(101)_b$ were assumed to be parallel, and regions of good atomic matching were sought in successive close-packed planes by introducing an array of regular steps or structural ledges up to 3 atom planes high. In later versions of the model,[35,36] steps up to 6 atom planes high were incorporated into the analysis. It was postulated that these might play a dual role, relieving the misfit strains normal or parallel to the terrace as well as optimizing the regions of good fit in the terrace plane. The energy associated with simple versions of these models has been determined by van der Merwe *et al.*[24] Although the structural ledge and O-lattice models were quite independent developments, recent studies have explored the relationship between these two approaches.[22,37] The application of moiré models, as first described by Bollmann,[27] is a particularly incisive way of analyzing the problem, which we shall follow.

A set of orthogonal axes (x,y,z) is used to define directions lying in and normal to the atomic habit plane, which is assumed to be $(1\bar{1}1)_f \parallel (101)_b$, the x-y plane. Superposition of layers of atoms corresponding to the arrangements in these two planes gives a moiré pattern. If one lattice is translated with respect to the other, the moiré pattern (O-lattice) shifts by an amount which can be simply calculated,[27] but the pattern itself does not change. The relative in-plane displacement ($\mathbf{d}_b$) between pairs of correlated lattice points (in the next bcc plane) can be described by

$$\mathbf{d}_b = \mathbf{b}_f - \mathbf{b}_b - \mathbf{K}\,\Delta d \qquad [16]$$

where $\mathbf{b}_b = A\mathbf{b}_f$ and $\Delta d = d_{(1\bar{1}1)_f} - d_{(101)_b}$. $\mathbf{b}_f$ and $\mathbf{b}_b$ are pairs of correlated lattice points in the fcc and bcc lattices, respectively (see also Eqs. [1] and [14]), $\mathbf{K}$ is a unit vector pointing in the z direction, and $d_{(1\bar{1}1)_f}$ and $d_{(101)_b}$ are the spacings of the close-packed planes.

The corresponding translation of the O-lattice with respect to $\mathbf{b}_b$ is $\mathbf{d}_{0_p}$, given by[27]

$$(I - A)\mathbf{d}_{0_p} = -A\mathbf{d}_b \qquad [17a]$$

or

$$T\mathbf{d}_{0_p} = \mathbf{d}_b \qquad [17b]$$

This translation with respect to the origin becomes

$$\mathbf{d}_0 = \mathbf{d}_{0_p} + \mathbf{b}_b \qquad [18]$$

(Note that in this description we have taken lattice (1) in Bollmann's formulation to be the bcc lattice.) Equations [16], [17], and [18] can be combined to yield

$$\mathbf{d}_0 = -\Delta d\, T^{-1}\mathbf{K} \qquad [19]$$

where $\mathbf{d}_0$ is a vector that defines the shift of the center of the "coherent patches," as described by Rigsbee and Aaronson,[23] in the subsequent layer of $(101)_b$ planes; *i.e.*, it defines the step vector associated with the structural ledge. Equation [19] provides a simple method to determine the orientation of the stepped interface, which is equivalent to the one constructed by the computer graphic technique.[23] This interface is defined by $\mathbf{d}_0$ and an O-lattice vector contained in $(1\bar{1}1)_f/(101)_b$ planes. Since $\mathbf{K}$ is parallel to $[1\bar{1}1]_f$, the interface plane must correspond to the O-lattice plane related to one of $(1\bar{1}\bar{2})_f$, $(121)_f$, or $(21\bar{1})_f$.

Equations [16] through [19] can be used to discuss a number of interesting cases from the literature.

(a) If $\Delta d = 0$ and $(1\bar{1}1)_f \parallel (101)_b$, a true invariant line ($\mathbf{x}$) inclined to the close-packed planes can be predicted from the structural ledge model, the O-lattice model,[27] or the invariant line approach.[25,36] If $\Delta d \approx 0$ and the close-packed planes are parallel, Eq. [19] can be expressed (in approximate form) as either

$$T\mathbf{d}_0 \approx 0 \quad \text{or} \quad A\mathbf{d}_0 \approx \mathbf{d}_0$$

As noted, these expressions can be used to define a "pseudo" or "near-invariant" line, $\mathbf{d}_0$, implicit in the original structural ledge model.[6,35]

(b) If $\Delta d \neq 0$, an invariant line can be obtained in the common close-packed plane by rotating the two lattices about the plane normal.[3] In this situation, there is an O-line passing through the origin, but the O-line may not exist in the next pair of correlated planes. The reason, as Bollmann[27] has demonstrated, is that the O-line disappears on translation unless the vector $\mathbf{d}_b$ is contained in the plane of the invariant normal. Since $\mathbf{d}_b = \mathbf{b}_f - \mathbf{b}_b - \Delta d\,\mathbf{K}$, and $\mathbf{b}_f - \mathbf{b}_b$ by definition lies in the plane of the invariant normal, this condition can only be satisfied in the special case when $\mathbf{K}$ lies in this same plane. For $a_f/a_b \sim 1.25$, this condition is not met for any rotation which leads to an invariant line in the close-packed planes.

(c) If $\Delta d \neq 0$ and $a_f/a_b \sim 1.25$, a small rotation ($\psi < 1$ deg) is necessary to produce a true invariant line lying out of the close-packed planes,[38] as discussed previously. Under these conditions, the solutions for the O-lattice vectors associated with $\mathbf{K}\,\Delta d$ always lie nearly parallel to the "effective" invariant line.* The predicted values of $\mathbf{d}_0$ found from the structural ledge approach are within 1 deg of the true invariant line determined from the O-line model. Provided the correct choice is made for the in-plane Burgers vector, effectively the same habit plane containing a single set of dislocations can be obtained from either the O-line or structural ledge models. Hall *et al.*[22] also compared the structural ledge and O-lattice model predictions of habit planes for fcc-bcc systems and demonstrated the close similarities in the two approaches, while Furuhara and Aaronson[36] recently have analyzed a similar situation in the hcp-bcc case.

It is interesting to note that a $(3\bar{3}5)_f$ habit plane can also be predicted from surface dislocation models for *N-W* related precipitates. Knowles and Smith[39] have shown that this interface plane is one of minimum net Burgers vector, and this is also the plane of minimum mismatch predicted from O-lattice analysis.[29] The correlations between these different approaches are impressive, but they also highlight the problems in making a clear choice between the models. As we have seen, any one of them predicts the near $(3\bar{3}5)_f$ habit plane for the *N-W* orientation relationship, but none can be used to predict the $\sim(1\bar{2}1)_f$ habit plane for *K-S*-oriented precipitates.

Clearly, the underlying assumption of these different models—that the interfacial energy varies monotonically with the average misfit—is not always valid. One of the obvious drawbacks to the O-lattice model is that the distribution of misfit on an atomic scale is not considered. This is simply illustrated by estimating the atom mismatch along the direction defined by $\mathbf{b}_f$. The atom match is best at a position nearest to the O-line, but it is *not* perfect. If $|\mathbf{b}_f| > |\mathbf{b}_b|$, the residual misfit ($\mathbf{d}_r$) at this location can be expressed as

$$\mathbf{d}_r = m\mathbf{b}_f - (m + 1)\mathbf{b}_b \qquad [20]$$

where m is an integer nearest to the distance from the origin along $\mathbf{b}_f$ to the adjacent O-line. If, as before, l is the O-line spacing and θ is the angle between the O-line and the Burgers vector, m is given by

$$(m + s)|\mathbf{b}_f| = \frac{l}{\sin\theta} \qquad [21]$$

where $|s| \leq 0.5$. By following the approach of Purdy and Zhang[40] in their analysis of misfit normal to the O-plane, one can then show that the magnitude of the residual misfit at the O-line is

$$|\mathbf{d}_r| = \left|\frac{s}{m + s}\right| \cdot |\mathbf{b}_b| \qquad [22]$$

If $m \gg s$, $|\mathbf{d}_r| \to 0$. This condition is satisfied when the spacing of the O-lines $\gg |\mathbf{b}_f|$ or the dislocations are in near-screw orientation. More detailed calculations for

*An effective invariant line was defined by Zhang and Purdy[29] as one satisfying the condition $A\mathbf{x} = (1 + \varepsilon)\mathbf{x}$ where $\varepsilon < 10^{-7}$. By analyzing orientation relationships very close to those needed to satisfy the exact invariant line criterion, one can then invert the matrix T (Eq. [17]) and find an effective invariant line.

lattice parameter ratios, $a_f/a_b \sim 1.25$, show that $m + s \approx 9$ and 40, respectively, for near *N-W* and *K-S* O-line orientations. From this point of view, the *K-S* orientation relationship would be favored. A second factor which is not considered in the O-line approach is the low density of lattice points on the invariant line, which usually is a high-index, irrational direction (see Tables II and III). On an atomic scale, the invariant line must be stepped. This again should favor *K-S*-oriented variants, as the invariant line is close to the common close-packed direction for this case (Table II), but not for the *N-W* case (Table III).

In concluding this section, the similarities found between the O-line and structural ledge models encourage us to believe that further refinements of the structural ledge model[24] (guided by the principles dictated by O-line theory) would be useful. These refinements include: (1) a small rotation (<1 deg) between $(1\bar{1}1)_f$ and (101_b), as discussed in the O-line model; (2) a recognition that the in-plane displacement, $\mathbf{d}_b$, need not be parallel to either the corresponding lattice vector $\mathbf{b}_b$, or the pattern translation vector, $\mathbf{d}_{0_p}$; and (3) substitution of the O-lattice translation condition, $T\mathbf{d}_{0_p} = \mathbf{d}_b$, as a more general cancellation criterion (for the relief of misfit within the interface) than that proposed by van der Merwe *et al.*,[24] which restricted $\mathbf{d}_{0_p}$ to take values equal to $2m\mathbf{b}_b$.

VI. SUMMARY

The factors which control the choice of the invariant line and the development of particular facet planes in diffusion-controlled precipitation reactions in fcc-bcc systems have been discussed from several different perspectives. Elastic strain energy considerations, based on plate-shaped transformation products and rotations about the normal to the close-packed planes, cannot account for the formation of an invariant line, as there is no reduction in energy at the invariant line position. On the other hand O-lattice (O-line) and structural ledge models are able to account for many of the experimental observations. In particular, we find that the O-line model is in excellent agreement with observations in Cu-Cr and Ni-Cr, if the misfit at the habit plane is accommodated by a single set of dislocations parallel to the invariant line which has Burgers vectors which lie in the habit plane. The habit plane is then an unrotated plane of the transformation. Many features of the structural ledge model are also implicit in the O-lattice approach, provided the spacings of the close-packed planes are similar and the habit plane is reasonably close to the common close-packed $(1\bar{1}1)_f$, $(101)_b$ planes. With these restrictions, we have demonstrated that one can define a set of structural ledges compatible with an O-line solution having a single set of dislocations parallel to the invariant line. However a number of problems remain with O-line or structural ledge models. Energy arguments, based on identifying the habit plane of minimum mismatch or minimum net Burgers vector content, can account for the near-{335} habit planes in Cu-Cr (*N-W*-related), but not the near-$\{121\}_f$ habit planes in Ni-Cr (*K-S*-related). In the latter case, although the observed habit plane does not contain the minimum net Burgers content, the set of dislocations predicted by the O-line analysis are in near-screw orientation. We have shown that the atom mismatch at the facet plane is then reduced. This factor, coupled with the observation that the invariant line lies close to the common close-packed directions in the two phases, might account for the formation of *K-S*-related precipitates in fcc-bcc systems.

APPENDIX

The Elastic Strain Energy of Misfit Dislocation Arrays

Consider first the two finite arrays shown in Figures 5(a) and (b). The elastic strain energy of the second array may be calculated by integrating the work done by the force $-\sigma_{yx}\, dx\, dz$ acting over the length, l, of the cut used to define the Somigliana dislocation.[31] σ_{yx} can be evaluated from the stress fields given by the separate contributions of the uniform array of infinitesimal dislocations and [for the case depicted in Figure 5(b)] the two discrete dislocations, $-\mathbf{b}/2$, located at either end of the facet.

For the uniform array of infinitesimal dislocations (Figure 5(a)), the elastic strain energy, E_1, per unit length (in the z direction) is given by

$$E_1 = 2E_h \quad \text{where } E_h = \frac{1}{2}\int_0^R \sigma_{yx}\,\Delta u_x\, dx$$

$$\sigma_{yx} = -\frac{\mu b}{4\pi(1-\nu)l}\ln\left[\frac{x-\frac{l}{2}}{x+\frac{l}{2}}\right]^2 \qquad [A1]$$

$$\text{and} \quad \Delta u_x = \frac{bx}{l} \quad \text{for } -\frac{l}{2} < x < \frac{l}{2}$$

$$= \frac{b}{2} \quad \text{for } x > \frac{l}{2}$$

$$= -\frac{b}{2} \quad \text{for } x < -\frac{l}{2}$$

The upper cut-off limit for the integral (R) has its usual meaning in dislocation theory. Evaluating Eq. [A1] leads to the result that

$$E_1 = \left[r_0\left(1-\frac{r_0}{l}\right)\ln\left(\frac{l}{r_0}-1\right) - \frac{l}{2} + r_0 + R\ln\left(\frac{R-\frac{l}{2}}{R+\frac{l}{2}}\right) + l\ln(l+r_0) + r_0\ln\left(1+\frac{l}{r_0}\right) - \frac{l}{2}\ln\left(R^2-\left(\frac{l}{2}\right)^2\right)\right]\left[-\frac{\mu b^2}{4\pi(1-\nu)l}\right] \qquad [A2]$$

If $r_0 \ll l \ll R$, then

$$E_1 = \frac{\mu b^2}{4\pi(1-\nu)}\left[\ln\frac{R}{l} + \frac{3}{2}\right] \qquad [A3]$$

For the Somigliana dislocation, $\Delta u_x = 0$ for $|x| > l/2$, but the stress field in the region $|x| < l/2$ now includes the contributions of the two discrete dislocations located at $x = \pm l/2$.[31] The energy of the array shown in Figure 5(b) is

$$E_2 = \frac{\mu b^2}{4\pi(1-\nu)}\left[\left\{\frac{r_0}{l}\left(1 - \frac{r_0}{l}\right) - \frac{1}{2}\right\} \cdot \ln\frac{r_0}{l - r_0} + \frac{r_0}{l} - \frac{1}{2}\right] \qquad [A4]$$

In the limit when $l \gg r_0$

$$E_2 = \frac{\mu b^2}{8\pi(1-\nu)}\left[\ln\frac{l}{r_0} - 1\right] \qquad [A5]$$

The energy of the third array (Figure 5(c)) has been considered by a number of authors.[41–44] Although the approach to the calculation of the total energy in these studies is different, the final form of the energy of the epitaxial array is very similar. By following the method of Willis[43] and Jain *et al.*[44] one can show that in the limit of $l \gg r_0$

$$E_3 = \frac{\mu b^2}{4\pi(1-\nu)l}\left[\ln\frac{l}{2\pi r_0}\right] \qquad [A6]$$

ACKNOWLEDGMENTS

The authors are grateful to NSERC (Canada) for their support of this work, to Dr. C.P. Luo for his permission to use Figure 3 from his studies of precipitation in Ni-Cr, and to Dr. J.P. Hirth for pointing out an error in an earlier version of the Appendix.

REFERENCES

1. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans.* AIME, 1953, vol. 197, pp. 1503-15.
2. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 29-37.
3. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-73.
4. U. Dahmen, P. Ferguson, and K.H. Westmacott: *Acta Metall.*, 1984, vol. 32, pp. 803-10.
5. U. Dahmen and K.H. Westmacott: *Acta Metall.*, 1986, vol. 34, pp. 475-82.
6. M.G. Hall, H.I. Aaronson, and K. Kinsman: *Surf. Sci.*, 1972, vol. 31, pp. 257-74.
7. G.C. Weatherly, P. Humble, and D. Borland: *Acta Metall.*, 1979, vol. 27, pp. 1815-28.
8. M.G. Hall and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, pp. 1409-18.
9. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, pp. 1963-72.
10. G.R. Speich and R.A. Oriani: *Trans. AIME*, 1965, vol. 233, pp. 623-31.
11. T. Fuji, T. Mori, and M. Kato: *Acta Metall.*, 1992, vol. 40, pp. 3413-20.
12. K. Ameyama, T. Maki, and I. Tamura: *J. Jpn. Inst. Met.*, 1986, vol. 50, pp. 10-19.
13. K. Ameyama, G.C. Weatherly, and K.T. Aust: *Acta Metall.*, 1992, vol. 40, pp. 1835-46.
14. P.R. Howell, P.D. Southwick, and R.W.K. Honeycombe: *J. Microsc.*, 1979, vol. 116, pp. 151-58.
15. Z. Nishiyama: *Sci. Rep.*, Tohoku Univ., Japan, 1934, vol. 23, pp. 637-49.
16. G. Kurdjumov and G. Sachs: *Z. Phys.*, 1930, vol. 64, pp. 325-43.
17. C.P. Luo and G.C. Weatherly: *Phil. Mag.*, 1988, vol. 58, pp. 445-62.
18. U. Dahmen: *Scripta Metall.*, 1981, vol. 15, pp. 77-81.
19. M. Kato and T. Mishima: *Phil. Mag.*, 1987, vol. 56, pp. 725-33.
20. M. Kato: *Mater. Sci. Eng.*, 1991, vol. A146, pp. 205-16.
21. W. Bollmann: *Phys. Stat. Solidi* A, 1974, vol. 21, pp. 543-50.
22. M.G. Hall, J.M. Rigsbee, and H.I. Aaronson: *Acta Metall.*, 1986, vol. 24, pp. 1419-31.
23. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 365-76.
24. J.H. van der Merwe, G.J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, pp. 1165-75.
25. J.M. Howe and D.A. Smith: *Acta Metall.*, 1992, vol. 40, pp. 2343-50.
26. S.H. Wen, E. Kostlan, M. Hong, A.G. Khachaturyan, and J.W. Morris, Jr.: *Acta Metall.*, 1981, vol. 29, pp. 1247-54.
27. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer-Verlag, Berlin, 1970.
28. W.-Z. Zhang and G.R. Purdy: *Phil. Mag. A.*, 1993, vol. 68, pp. 279-90.
29. W.-Z. Zhang and G.R. Purdy: *Phil. Mag. A.*, 1993, vol. 68, pp. 291-303.
30. K.M. Knowles: *Phil. Mag. A*, 1982, vol. 46, pp. 951-69.
31. R. Bonnet, G. Marcon, and A. Aiti: *Phil. Mag. A*, 1985, vol. 51, pp. 429-48.
32. R. Bonnet, M. Loubradon, A. Catama, and P. Stadelmann: *Metall. Trans. A*, 1991, vol. 22A, pp. 1145-58.
33. R. Ecob and B. Ralph: *Acta Metall.*, 1981, vol. 29, pp. 1037-46.
34. K.M. Knowles, D.A. Smith, and W.A.T. Clark: *Scripta Metall.*, 1982, vol. 16, pp. 413-16.
35. T. Furuhara, J.M. Howe, and H.I. Aaronson: *Acta Metall.*, 1991, vol. 39, pp. 2873-86.
36. T. Furuhara and H.I. Aaronson: *Acta Metall.*, 1991, vol. 39, pp. 2887-99.
37. Y. Mou: Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, PA, 1992.
38. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, 1029-34.
39. K.M. Knowles and D.A. Smith: *Acta. Crystallogr.*, 1982, vol. A38, pp. 34-40.
40. G.R. Purdy and W.-Z. Zhang: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1875-83.
41. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. London A*, 1949, vol. 198, pp. 205-25.
42. L.B. Freund: *J. Mech. Phys. Solids*, 1990, vol. 38, pp. 657-65.
43. J.R. Willis: S.C. Jain and R. Bullough, *Phil. Mag. A*, 1990, vol. 62, pp. 115-29.
44. S.C. Jain, T.J. Gosling, J.R. Willis, D.H.J. Totterdell, and R. Bullough: *Phil. Mag. A*, 1992, vol. 65, pp. 1151-68.

On the Role of Interphase-Boundary Structure in Plate Growth by Diffusional Mechanisms

GARY R. PURDY and WENZHENG ZHANG

The structures of planar phase interfaces and of interfacial defects responsible for their diffusional migration are discussed in terms of extensions of the O-lattice concept, in which the intersection of the two structures is treated analytically. Two cases are considered and illustrated with well-characterized experimental examples: one in which two-dimensional structural matching leads to O-planes, and a second in which linear matching yields an array of O-lines. It is suggested that growth ledges moving normal to the O-lines will often require lateral kink formation and motion for their propagation. The misfit associated with transformation ledges is modeled in terms of real (screened) dislocations, which may coexist with virtual (unscreened) dislocations representing a residual component of misfit. A macroscopic shear can result from the cumulative action of transformation ledges with shear components parallel to the habit plane.

I. INTRODUCTION

THIS article focuses on the role of interfacial structure in the formation of platelike precipitates on initial growth from supersaturated solid solution. Faceted precipitates formed on initial growth can take shapes far from equilibrium, and we believe that these shapes are dictated primarily by anisotropies of interfacial mobility.

Thus, the question of solid-solid interfacial structure and its relationship to interfacial mobility (and anisotropy of interfacial mobility) takes preeminence. Many of the main points of this article will be illustrated through specific examples; the characterization of interfacial structure through high-resolution microscopy has developed into a high art in the past few decades, but the number of observations that speak directly to the question of migration modes and mobilities of the same interfaces is remarkably few. Many of these studies are due, directly or indirectly, to Aaronson and his coworkers.[1]

When one speaks of mobility as a property of an interphase boundary, it is implied that there exists some unique response function that relates the speed of boundary migration to an imposed normal force per unit area, or pressure. This function in general will be influenced by the structure and migration modes of the boundary, and these will often in turn be strong functions of the crystallographic parameters (structures and lattice parameters of parent and daughter phases, orientation relationship, and interface orientation) attending the phase transformation. In certain cases, the form of the response function will also depend on the magnitude of the force itself.[2] Finally, it seems that finite perturbations or singularities are often needed to permit the migration of simple interphase boundaries.

GARY R. PURDY, Professor of Materials Science and Engineering, and WENZHENG ZHANG, Postdoctoral Research Fellow, are with the Department of Materials Science and Engineering, McMaster University, Hamilton, ON L85 4L7, Canada.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

The concept of a "force" for interface migration also requires examination. It is conventional practice to regard the components of a normal force for interface migration as being additive and derived from several distinct sources. Thus, for example, one speaks of a capillary force, expressed as the product of the specific surface free energy σ and the interface curvature K. Other forces of a chemical, mechanical, or frictional nature have been invoked and a local force balance applied to the moving boundary.[3,4] This concept has been extended in several recent treatments to include both the finite thickness of the interface and the nature of diffusional processes within the interfacial region.[5] It has also been challenged,[6] on the basis that the coupling mechanism of the force to the migration process is not always evident. The formalism remains of conceptual value, but clearly does need further examination in terms of its generality of application.

This article addresses the nature of the discontinuity at an interface between structurally dissimilar crystalline phases. However, in the context of diffusional phase transformations, it may be important to note first that a chemical concentration difference will generally be present, and coincident or nearly coincident with the structural discontinuity. The "pure" chemical phase boundary, perfectly coherent, has been described in its equilibrium condition by Cahn and Hilliard,[7] and its kinetic response has been modeled by Langer and Sekerka.[8] These elegant treatments, arguably the most complete descriptions of a phase interface, yield insight into the behavior of the fully coherent interface between structurally similar or identical phases. They represent the opposite extreme of the cases dealt with here, and they remind us of the existence of a chemical component capable of contributing to the anisotropy of interface properties, and perhaps not generally separable from the structural component. For the present purpose, we will isolate the "pure" structural boundary, so that the problem becomes geometrical. Because it ignores the chemical differences, it is unlikely that any such approach, informative as it may be, will yield results of complete generality.

II. SYSTEMS EXHIBITING TWO-DIMENSIONAL MATCHING

We will first consider systems where the habit plane is one which admits the perfect, or nearly perfect, matching of the two structures. Two prototypical systems are chosen for discussion: θ' in Al-Cu and γ' in Al-Ag.

Precipitation from α Al-Cu solid solution has received perhaps more attention from experimentalists than any other system. The planar interfaces (habit planes) of tetragonal θ' precipitates are often found to be fully coherent. These interfaces hold rigidly to a $\{100\}_\alpha$ orientation; extrinsic defects in the form of ledges are necessary for the motion of the habit plane, which occurs by a lateral mechanism. Growth is therefore extremely anisotropic, so much so that newly grown plates have aspect ratios far from equilibrium and determined in part by the anisotropy of mobility of the interphase interfaces. The origin and nature of the growth ledges (or the similar edges of the platelike precipitates) have been studied in some detail.[9,10,11] In recent discussions,[12,13] the elastic properties of growth ledges on these precipitates and the nature and magnitude of the elastic interactions between ledges were examined as a function of elastic misfit components perpendicular to and parallel to the habit plane.

By first considering systems where the misfit is one-dimensional and perpendicular to the habit plane, we can demonstrate certain useful principles. Referring to Figure 1, which is a schematic intersection of two structures with one-dimensional misfit perpendicular to the habit plane, it is seen that periodicities may occur in the direction of the misfit. Adopting Bollman's terminology,[14] the planes of best matching are termed O-planes. The spacing of these planes will give an indication of the plate (or growth ledge) thickness (or height) corresponding to optimum matching between structures in the direction normal to the O-planes, and this distance will correspond to an absolute minimum in the elastic misfit associated with a growth ledge.

Thus, in the case illustrated in Figure 1, where there is a perfect matching of every five interplanar spacings of α with four spacings of β, we can expect the misfit to be accommodated completely by a prismatic dislocation loop. There will be no long-range strain field associated with the growth ledge of height h, however, and we will therefore term this misfit-accommodating dislocation "screened."

Note that it is not generally possible to achieve perfect matching in the direction of misfit (especially if growth-ledge heights are restricted to relatively small numbers of structural units) and that a residual elastic misfit (ε_r) will exist, as suggested by the schematic diagram shown in Figure 2. Precipitate thicknesses or ledge heights that minimize this elastic component of the energy will tend to be favored, especially for interfaces nearer equilibrium with respect to diffusion, *e.g.*, those formed during coarsening, as studied by Sargent and Weatherly.[9] In a study of the distribution of ledge heights on θ' precipitates during initial growth from solid solution, Stobbs and Purdy encountered a number of ledges that did not correspond to this minimum in the elastic energy, but which were identified with other, shallower minima.[11] Following the procedures used in a previous treatment of elastic misfit and its effect on interactions among growth ledges, we can simulate the residual elastic misfit normal to the habit plane in terms of virtual dislocations (with $b = \varepsilon_r d$, where d is the distance between virtual dislocations) and, in the current nomenclature, these will be termed "unscreened" to indicate that they carry a long-range strain field. For smaller ledge heights, it is appropriate to associate a single virtual dislocation with

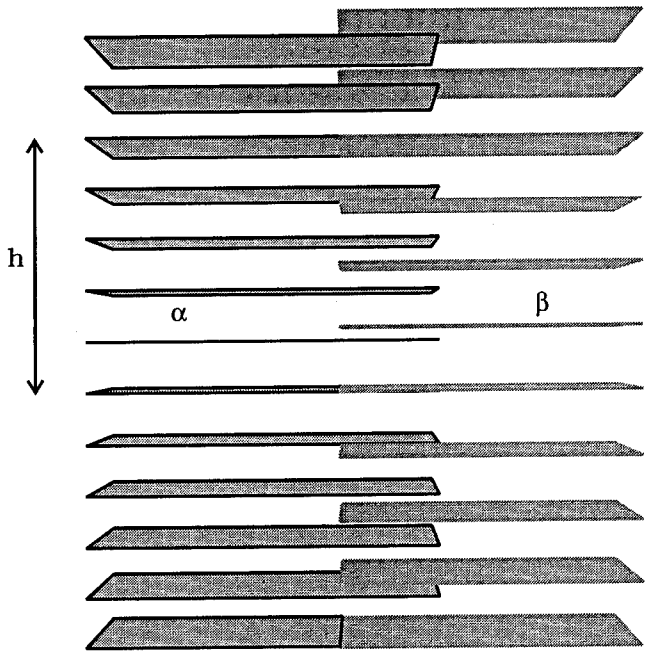


Fig. 1—Schematic intersection of two structures exhibiting planar structural matching. Five interplanar spacings in α correspond exactly to four such spacings in β, leading to O-planes spaced h apart. The misfit in the vertical c direction can be accommodated entirely by a "screened" edge dislocation, with Burgers vectors aligned in the c direction. A transformation ledge of height h would then require the climb of the misfit dislocation for its motion.

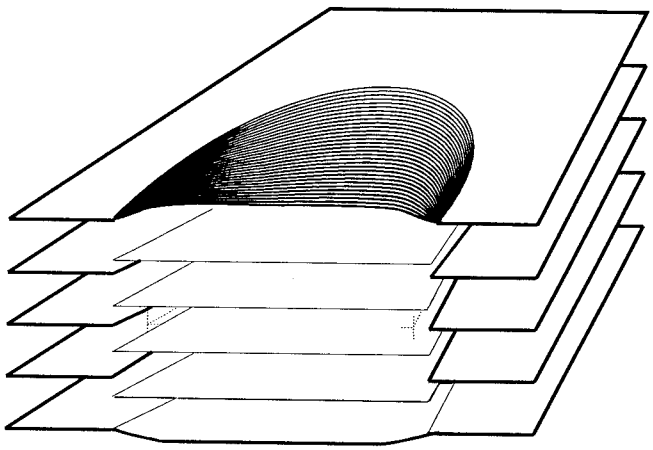


Fig. 2—In this case, the misfit is taken up by an edge dislocation, as in Fig. 1 (dashed line), which coexists with a residual elastic misfit (ε_r). This residual component of misfit can be simulated in approximation by a virtual "unscreened" edge dislocation loop with Burgers vector of magnitude $|\varepsilon_r h|$.

each growth ledge, in which $b = \varepsilon_r h$. Then a transformation ledge might have an associated misfit-compensating real (screened) dislocation as well as a residual misfit, represented by an unscreened virtual dislocation.

Treating the systematics of the situation more formally, we begin with the general definition of an invariant plane transformation[15]

$$S = I = \mathbf{e}\mathbf{v}' \qquad [1]$$

where $\mathbf{v}$ is the normal to the invariant plane; if $\mathbf{v}$ is a unit vector, then $\mathbf{e}$ defines the displacement associated with $\mathbf{v}$. Indeed, the displacement due to the transformation of any vector is parallel to $\mathbf{e}$ and I is the identity matrix.

Suppose a Burgers vector $\mathbf{b}_1$ is parallel to $\mathbf{e}$. An O-plane defined by any two vectors from the set $\mathbf{x}^O$ can then be determined from

$$T\mathbf{x}^O = \mathbf{b}_1 \qquad [2a]$$

with

$$T = I - S^{-1} \qquad [2b]$$

if lattice 1 is taken as the reference lattice. It can be shown that the O-plane is normal to $\mathbf{v}$ and has a spacing

$$h_O = \frac{(1 + \Delta)\,|\mathbf{b}_1|}{|\mathbf{e}|} \qquad [3]$$

where $\Delta = \mathbf{v} \cdot \mathbf{e}$. For the special case where $\mathbf{b}_1 \parallel \mathbf{v}$, and hence $\mathbf{b}_1 \parallel S\mathbf{b}_1$, the spacing of the O-planes is simply

$$h_O = \frac{|\mathbf{b}_1|\,|\mathbf{b}_2|}{|\Delta\mathbf{b}|} = \frac{|\mathbf{b}_2|}{|\mathbf{e}|} = \left(\frac{1}{|\mathbf{e}|} + \frac{\Delta\mathbf{b}}{|\Delta\mathbf{b}|}\right)|\mathbf{b}_1| \qquad [4a]$$

where

$$\Delta\mathbf{b} = |\mathbf{b}_2| - |\mathbf{b}_1| = \mathbf{b}_1 \cdot \mathbf{e} \qquad [4b]$$

Here, $\mathbf{e}$ represents the average strain normal to the habit plane. We next define $1/|\mathbf{e}| = m + s$, with $|s| < 0.5$ and m a positive integer. Then, with $\Delta\mathbf{b} < 0$, $h_O = (m + s)|\mathbf{b}_2| = (m - 1 + s)|\mathbf{b}_1|$.

The minimum height of the semicoherent step is likely to be $m|\mathbf{b}_2|$ *vs* $(m - 1)|\mathbf{b}_1|$. When $s \neq 0$, these quantities are not equal; there will be a residual strain $(s/m)(\Delta\mathbf{b}/|\mathbf{b}_2|)$ normal to the terrace. (This residual strain will generally be smaller than that of a coherent single-layer step.) The step may then be represented by a real screened dislocation, and an unscreened virtual dislocation with Burgers vector $s\Delta\mathbf{b}$ normal to the terrace.

A more stable semicoherent step may consist of several layers of near-matching planes such that the residual strain is further reduced. We further defined $s = (l + q)/n$, with n and l integers and $|q| \ll 1$. The height of the multilayer is

$$h_m = (mn + l + q)|\mathbf{b}_2| = (mn - n + l + q)|\mathbf{b}_1| \qquad [5]$$

The residual strain associated with a step of height $(mn + l)|\mathbf{b}_2|$ *vs* $(mn - n + l)|\mathbf{b}_1|$ is $\{q/(nm + l)\}\{|\Delta\mathbf{b}|/|\mathbf{b}_2|\}$ with $|q/(nm + 1)| \ll 1$. It seems that much of the known behavior of θ' in Al-Cu is covered by these considerations. Assuming that the atomic correspondence within the habit plane is $[002]_{\theta'}$ *vs* $[001]_\alpha$ and therefore $|\mathbf{b}_2| =$ 2.9 Å, $|\mathbf{b}_1| = 4.04$ Å, we have estimated the first interplanar spacing as 10.27 Å ($|\mathbf{e}| = 0.28$, $m = 4$, $s = -0.46$). Rather large step heights of approximately 2 nm are commonly observed, and these correspond to two planes ($n = 2$, $l = -1$, $q = 0.088$) with an elastic strain of 0.5 pct. This strain is concentrated at the growth ledge, leaving the terraces essentially strain free. Smaller steps have also been reported, with considerably higher strains (−4.5 pct) corresponding to more "metastable" elastic energy minima.[11] These observations are confirmed by the more recent studies of Dahmen and Westmacott,[16] who further consider possible detailed unit processes, conservative and nonconservative, of ledge migration in this system. They note that the basic principles underlying the proposed growth modes are the minimization of shape and volume strain. Thus, certain of their models involve coupled unit ledges with equal and opposite shear components, which cancel one another, and net displacements in the direction perpendicular to the habit planes.

As noted previously, the residual elastic strains can be modeled in approximation using "unscreened" virtual dislocations. They will often be responsible for long-range elastic interactions among growth ledges and for the main strain contrast associated with images of ledges in transmission electron microscopy (TEM).

We also note in passing that there is some evidence in this system for faceting of the precipitate edges and those of ledges, both of which tend to lie along $\langle 100\rangle_\alpha$ directions, particularly for growth under conditions of low supersaturation. This point invites further systematic experimental study, as it appears that this faceting behavior is lost at higher supersaturations.

A second well-studied case with planar matching is the precipitation of hexagonal close-packed (hcp) γ' plates in α Al-Ag face-centered cubic (fcc) solid solution.[17] Here, the structural matching is in some ways simpler than that for θ', discussed earlier. The stacking of closed-packed planes in the precipitate differs from that of the matrix, in the sense that two-thirds of the precipitate planes are sheared with respect to their matrix counterparts; unlike the previous example of θ' in Al-Cu, the main atomic displacements are parallel to the habit plane segments, rather than normal to them. The plates were found by Liu and Aaronson to give rise to reliefs on a prepolished surface[18]—clear indication that shears can result from the cumulative diffusional motion of ledges possessing an associated shear component. However, within the interior of a specimen, it seems that there is no macroscopic shear associated with the transformation, and that the individual $a/6\langle 112\rangle$ Shockley partial transformation dislocations are constrained to have their Burgers vectors distributed in such a way that they sum to zero.[17,19]

One next asks the question: What can the study of the geometrical intersection of the two structures tell us about the growth process? Noting that the interplanar (basal plane) spacing (here termed ϕ) in the γ' phase is rather close to the corresponding {111} spacing in the parent phase (within about 2 pct[20]), and that the two structures are brought into near-coincidence every 6ϕ, we might expect the unit transformation ledge height to be equal to 6ϕ. This is a very incomplete picture, as

demonstrated by a number of TEM studies.[21,22] The most recent and elegant of these is due to Howe, Dahmen, and Gronsky,[19] who confirm that this element of transformation is dissociated into Shockley partial dislocations spaced 2ϕ apart in the c direction and distributed such that their shear components sum to zero. Thus, we can consider that the individual Shockley growth ledges are unscreened, but that the shear components of the precipitate edges or of ledges of height 6ϕ will be self-screened. Howe *et al.*[19] also present evidence for a terrace-ledge-kink growth structure, indicating that the edges of the ledges are coherent, and suggest that the limiting process is probably related to the diffusion of silver atoms across kinks in partial dislocation ledges.

We conclude that certain features of transformation ledges can often be inferred directly from the study of the intersection of the two structures, but that the detailed transformation process cannot generally be predicted. Bollman's O-lattice approach has been of value in the rationalization of precipitate habit planes and, as we will note in Section III, as an aid in envisioning in a rather idealized and simplified way the essential elements of interfacial structures. When coupled with the general requirement that strain energy be minimized within the degrees of freedom available to the system, the geometrical theory can give limited but useful insights into the probable modes of growth of faceted precipitates.

III. SYSTEMS EXHIBITING LINEAR MATCHING

It is often observed that the habits of platelike precipitates formed by diffusional processes from structurally dissimilar parent crystals contain an invariant line of the transformation.[23,24] Indeed, many of the articles in this symposium are devoted to this topic.

We have recently reported on experimental studies of the structure of the Widmanstätten α/β habit plane interface in a Zr-2.5Nb alloy and on the crystallographic data relevant to this transformation.[25] In analyzing these results, we have hypothesized that the optimum orientation relation is that (invariant line orientation relation) which minimizes the interfacial misfit in the habit plane.[26,27] In the general invariant line orientation, it is quite possible that no periodic array of O-lines (lines of best matching, or extended O-lattice points) exists. The search for an optimum misfit is therefore confined to those orientation relationships that admit of a habit plane containing O-lines; the habit plane selected is the primary O-lattice plane with the smallest interplanar spacing and, by inference, the largest interdislocation spacing.* In this plane, the misfit is highly concentrated and can in principle be accommodated completely by a single set of dislocations which alternate with the O-lines, as suggested by Figure 3. Figure 4 shows a cross section of a lath-shaped α precipitate.

*The primary O-lattice plane is defined as that O-lattice plane associated with the closest-packed plane of the reference lattice.[26]

The comparison of this analysis with the experimental data is summarized in Table I. The (irrational) habit plane is indeed characterized by a single set of misfit dislocations, whose direction is coincident with the optimal invariant line and whose spacing in the habit plane is equal to the calculated distance between the O-lines (~10 nm). The micrographs in Figures 4 and 5 attest to the rather perfect dislocation structure of the undeviated habit plane interface. These dislocations are found to have Burgers vectors of the type $\langle 100 \rangle_\beta$, nearly perpendicular to the dislocation line direction. The full geometrical characterization of the habit plane interface[28] indicates that there exist "structural ledges" running approximately perpendicular to the invariant line.

We note that it may not always be possible to find appropriate direct interatomic vectors in the habit plane to serve as Burgers vectors for the misfit-accommodating dislocations and that significant out-of-plane components may be geometrically necessary, even

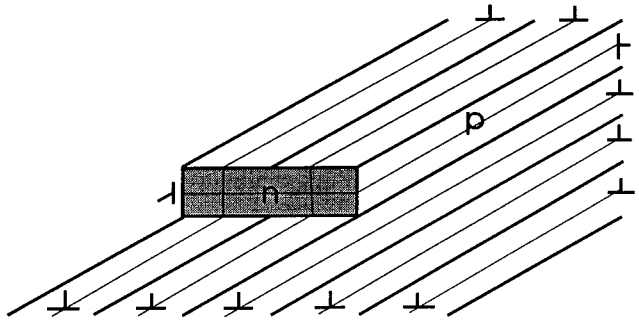


Fig. 3—An array of O-lines (shown here as full lines), generated through the geometrical intersection of the two structures. The O-lines alternate with misfit-accommodating (screened) dislocations in the habit plane. A transformation ledge is also shown, which carries its own edge dislocation (also screened), and whose lateral motion is required for the advance of the habit plane interface. Ledge faces lying parallel to the O-lines (p) are expected to experience stronger periodicities in their migration than those (shaded) lying normal to the O-lines (n).

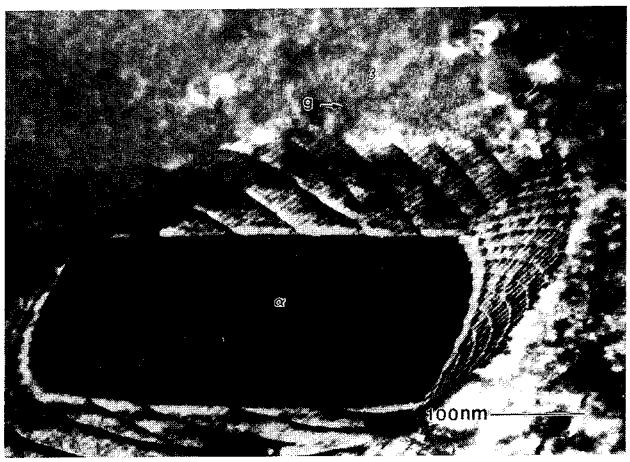


Fig. 4—TEM micrograph of an α lath, isothermally grown from a β at 650 °C in a Zr-11Nb alloy. The top surface is the habit plane, characterized by regularly spaced linear defects approximately 10 nm apart and lying parallel to the invariant line (see Table I). A series of widely spaced growth steps in the habit plane is also imaged, as is a more complex array of linear and curvilinear defects in the lath edge lying parallel to the invariant line. $g = (01\bar{1})_\beta$.

Table I. Comparison of Experimental and Calculated Results for the Widmanstätten Habit Plane Interface; α in β Zr-Nb

hcp(α)/bcc(β)	Orientation Relationship	Habit Plane	Invariant Line (Dislocation Direction)	Burgers Vector	Dislocation Spacing
Experimental	$(0\ 0\ 0\ 1)_\alpha \sim //(1\ 1\ 0)_\beta$ $[1\ 1\ -2\ 0]_\alpha \sim //[1\ -1\ 1]_\beta$	$(3.0\ -2.2\ -3.0)_\beta$	$[2.0\ -1.0\ 2.5]_\beta$	$[0\ 1\ 0]_\beta$	~9.5 to 10 nm
Theory	$\gamma = -0.290$ deg; $\phi = 1.015$ deg; $\theta = -2$ deg	$(3\ -2.16\ -3.05)_\beta$	$[2\ -1.02\ 2.69]_\beta$	$[0\ 1\ 0]_\beta$	9.9 nm

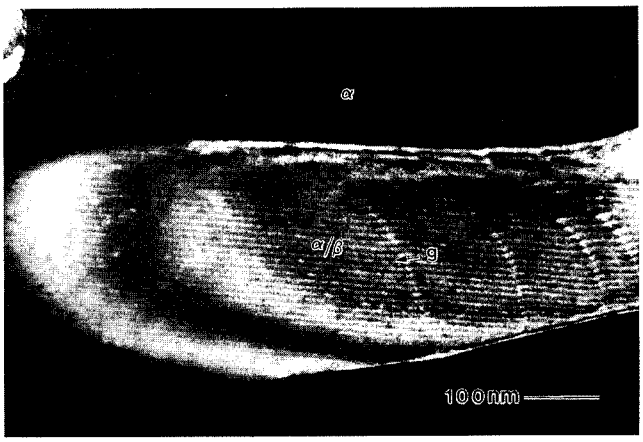


Fig. 5—An undeviated habit plane interface, viewed in plan, indicating the extremely regular nature of the linear interfacial defects. $g = (101)_\beta$.

though this will not contribute to the accommodation of misfit. Thus, just as we have insisted earlier that the invariant line correspond physically to an O-line, we now require that real interatomic vectors be available to accommodate misfit efficiently. In some cases, the choice of habit plane may be influenced by this second criterion.

Continuing in a general sense, we note that the habit plane interface defined by these geometrical considerations will normally possess very low symmetry. One needs only to examine the symmetry of intersection of, say, fcc and body-centered cubic (bcc) crystals in the Kurdjumov-Sachs orientation relationship to see that nearly all of the symmetry elements of the individual structures are lost in the intersection. Thus, the two examples of Section II, in which two-dimensional structural matching occurs on high-symmetry of the habit plane is that crystallographic degeneracy of Burgers vectors of the transformation dislocations (as found in the Al-Ag system) will also be the exception, rather than the rule. Further consequences arise when we consider the nature of possible transformation ledges in these invariant-line transformations.

We turn next to consideration of these ledges and their properties. Since we have selected the habit plane as that primary O-lattice plane containing maximal spacing between O-lines, the normal spacing between primary O-lattice (habit) planes would seem a reasonable choice for the transformation ledge height, as suggested by Figure 3. However, as in the previous discussion of plane-matching systems, we must distinguish between the cases where

(a) the spacing of adjacent habit planes and any microscopic translations parallel to the habit plane are such that the misfit can be accommodated completely by (screened) misfit dislocations and
(b) real (unscreened), misfit dislocations coexist conceptually with virtual (unscreened) dislocations.

First, let us consider case (a). Figure 6 is a schematic section of an O-lattice, in which the O-lines are viewed end-on and the traces of the primary O-lattice planes of least interplanar spacing are shown as bold lines. We see immediately that the ledge structure is not symmetrical; growth ledges moving to the right and to the left have

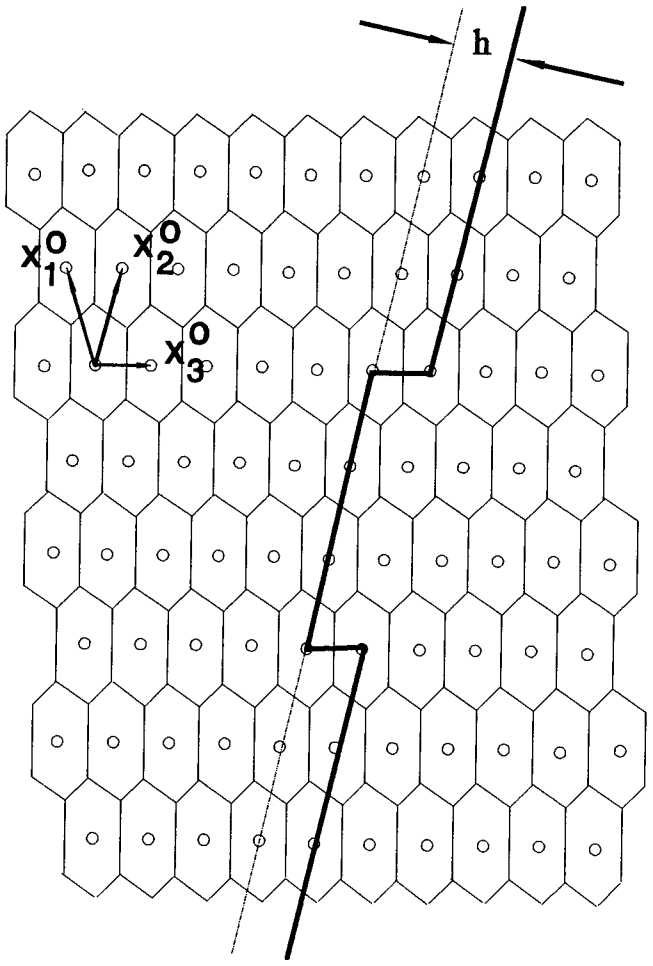


Fig. 6—A lattice of extended O-points (O-lines) viewed end-on. The habit plane is shown as the primary O-lattice plane with the smallest interplanar (and largest interdislocation) spacing. A transformation step of height h would contain a single dislocation and possess no long-range strain field.

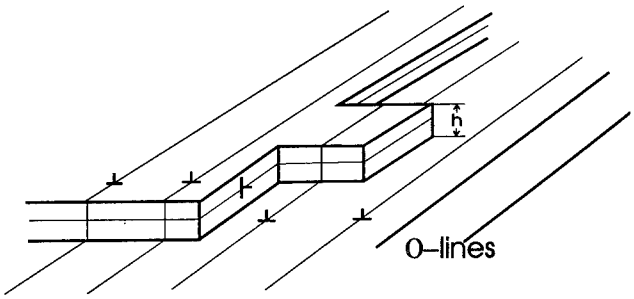
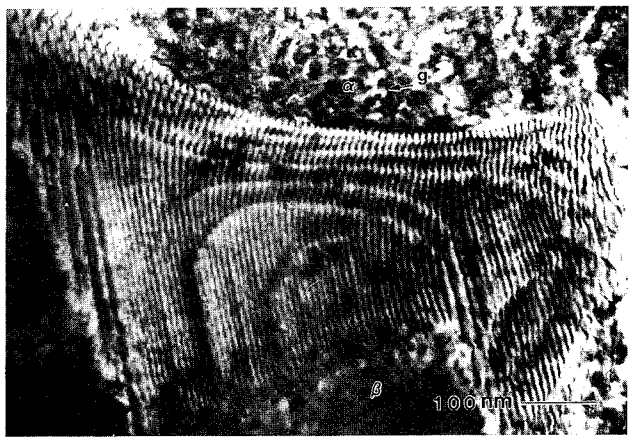


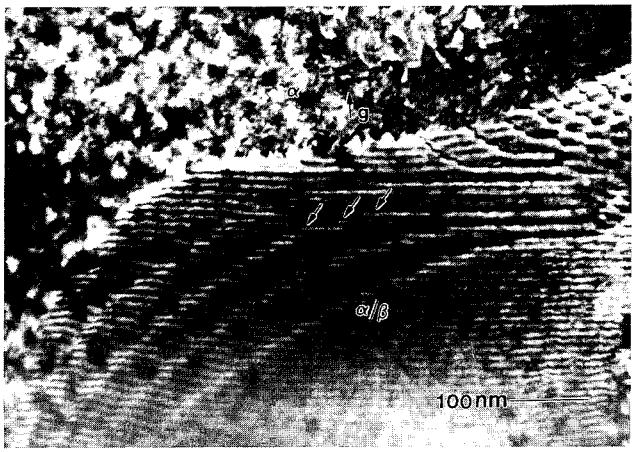
Fig. 7—Schematic illustration of the ledge structure of a habit plane interface, suggesting that for a ledge face lying parallel to the invariant line, double-kink nucleation is required to bring the ledge element into a new equilibrium position.

different rise orientations with respect to the habit plane, but the same height, h. It is also apparent that in the general case, the face of the transformation ledge joining O-lines is not perpendicular to the habit plane.

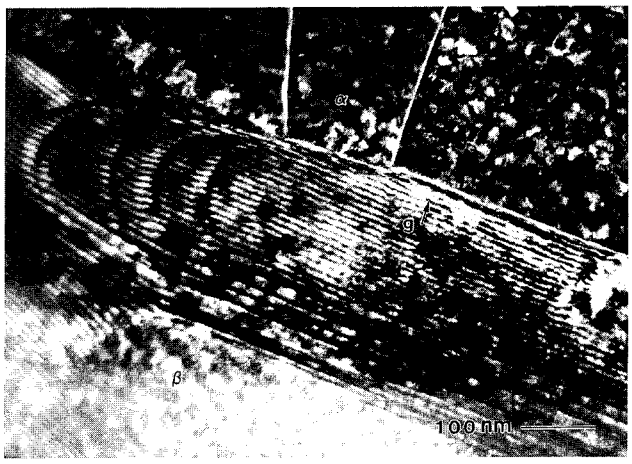
Again, from Figure 6, which shows two orientations of a transformation ledge on the habit plane, along with the misfit-accommodating dislocations, a further distinction becomes apparent. Those ledge risers which lie parallel to the O-lines (and to the primary misfit dislocations) will necessarily experience strongly periodic energy profiles as a function of the distance from the primary misfit dislocations. We expect that the equilibrium configurations for ledges in this orientation will be straight and parallel to the O-lines by torques whose magnitudes depend as usual on the variation of ledge energy with orientation. The main effect of these torques will be to see that the motion of these ledges tends to proceed by the motion of kinks, as suggested by Figure 7 and by the micrographs of Figure 8.



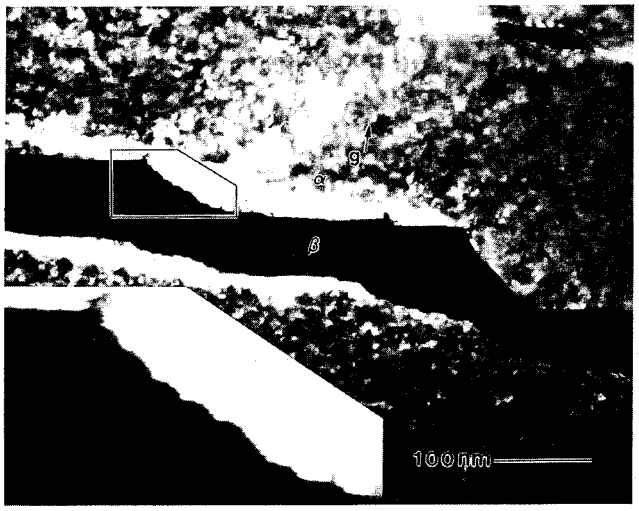
(a)



(c)



(b)



(d)

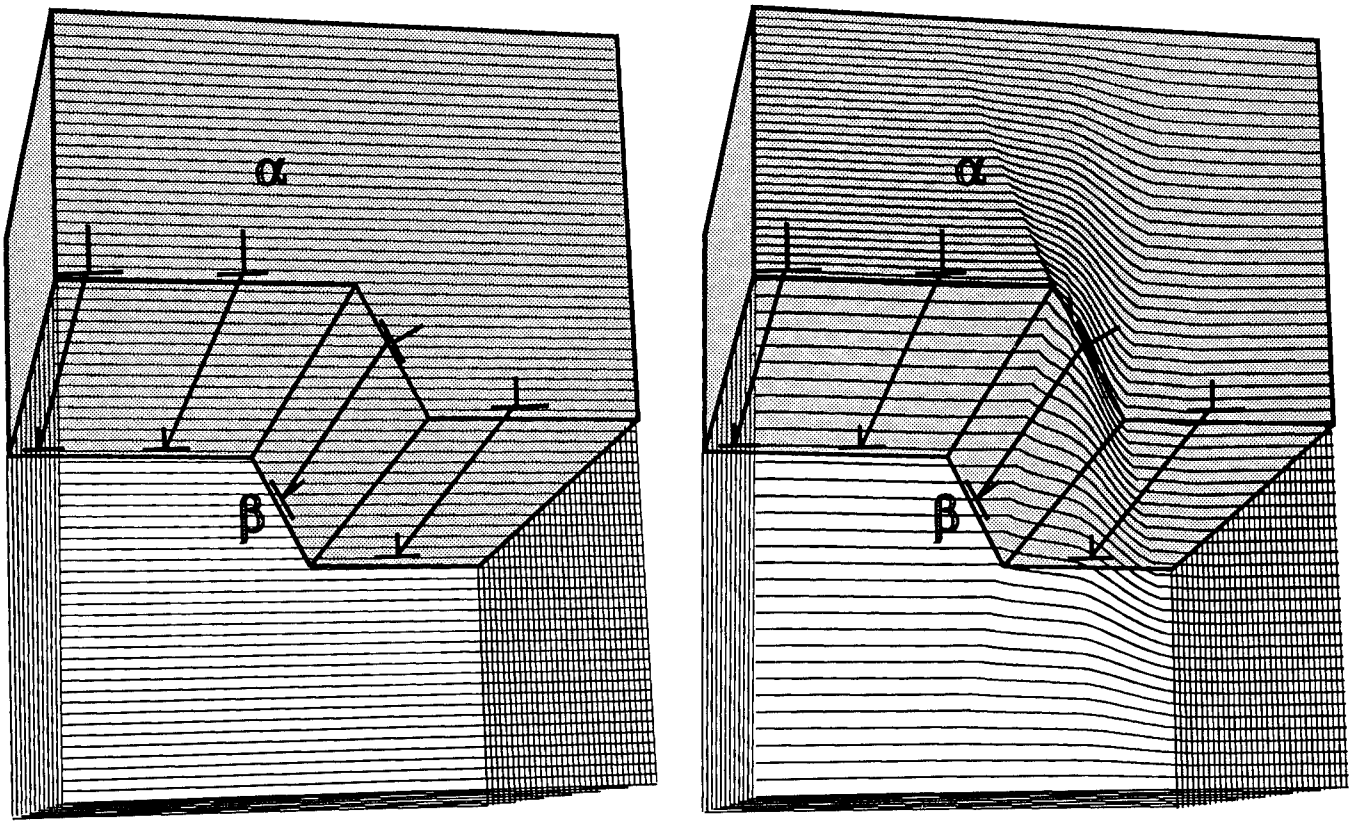
Fig. 8—TEM micrographs illustrating the morphologies of growth ledges on Widmanstätten α lath habit plane interfaces in Zr-2.5Nb. (*a*) A regular array of growth ledges; in general, these ledges are observed to terminate at the foil surface, or at the matrix dislocation (out of contrast in this image). $g = (110)_\beta$. (*b*) Growth ledges on a lath α precipitate. Note that the ledges appear to "lock" into the invariant-line orientation at the sides of the lath. $g = (110)_\beta$. (*c*) Images of kinks (arrows) on growth ledges. The kinks become more pronounced as the invariant-line orientation is approached, and the kink height appears to be equal to the spacing between interfacial dislocations. $g = (110)_\beta$. (*d*) An α/β interface viewed along the invariant line, demonstrating the repeatability of the transformation ledge height (approximately 2 nm). $g = (1\bar{2}12)_\alpha$.

In contrast, those ledge segments lying perpendicular to the O-lines (including the kink interface) will experience much weaker periodicities in their motion and will more likely be capable of uniform diffusional motion. Ledge interfaces rotated from the O-line orientation will contain kinks as a (geometrically) necessary part of their structure, but ledge interfaces lying parallel to the O-lines must somehow acquire kinks, *e.g.*, by a nucleation process. This strongly suggests that there must be different mobilities for the two orientations of ledge, and this difference should tend to favor the development of lath morphologies with the long axis of the lath parallel to the invariant line. (Of course, this tendency might also be favored by other factors, such as the elastic misfit of the entire precipitate.) Similar arguments have been advanced by Prabhu and Howe[29,30] and supported by reference to experiments on Al-Ag alloys.

A second inference that can be drawn from the model concerns the possibility of a macroscopic shear associated with the growth of the plate. It is known that many diffusional precipitates generate shear strains, or tilts, where they intersect a previously planar surface. If the transformation ledges contain a shear component, the cumulative effect of the emergence of like ledges at a free surface will be to generate shear-type reliefs, regardless of the actual mechanism of ledge migration. If the (plastic) deformation causing the relief is accompanied by substantial elastic deformation, creep processes will tend to moderate the surface relief in time.[31]

We turn next to case (b). Recalling that the preceding discussion has assumed that a lattice of extended O-points (O-lines) exists (in the sense of Figure 3) and that the semicoherent ledges have no unscreened elastic strain fields, we ask if this condition is likely to be general. By analogy with the case of θ' in Al-Cu, it seems that it is much more likely that some unscreened elastic strain is necessary to recreate the terrace structure on either side of the transformation ledge. The choice of transformation ledge character then becomes ambiguous. Figure 9(b) is a schematic indication of the need for further components of deformation required to accomplish the condition of structural similarity; the ledge becomes a center of unscreened elastic strain. We believe that this is the case for the example of α/β Zr-Nb for two reasons. First, our calculations admit of no opportunity for establishing structural equivalence on adjacent terraces in the absence of residual strain; and second, the ledges do possess long-range strain fields, as determined using appropriate images in conventional TEM.

To continue with discussion of the calculations, based on studies of the geometry of several hcp/bcc and fcc/bcc systems, we have determined that a



(*a*) (*b*)

Fig. 9—(*a*) Schematic illustration of the general transformation ledge for the case of no residual distortion. The habit plane is again characterized by misfit dislocations lying parallel to an invariant line, and the ledge is shown in its equilibrium position. (*b*) For the case where the ledge defect is insufficient to accommodate the total displacement required to replicate the terrace structure on either side of the transformation ledge, a residual long-range elastic strain field is developed.

two-dimensional array of O-lines is not generally available. We can, however, always obtain a single set of O-lines. The habit plane, which is the plane containing these O-lines, can be determined from $\Delta\mathbf{g}^*_{cp}$, with $\Delta\mathbf{g}^*_{cp} = \mathbf{g}^*_\alpha - \mathbf{g}^*_\beta$, where $\mathbf{g}^*_\alpha$ and $\mathbf{g}^*_\beta$ represent the correlated close-packed (or nearly close-packed) planes in the two crystals. Either of the $\mathbf{g}^*$'s must contain the Burgers vector that defines the displacement between adjacent O-lines; we can demonstrate that the minimum step height is given by the smallest value of $1/|\Delta\mathbf{g}^*_{cp}|$. In addition to the strain accommodated by the dislocations, there will generally exist unscreened components of strain, both parallel and perpendicular to the terrace, and necessary to recreate the terrace structure on either side of the ledge.

We therefore take the probable step height as $1/|\Delta\mathbf{g}^*_{cp}|$. We have previously noted that the O-lattice planes are described by the same expressions as "moiré planes," defined by $\Delta\mathbf{g}^* = \mathbf{g}^*_\alpha - \mathbf{g}^*_\beta$ when the two sets of planes are related by the transformation strain.[26,27] Thus, the experimental study of moiré patterns can be of help in identifying the probable transformation step height as well as the habit plane. The residual components of misfit will be responsible, as before, for a long-range elastic strain field associated with (and concentrated at) the growth ledge.

We believe that it is now of great interest to study the characteristics of growth ledges (in terms of structure as well as strain fields) in more complex systems, such as Zr-Nb, where the characterization of the undeviated habit plane interface is now well advanced.

IV. CONCLUSIONS

1. The structural contribution to interphase-boundary properties can often be understood on purely geometrical grounds. The habit plane of a platelike precipitate is postulated to be one of optimum interfacial misfit. Anisotropy of mobility (of habit planes *vs* precipitate edges and ledges, and of growth ledges within the habit plane) is considered an important determinant of growth morphology.
2. For systems exhibiting planar structural matching, it is found that transformation ledges are often centers of "unscreened" elastic strain, joining structurally equivalent habit plane terraces. A similar conclusion is reached for systems with linear matching.
3. Invariant line transformations are modeled in terms of an extended O-lattice, in which O-lines alternate with misfit-accommodating dislocations lying parallel to the invariant line. The habit plane is considered to be one of optimum misfit, with consequent maximal spacing between misfit-accommodating defects. Transformation ledges will experience torques that tend to rotate the ledges into the invariant-line orientation. In the general case, the transformation ledges will possess a short-range strain field, and a longer-range strain field due to a residual component of misfit. The transformation ledges can interact with habit plane misfit-accommodating defects, and this interaction may be important in considerations of double kink nucleation for ledges lying parallel to the invariant line. An operational selection rule has been proposed for the transformation ledge height.

ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada and by Ontario Hydro Research. We are indebted to Drs. V. Perovic, L.M. Brown, G.C. Weatherly, and Y.J.M. Brechet, for valuable discussions, and to Professor H.I. Aaronson and the organizing committee for the invitation to develop and present this paper.

REFERENCES

1. H.I. Aaronson and M.G. Hall: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1797-1819.
2. J.W. Cahn: *Acta Metall.*, 1960, vol. 8, pp. 554-62.
3. M. Hillert: in *The Mechanism of Phase Transformations in Crystalline Solids*, The Institute of Metals, London, 1969, pp. 321-47.
4. G.R. Purdy: in *Materials Science and Technology: Phase Transformations in Materials*, P. Haasen, ed., VCH, Weinheim, Germany, 1991, vol. 5, pp. 306-08.
5. Y.J.M. Brechet and G.R. Purdy: *Acta Metall.*, 1989, vol. 37, pp. 2253-59.
6. J.W. Cahn: private communication, Stockholm, Sweden, October, 1990.
7. J.W. Cahn and J.E. Hilliard: *J. Chem. Phys.*, 1958, vol. 28, pp. 257-67.
8. J.S. Langer and R.F. Sekerka: *Acta Metall.*, 1975, vol. 23, pp. 1225-37.
9. C.M. Sargent and G.C. Weatherly: *Phil. Mag.*, 1970, vol. 22, pp. 1049-55.
10. C. Laird and H.I. Aaronson: *Acta Metall.*, 1967, vol. 15, pp. 73-85.
11. W.M. Stobbs and G.R. Purdy: *Acta Metall.*, 1978, vol. 26, pp. 1069-72.
12. L.M. Brown, J.R. Dryden, V. Perovic, and G.R. Purdy: *Metall. Trans. A*, 1991, vol. 22A, pp. 1159-64.
13. J.R. Dryden, L.M. Brown, and G.R. Purdy: *Scripta Metall. Mater.*, 1991, vol. 25, pp. 2075-78.
14. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer, Berlin, 1970.
15. J.W. Christian: *The Theory of Transformation in Metals and Alloys*, Pergamon Press, Oxford, 1975.
16. U. Dahmen and K.H. Westmacott: *Phys. Stat. Sol. A*, 1983, vol. 80, pp. 248-63.
17. P.G. Shewmon: *Transformations in Metals*, McGraw-Hill, New York, NY, 1969, p. 279.
18. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, pp. 845-56.
19. J.M. Howe, U. Dahmen, and R. Gronsky: *Phil. Mag.*, 1987, vol. 56, pp. 31-61.
20. L.F. Mondolfo: *Aluminum Alloys, Structure and Properties*, Butterworth's, London, 1979, p. 213.
21. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 639-48.
22. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 649-58.
23. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-73.
24. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, pp. 1963-72.
25. W.-Z. Zhang and G.R. Purdy: *Acta Metall. Mater.*, 1993, vol. 41, pp. 543-55.

26. W.-Z. Zhang and G.R. Purdy: *Phil. Mag.*, 1993, vol. 67, in press.
27. W.-Z. Zhang and G.R. Purdy: *Phil. Mag.*, 1993, vol. 67, in press.
28. W.-Z. Zhang: Ph.D. Thesis, McMaster University, Hamilton, ON, Canada, 1991.
29. J.M. Howe and N. Prabhu: *Acta Metall. Mater.*, 1990, vol. 38, pp. 881-87.
30. N. Prabhu and J.M. Howe: *Acta Metall. Mater.*, 1990, vol. 38, pp. 889-96.
31. J.W. Christian: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1821-39.

Ledges and Dislocations in Phase Transformations

J.P. HIRTH

The coupling of dislocations and ledges in phase transformations is considered in terms of the symmetry of phases meeting at an interface. Growth and structural ledges and a variety of dislocation arrays are treated. Symmetry and the absence of long-range strain fields alone are shown to produce few restrictions on possible interface structures.

I. INTRODUCTION

STEMMING from the work of Gibbs,[1] Kossel,[2] Stranski,[3] and Burton *et al.*[4] on the role of ledges in crystal growth from the vapor, ledge models have been developed for a variety of diffusional and diffusionless phase transformations in crystalline solids. The concept of structural ledges, introduced to minimize the interfacial free energy, has also appeared.[5] In many cases, interfacial dislocations are associated with the ledges and an extensive literature has evolved in the description of the ledge/dislocation interactions as reviewed and discussed in References 6 through 12.

These concepts are related to the prediction of the habit plane and orientation relationship of two adjoining phases. The general guidelines to such theories are that the mechanism of transformation should be simple (ledges, dislocations, shuffles) and that the habit planes should have low interfacial free energy. In turn, low interfacial free energy has been associated with favorable lattice matching at the interface. The theories broadly fall into three categories. The first, based on the phenomenological theory[13,14] of martensite formation, is the invariant line theory.[15,16,17] The second is the theory of structural ledges,[5,18,19] which minimizes the atom mismatch in the sense of broken bonds at the interface. The third derives from the concepts of the O-lattice model[20,21] and of the coincidence site lattice and displacement shift complete models.[6,8,9] A recent symmetry theory developed by Pond[22] and based on the union of the crystal symmetry of the two crystals, *i.e.*, the dichromatic complex, provides a general description of defects, faults, and habit planes that, in some cases, reduces to the above lattice models.

A succinct comparison of these theories, except for the latter, is provided by Smith and Shiflet[23] and by Howe and Smith,[24] a treatment that is the starting point for the present work. Here, we focus on the equilibrium dislocation structure for interfaces, including those containing ledges. In some cases, this leads to the expectation of deviations from a simple habit plane and orientation relationship. We first review salient features of grain boundary theory and then turn to dislocation/ledge structures for several classes of interfaces, consider ledge bunching, and show that second-order coherency effects can exist. We follow many earlier workers and treat interfaces that are, on the average, planar and with unidirectional ledge/dislocation structures. The latter restriction is removed in some instances and, in any case, the two-dimensional (2-D) arrays follow the same physical principles. Consideration of three-dimensional, curved, or faceted second phases, with attendant more complicated considerations of dislocation formation and matrix dislocation configurations, is beyond the scope of the present work.

II. ARRAYS OF DISLOCATIONS

The treatment of the equilibrium structure of grain boundaries[25,26] requires some modification for interfaces. For illustration, we follow the specific example in Reference 27. We imagine two symmetrically ledged vicinal surfaces (Figure 1(a)) brought together to form a symmetric tilt boundary with angle of tilt θ (Figure 1(b)). As discussed subsequently, the Burgers vector content of the boundary is determined by means of a circuit in a perfect reference lattice. The actual dislocation spacing D is related to that in the reference lattice D_R, as shown in Figure 1(c). The skewed dislocations in Figure 1(b) can react to form grain boundary dislocations, as in Figure 1(d). As discussed by Bullough,[28] the relationships among the above quantities, which are forms of Frank's formula,[27] are

$$D_R = b/2 \tan(\theta/2)$$

$$D = b/2 \sin(\theta/2) = b_g/\sin\theta \qquad [1]$$

where b and b_g are the Burgers lengths of lattice and grain boundary dislocations, respectively.

For an interface, ledges may be more easily formed in one of the phases (Figure 2(a)). The interface dislocation array formed in this case (Figure 2(b)) is no longer symmetric. Indeed, an array of dislocations of type A no longer forms an equilibrium array with no long-range stresses, and dislocations of type B must be added to create an equilibrium array. A specific example of this sort[27] is depicted in Figure 3(a). Simple counting indicates that there are $n = 7$ dislocations of type A in the interval of boundary shown. Construction of an *FS/RH* Burgers circuit[29] and its equivalent in the perfect reference lattice (Figure 3(b)), gives the total Burgers vector $\mathbf{b}_T$ of the boundary. Consistent with counting, there are seven dislocations with Burgers length b_A and one with length b_B in the boundary. These suffice to establish

J.P. HIRTH, Professor, is with the Department of Mechanical and Materials Engineering, Washington State University, Pullman, WA 99164-2920.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

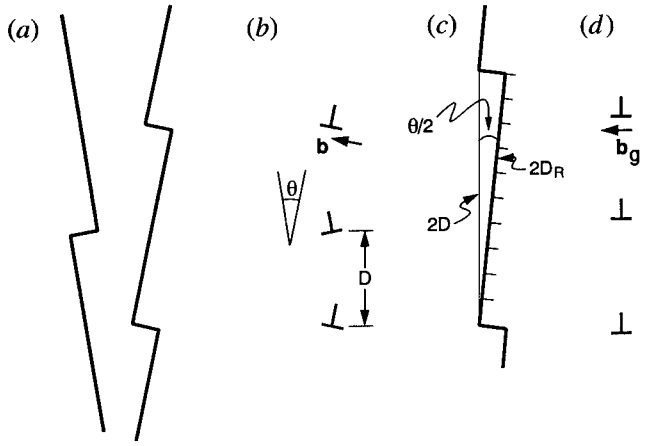


Fig. 1—Joining of ledged surfaces (*a*) to form (*b*) a symmetric tilt boundary, (*c*) detail of section of boundary, and (*d*) boundary represented by grain boundary dislocations.

an equilibrium array in the sense of no long-range stresses in the isotropic elastic case. The array would have one dislocation of type B every eighth dislocation in the example shown. Also, because of the difference in both ledge propensity and atomic bonding in the two phases, a symmetric tilt array may not be the equilibrium structure at an interface.

The spacing of dislocations is related to $\theta = \theta_1 + \theta_2$ and $\phi = \theta_1 - \theta_2$ by the relations[27]

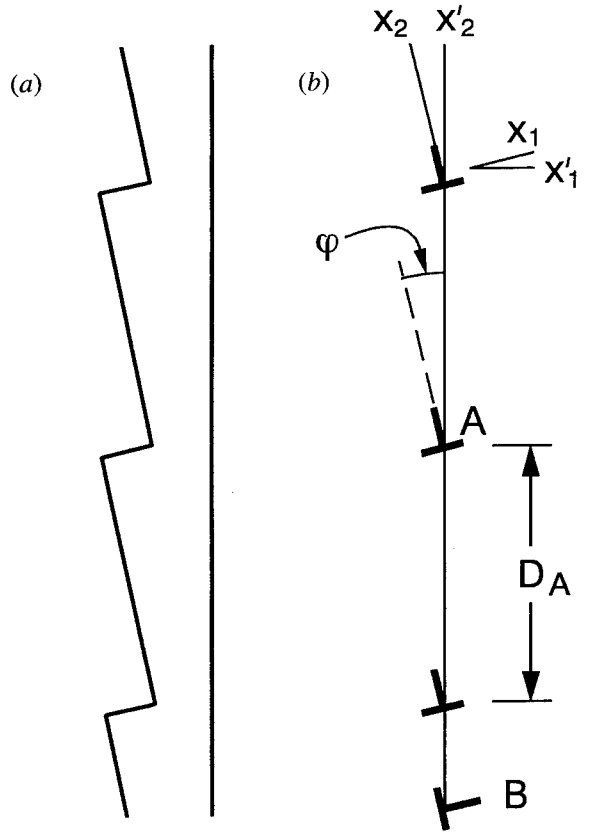


Fig. 2—Joining of ledged surfaces (*a*) to form (*b*) an asymmetric tilt boundary.

$$D_A = b/2 \sin(\theta/2) \cos\phi$$
$$D_B = b/2 \sin(\theta/2) \sin\phi \quad [2]$$

In the extension to interfaces, the Burgers vectors of type A dislocations, in general, will differ in the two phases, as illustrated in Figure 4, identical to Figure 3 except for the magnitude of b_α. In this case, the total Burgers vector in α and β scales as $\tan\theta_1$ and $\tan\theta_2$, respectively, so D_A further partitions to[27,30]

$$D_{A\alpha} = \left[\frac{\tan\theta_1 + \tan\theta_2}{\tan\theta_1}\right] b_\alpha/2 \sin(\theta/2) \cos\phi$$
$$D_{A\beta} = \left[\frac{\tan\theta_1 + \tan\theta_2}{\tan\theta_2}\right] b_\beta/2 \sin(\theta/2) \cos\phi \quad [3]$$

For a twist boundary, a square array of screw dislocations, as in Figure 5(d), satisfies equilibrium in the isotropic elastic case. If the Burgers vectors of one of the sets of dislocations were larger by some factor, the spacing of that set would be larger by the same factor in the equilibrium array. An equilibrium array could also be produced by a hexagonal array of three sets of screws with equal Burgers lengths.

For low symmetry crystals, the examples of Figures 3 through 5 may not be possible. For example, if the β crystal in Figure 3 were monoclinic, the only dislocations with a tilt component might also perforce have a screw component. An added set of screw dislocations would then be required to compensate for the elastic field of the screw components of the mixed β dislocations. In such a case, three sets of dislocations might be needed, but three would suffice for a general boundary if the strain components with plane normals in the boundary were unconstrained. In general, five sets of dislocations could be required.

Another case of an equilibrium array is that of misfit dislocations. In the well-known case[31] of crystals with cube habit planes misfitting in biaxial tension/compression, the equilibrium array is a square grid of like-sign edge dislocations with spacing $\lambda = b\langle a\rangle/\Delta a$. Here, $\langle a\rangle$ and Δa are the average and difference in the two respective lattice (or sub/superlattice) parameters. In the rare case when Δa is large, $\gtrsim 0.1\langle a\rangle$, a more complicated exact expression is needed for λ.[32] For tetragonal or orthorhombic habit planes, the array would be rectangular with λ, b, a, and Δa scaling with the respective in-plane $\langle 100\rangle$ direction.

For lower symmetry habit planes or crystals, in-plane shear misfit can also be present. The misfit relieving array of dislocations can be described by the reversible hypothetical path of formation illustrated in Figure 5. The same model is useful in the computation of the misfit. We consider two coincident tetragonal crystals α and β with matching (001) surfaces (Figure 5(a)). These crystals are then deformed to a cubic reference state with a lattice parameter d that is the average of the parameters a and c of the (001) surface (Figure 5(b)). Crystal β is then rotated with respect to α by an angle θ (Figure 5(c)). This rotation can be accomplished without added lattice strain by the generation of an orthogonal set of right-handed screw dislocations (Figure 5(d)). The spacing D

of the dislocations is given by Eq. [1]. The line directions of the dislocations are inclined to crystal α by $(\theta/2)$, and the spacing decreases with increasing θ to a minimum value at $\theta = \pi/4$. For greater rotations, symmetry indicates that θ in Eq. [1] should be replaced by $(\pi/2) - \theta$ and D increases with increasing θ until a dislocation free state is restored at $\theta = \pi/2$.

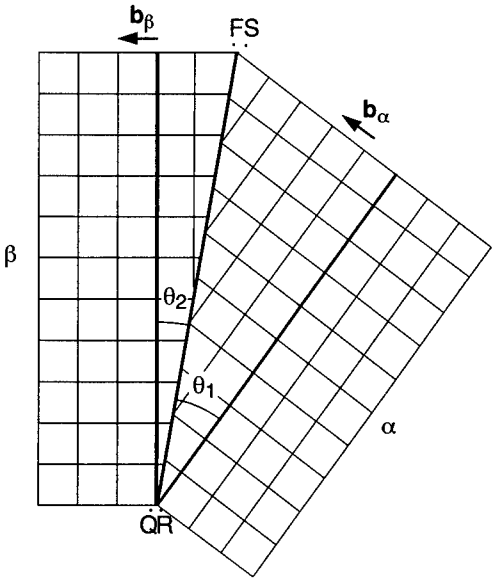
The crystals are then returned to their equilibrium tetragonal shape (Figure 5(e)). This does not change the spacing or character of the twist array in Figure 5(d). However, because of the prior rotation, the interface strain of the two crystals relative to the reference cubic crystal is incompatible, producing misfit. The misfit can be removed by a set of orthogonal misfit dislocations (Figure 5(f)). According to standard grain boundary theory,[30,33] the line directions are rotated relative to the cube axes of crystal α by an angle of $(\theta/2)$ and they are, in general, mixed dislocations. For small θ, however, they are close to screw orientation and, as indicated in Figure 5(f), are of mixed sign, right-, and left-handed screws. The spacing λ of the misfit dislocations, to be determined subsequently, decreases monotonically with increasing θ to a minimum value at $\theta = \pi/2$. Near $\theta = \pi/2$, as shown in Figure 5(g), the misfit is a state of balanced biaxial tension and compression. The misfit dislocations in this case are almost pure edge in character and are of opposite sign.

For an arbitrary degree of misfit at the interface, the in-plane misfit strain tensor can be decomposed into equal biaxial stresses and a pure shear stress. The appropriate misfit dislocation array is then the superposition of the traditional edge array[31] and that of Figure 5.

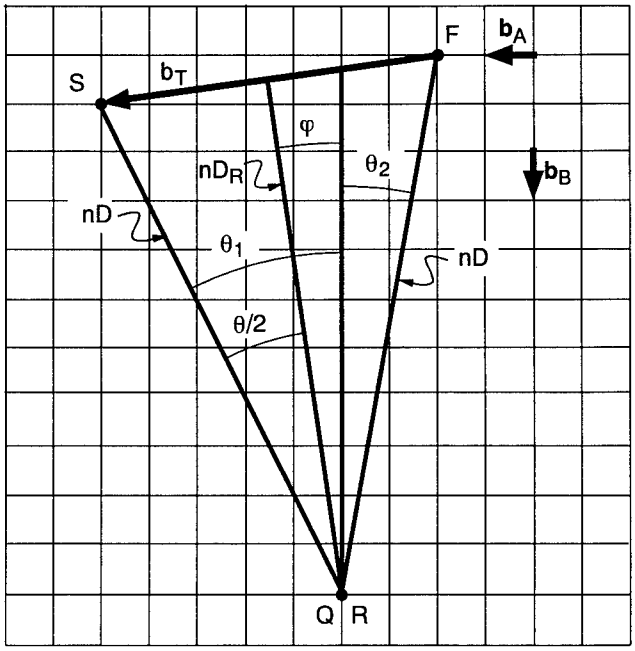
For either tilt or twist boundaries or for misfit arrays, a regular spacing is not stable unless D or λ, respectively, is an integral number of lattice spacings. If D differs, say by $0.011b$, from such an integral number, an extra misfit dislocation would be needed every 100 lattice spacings and every 1000 spacings for equilibrium to hold. A series of such dislocations of ever increasing spacing would be required if the above difference were irrational. Locally, however, intervals with a regular spacing would be present.

III. STEPPED INTERFACES

We follow the clear exposition of Howe and Smith[24] as a starting point. Figure 6(a) shows their view of a ledge in an otherwise perfectly matching interface between a simple cubic β crystal and a tetragonal α crystal. They note that such a ledge produces atom mismatching that can only be relieved by a dislocation and suggest that, therefore, there is no reason for such a ledge to be present in an equilibrium array, though it could be present as a growth ledge. We pursue this point and demonstrate that the ledge can be part of an equilibrium interface structure, albeit perhaps not the lowest energy one, if a rotation is involved in creating the equilibrium as they suggest. There is no constraint against the relaxation of the ledge to the configuration of Figure 6(b), which has ledge character with height $h = (a_\alpha + a_\beta)/2$, translational dislocation character[34] with Burgers vector $\mathbf{b}_t = \delta a$, and corresponds to a special case of a structural



(*a*)



(*b*)

Fig. 3—Representation of asymmetric tilt boundary in (*a*) real crystals α and β and (*b*) perfect reference crystal.

ledge.[5] As an explicit example, we consider $\Delta a = 0.1\ \langle a \rangle$. In terms of the symmetry theory of Pond,[22] see also Pond and Hirth,[34] the structurally identical regions on either side of the ledge in Figure 6(b) are related by the exchange symmetry operation $\mathbf{W}(\mathbf{R}, \mathbf{t})$. In the present case, the rotation $R = \eta$, where $\eta \rightarrow 0$ for a single ledge on a large interface region and the translation length is given by $t = \langle a \rangle - b_t$.

An array of such ledges is depicted in Figure 7. The dislocations are misorientation dislocations in the sense that they are in a tilt array. Locally, between the ledges, the (001) habit plane is retained. Yet the tilt array produces a global rotation of 0.716 deg so that the average

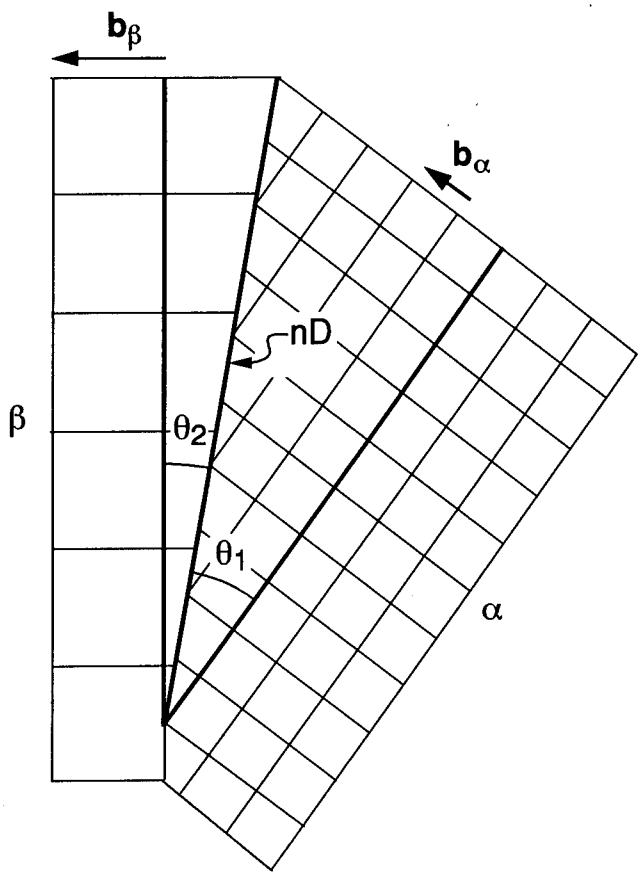


Fig. 4—Version of Fig. 3(a) with a different lattice parameter for β.

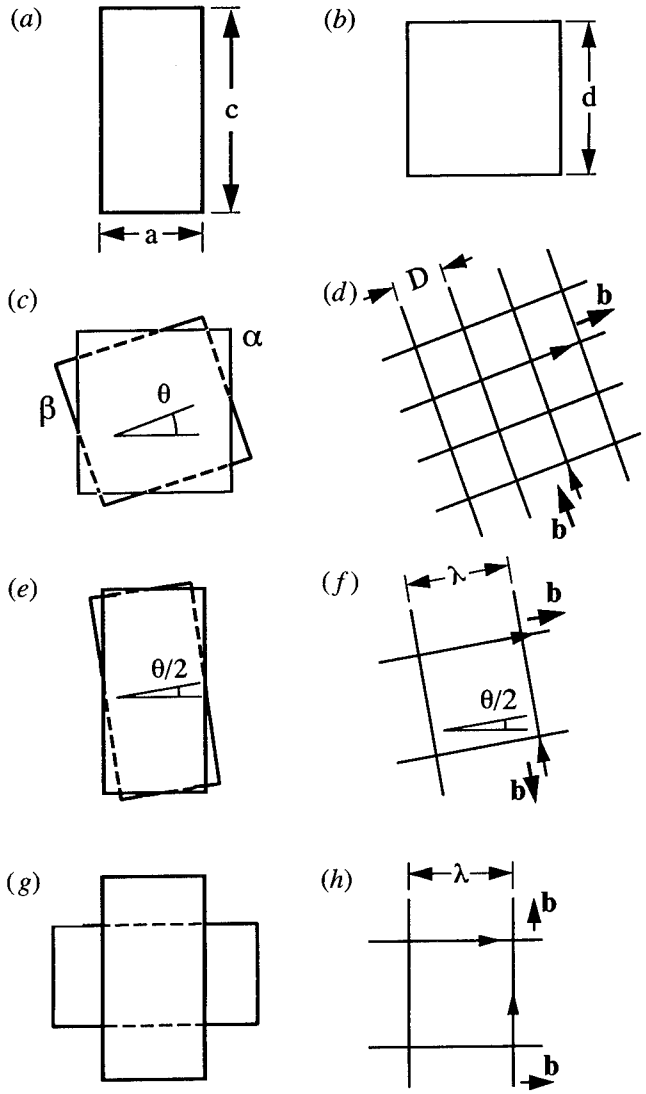


Fig. 5—Tetragonal crystals α and β: (*a*) with aligned (001) faces; (*b*) strained to a cubic shape: (*c*) rotated about [001] by an angle θ; (*d*) twist boundary array of right-handed screws at interface; (*e*) configuration (*d*) relaxed back to tetragonal shape; (*f*) misfit dislocations array to accommodate relaxation; (*g*) relaxed configuration for large $\theta \sim \pi/2$; and (*h*) corresponding misfit dislocation array.

boundary has become a (1 0 80) boundary. The ledge offset of the dislocations, *i.e.*, the rotation of the average interface relative to a pure tilt boundary, has the consequence that the array is of the type illustrated in Figure 2(b) with $\phi = 7.13$ deg. In order that the boundary be an equilibrium array with no long-range stress field, perfect dislocations of the type denoted as B in Figure 7 are necessary, with D given by Eq. [1] as $645a$. The magnitude of the global rotation scales directly with b_t and inversely with the ledge spacing D_A, while D_B scales inversely with b_t. Hence, arbitrary rotations about the low-index (001) habit plane are possible for interfaces in equilibrium in the sense of having no long-range strain fields. An unstepped boundary would probably have the lowest energy in the example given, but the opposite is not impossible. Indeed, the analysis of van der Merwe *et al.*[35] indicates that unstepped boundaries are favored for large misfits and small values of the misfit dislocation Burgers vector, while stepped boundaries are favored for small misfits and large values of the Burgers vector.

As discussed in Section IV, unit height ledges are those with translation vectors of the coset of the symmetry of the dichromatic complex,[22] the interpenetrating pair of lattices. In Figure 7, these are equivalent to the unit cell parameter of the two crystals; in general, the unit ledges in the complex can be several multiples of the interatomic spacing in either crystal.

The short-range elastic field can be relieved by faceting. For the above case, a 10 unit high ledge would accumulate a translational dislocation strength of $10b_t = b$. Thus, a superposed perfect misfit dislocation with Burgers length b would cancel the field of the translational dislocations, as in Figure 8. If the step were 20 units high, two perfect misfit dislocations would be needed and so forth. This would remove the tilt wall strain field, extending over a distance $\approx D_A$ in Figure 7, and replace it with a dipole field extending over a distance $\approx H$. Also, the rotation associated with the tilt wall would be removed and the (001) habit restored. If b_t were, for example, irrational, there would be a small residual δb_t after the process of Figure 8 occurred, leaving a small residual tilt which could, in turn, be removed at ledges with a larger period, and so forth. The change in strain energy in achieving the configuration of Figure 8 could be one of the reasons for the typical observation of multiple height growth ledges in certain diffusional phase transformations.[36–39]

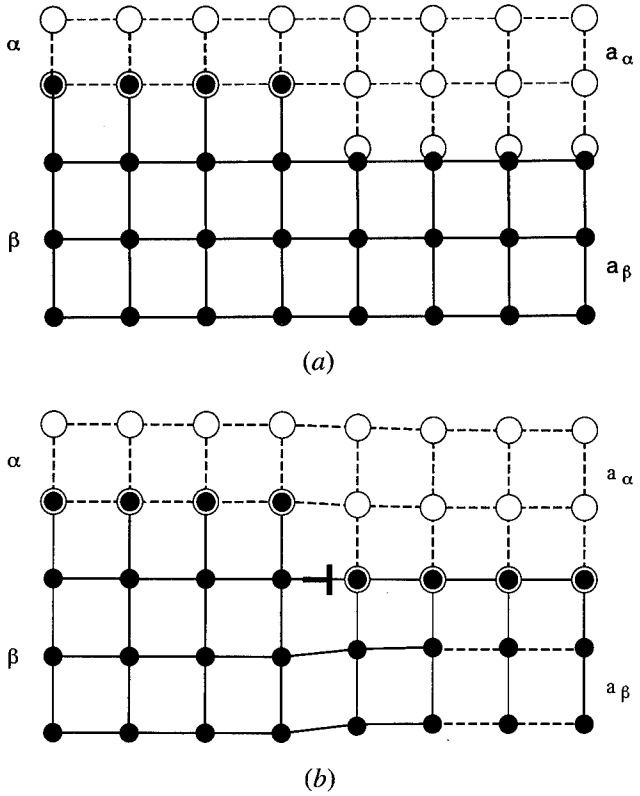


Fig. 6—(*a*) Unrelaxed ledge at interface;[24] (*b*) and relaxed ledge with translational dislocation.

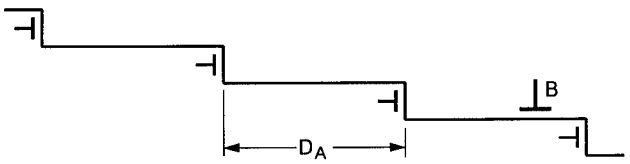


Fig. 7—Array of ledges of the type in Fig. 6(b).

In both Figures 7 and 8, a pattern advance is accomplished by the motion of ledges and their associated dislocations. In symmetry theory,[22] this pattern advance is described by means of the exchange operations $\mathbf{W}(\boldsymbol{\eta}, \mathbf{t})$ and $\mathbf{W}(0, \mathbf{t})$ for the cases of Figures 7 and 8, respectively.

The other simple case treated by Howe and Smith[24] is that of a translation in the habit plane with no rotation (Figure 9(a)). The symmetry description is $\mathbf{W}(0, \mathbf{t})$, with $\mathbf{t}$ in the habit plane. Here, the translational dislocation (Figure 9(b)) is also a transformation dislocation with $\mathbf{b} = (1/3)$ [100] in the depicted example. Again, the ledges and associated dislocations separate interface regions of the same atomic structure and, hence, would produce pattern advance if they were to move. A simple example of this case is the Σ 3, (111) twin in fcc metals, where the transformation (twinning) dislocation is a Shockley partial. An array of such dislocations would generate a long-range strain field. An equilibrium array can be restored in two ways. First, opposite-sign, perfect dislocations can be introduced with a spacing 3λ (Figure 9(c)). Second, ledges with a height equal to 3 unit ledges can be formed and a perfect dislocation superposed at the ledge. The result is essentially a shuffle

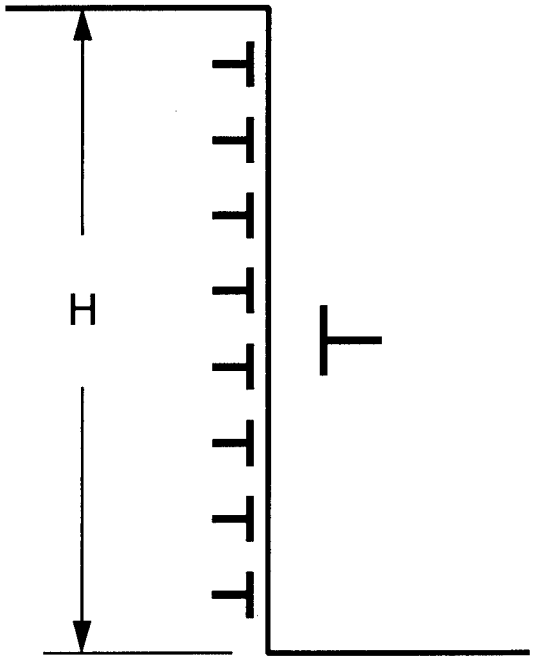


Fig. 8—Multiple height ledge variant of Fig. 7.

set of dislocations with no net Burgers vector (Figure 9(d)). Any lattice mismatch in the habit plane can be accommodated in an analogous manner by the superposition of misfit dislocations on the arrays of Figure 9.

The preceding discussion shows that there is no symmetry restraint to the glide-like motion of the unit ledges, with their translational dislocation character, in which case they could act both as structural and growth ledges. Whether the unit ledges can move, however, is an open question that has led to differing viewpoints in the literature, as reviewed in the present conference proceedings. While a detailed discussion of this issue is beyond the scope of the present article, we can outline possible mechanisms for dislocation motion or pinning. If they were to move, analogous to lattice dislocation motion, they would likely move by the formation of double-kinks over a Peierls barrier, as seen by Howe and Prabhu,[40] a process that would require thermal activation. Lateral kink motion could also involve a secondary Peierls barrier and could require thermal activation. With large driving forces (supersaturations), the barriers could be overcome athermally, in which case photon/electron damping, jog drag, or solute drag could become rate controlling. Solute pinning could also influence kink motion. With very low driving forces, the lateral motion of thermal kinks, present in equilibrium concentrations, could control dislocation motion. Alternatively, any of these constraints could cause the unit ledges/dislocations to be immobile. All of these processes are discussed, and model calculations for them are presented in Reference 27. With regard to the synchronous motion of ledge trains, motion by kink nucleation/lateral motion tends to lead to the same velocity of dislocation motion, so in the absence of drag effects, considered later in connection with multiple height ledges, a nominally equally spaced array of ledges would tend to persist, on the average, as a regular array. Moreover, both diffusional interactions[4]

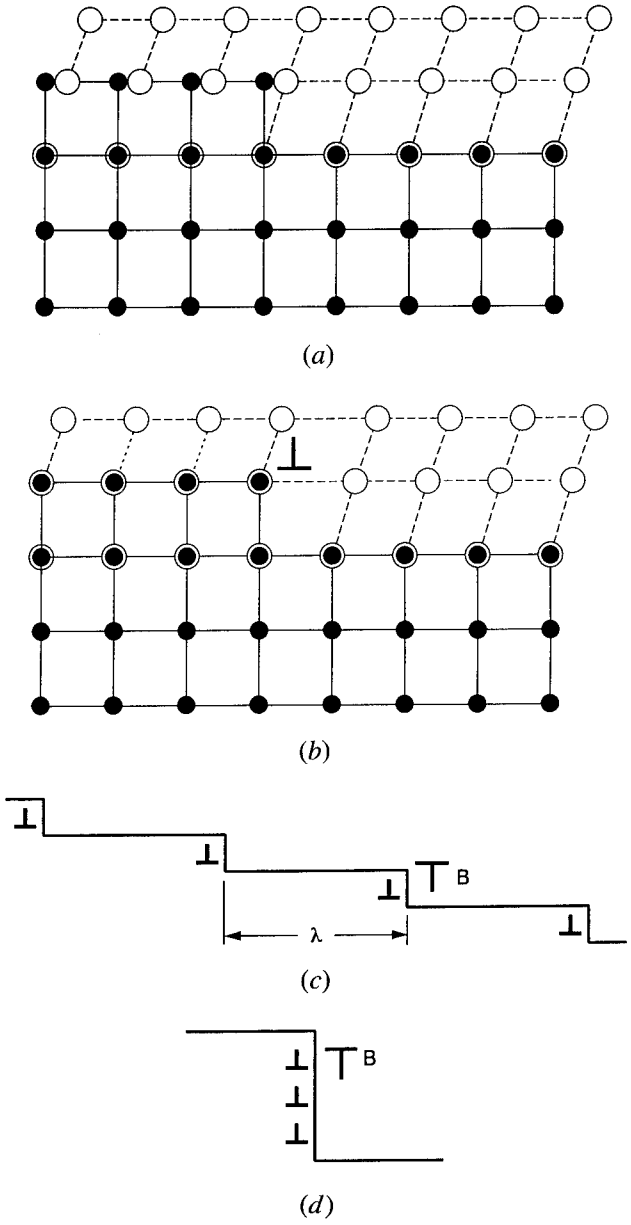


Fig. 9—(*a*) Ledge at interface,[24] (*b*) relaxed version of ledge, (*c*) array of such ledges with misfit relaxed, and (*d*) multiple height ledge with superposed misfit dislocation.

and elastic interactions of the translational dislocations would tend to stabilize the ledges at a regular spacing. Hence, there are reasons for ledges to fluctuate locally in spacing because of the stochastic nature of kink nucleation and lateral motion but to maintain a nominally regular average spacing.

Since the most general 2-D transformation involves the superposition of a simple shear and an in-plane and out-of-plane expansion or contraction, the two cases discussed previously can be added to produce equilibrium interface structures in the general case. The same concepts apply: that single or multiple ledge sequences can be equilibrium structures and can act as growth or structural ledges or both. All that is necessary is that the local interface regions on either side of the ledges be structurally identical and that the exchange operation $\mathbf{W}(\mathbf{R}, \mathbf{t})$ is a proper vector of the coset of the symmetries of the two crystals, *i.e.*, the dichromatic complex.[22] In terms of matrix transformation theory,[13,14] the general transformation matrix is

$$\mathbf{A} = \boldsymbol{\theta}\mathbf{RBS}\boldsymbol{\delta} \qquad [4]$$

where **B** is the Bain strain, $\boldsymbol{\theta}$ is the lattice invariant rotation associated with any translational dislocations in a misorientation tilt array, **R** is the remaining rigid body rotation accompanying the transformation, **S** is the lattice invariant shear, and $\boldsymbol{\delta}$ is a uniform lattice dilatation. The separate definition of $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$ is not necessary: $\boldsymbol{\theta}$ could be incorporated in **R** and $\boldsymbol{\delta}$ in **S**. However, they are separated here to emphasize the physical significance of these extra factors.

From the aspect of defects, $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$ represent deviations from a perfect transformation **RBS** that results in a lattice-matched, equilibrium interface. The quality $\boldsymbol{\delta}$ allows for an in-plane lattice mismatch that is accommodated by misfit dislocations for an equilibrated system and permits a broader range of habit planes than the simple theory. This role of $\boldsymbol{\delta}$ has been demonstrated and exemplified in the work of Srinivasan and Wayman.[41,42] The quantity $\boldsymbol{\theta}$ allows for lattice rotation in addition to that associated with **RBS**, is accommodated by misorientation dislocations for an equilibrated system, and permits a broader range of lattice orientation relationships than the simple theory. If the systems in **S** correspond to transformation dislocations,[43,44] for example, residual long-range stresses are left as a consequence of the geometry, as discussed here. The added dislocations needed to remove the remaining stresses and leave the final boundary free of long-range stresses contribute indirectly, as a geometric necessity, to $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$. In a sense, there can be a tradeoff between these parameters: a coherent, planar interface would have maximum $\boldsymbol{\delta}$ and zero $\boldsymbol{\theta}$, while as ledges were introduced, $\boldsymbol{\theta}$ would increase and $\boldsymbol{\delta}$ would decrease.

The various equilibrium arrays discussed here are consistent with Eq. [4]. Among these arrays, those that have the lowest energy would be expected in equilibrated microstructures and those that naturally follow from the transformation mechanism would be expected in unequilibrated microstructures. There is no fundamental symmetry principle that bars any of these structures.

Concepts, such as the invariant line model,[15,16,17] and the structural ledge model[5,18] have capability in predicting the lowest energy structures, but other structures seem possible. Indeed, if one were to relax the requirements of these models to predict habit plane/orientation relationships within a few degrees of the strict predictions of the models, essentially all of the structures considered here would be included. Such relaxations could be associated with lattice rotations, as discussed previously, with incomplete misfit accommodation, with mixed misfit dislocation/structural ledge accommodation, and with matrix deformation, to give some examples. In particular, if the difference between the **t** vector and a perfect vector of the dichromatic complex is small, the above models would be favored with the small extra effects accommodated, as discussed previously. Hence, if b_t were, *e.g.*, $0.01a$, single ledges in a configuration close to that predicted in the preceding models would be favored from

a combined minimum bond-breaking/low strain energy criterion. On the other hand, if b_t were $0.3a$, a triple height ledge with an accommodating misfit dislocation would be favored by a minimum strain energy criterion, since the strain energy of single ledges would be much larger than in the previous example.

IV. STRUCTURAL AND GROWTH LEDGES

A. *Small Ledges*

Howe and Smith[24] have commented on the similarity of the structural ledge and invariant line models in some cases. We present here a somewhat different viewpoint, in terms of dislocations, that leads to the same result. As discussed in the original developments,[5,18] structural ledges replace misfit dislocations as a way of retaining low-index terraces between the respective defects. From the dislocation viewpoint, the misfit dislocations are still present but have been lifted into the ledges. An illustration of this is the interface in Figure 10 of a lattice mismatched hcp-fcc pair with low-index terraces.

Figure 10(a) shows a semicoherent interface with perfect lattice dislocations, indexed on the fcc Thompson tetrahedron, acting as misfit dislocations to relieve coherency strain. The alternating **AB-AC** dislocations also have extraneous alternating sign screw components, a possible configuration chosen because it can directly relax to that of Figure 10(b). In this view, the misfit dislocations are pure edge, **Aδ** partial dislocations separating a repeating sequence of interfaces in which every third region is the interface equivalent of an extrinsic stacking fault with one adjacent intrinsic fault and every third partial has an extended core.[45] The faulted regions, discussed in more detail in the next section, arise

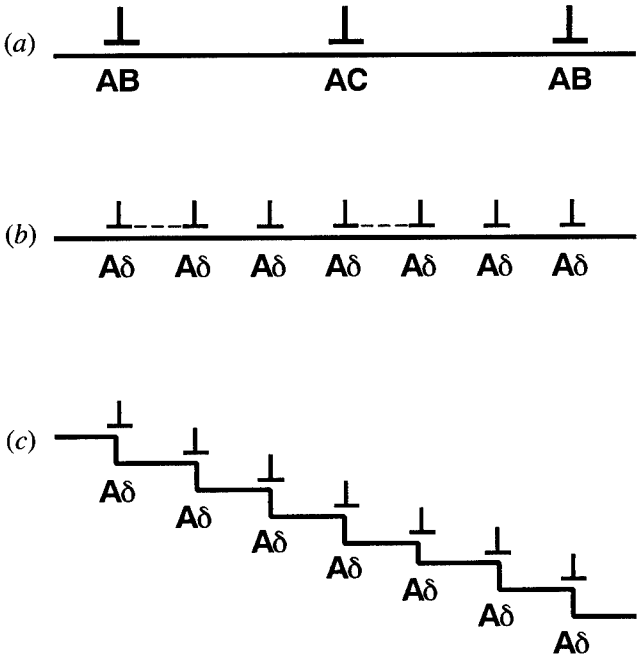


Fig. 10—(*a*) Perfect misfit dislocations at fcc-hcp interface, (*b*) partial dislocation version with dashed line indicating extrinsic-type fault, and (*c*) structural ledge version of crystals with same orientation relationship: the ledges are two planes high in this case.

because the **t** vector corresponding to $\mathbf{b} = \mathbf{A\delta}$ in $\mathbf{W}(\eta, \mathbf{t})$ is not a proper vector of the dichromatic complex. Continuous infinitesimal coherence dislocations of opposite sign[7,10] are present at the interfaces in Figures 10(a) and (b) but are not shown to simplify the presentation.

If (double height) structural ledges are present, as required for this fcc/hcp case[46,47] (Figure 10(c)), they incorporate the misfit dislocations. If there is no mismatch in lattice spacing normal to the terraces, Figure 10(c) also corresponds to the invariant line habit plane of Dahmen's[16] Figure 9(c), adjusted to have low-index terraces. In a generic sense, the same duality is illustrated in Figure 4 of Reference 24, but there, Figure 4(b) is the unadjusted invariant line view equivalent to Dahmen's Figure 9(b) and requires local rearrangement on the terraces to create the Figure 9(c) configuration of coherent terraces and dislocated ledges (the rearrangement is equivalent to the combination of distributed misfit dislocations into discrete ones). Construction of Burgers circuits[7] in the cases illustrated in References 16 and 24 verifies that the misfit dislocations are present along with the ledges. Indeed, since a circuit can be constructed remote from the boundary, a local rotation of the boundary cannot change the net misfit Burgers vector (although it can lead to the need for misorientation dislocations). Since misfit dislocations perforce have a spacing related to the lattice vectors of both lattices, the structural ledges are also spaced in such a manner. Depending on Δa, a regular array may satisfy equilibrium or occasional added dislocations may be needed with a regular array applicable only locally. Whatever the situation for the misfit array, so it is for the structural ledges.

In terms of Eq. [4], the structural ledges allow an in-plane dilatation related to **δ**, accommodated by misfit dislocations, that permits low-index terraces while at the same time producing a rotated habit plane as a consequence of the ledges. When the out-of-plane strain is zero, the invariant line model and the invariant plane strain model leading to Eq. [4] coincide, and if **δ** is small, the prediction of the invariant line and structural ledge models should then be the same, as demonstrated by Howe and Smith.[24]

In the case of Figure 10(c) the terraces are all identical in structure. In the symmetry operation relating them, $\mathbf{W}(\eta, \mathbf{t})$, the translation vector $\mathbf{t} = \mathbf{b} + \mathbf{h}$, where **h** is the step height vector,[34] is a proper vector of the dichromatic complex even though neither **b** nor **h** is. Hence, a dislocation interpretation of the structural ledges is that they are ledges of the proper height **h** to convert a faulted configuration, such as Figure 10(a), to that where terraces are perfect by converting **t** to a proper vector of the dichromatic complex. Moreover, the dislocation/ledge defects in Figure 10(c) are precisely the transformation dislocations[48] for the fcc-hcp phase transformation. These parallels show, as already stated by Smith,[49] that, geometrically, the phenomenological theories of martensite/invariant line models, the structural ledge models, and the transformation dislocation models are closely related for single ledge arrays, such as Figure 10(c). In addition, there is no symmetry constraint to the structural ledges acting as growth ledges as well in this case.

There are starting configurations other than Figure 10(a)

that can lead to the same result. If all the initial dislocations were of one type, say **AB**, they could relax to Figure 10(b) or (c) provided that emissary **BC** screw dislocations[50] were emitted from the interface to shed the screw components of the initial dislocations. Also, transformation could, in principle, be produced by alternating **Aδ**, **δB**, and **δC** transformation dislocations.[43] These would have no net misfit relieving component, however, and in the case considered here, would again have to emit emissary dislocations to produce the misfit relieving array of Figure 10(b) or (c). Alternatively, **CB** screw dislocations could be absorbed from the matrix with the same consequences. The emission or absorption of matrix dislocations would, of course, change the strain field in the matrix.

B. *Multiple Height Steps*

Of the unit (or double for Figure 10) height ledges illustrated here, there is no symmetry constraint to them acting as growth ledges. Those of Figure 7 could move by local diffusion jumps (most likely at kink sites) and would produce no large shear-type offset but would produce a weak lattice rotation. Those of Figure 9(a) or (b) or Figure 10(c) could move under diffusion control at kinks or by standard kink motion, would produce a shear-type offset, and would produce a large lattice rotation. Why then are multiple height ledges often observed in growth? There are at least two possibilities for the bunching of unit steps into multiple height steps, although these would be opposed by elastic interaction effects for cases such as Figures 9(b) and 10(c). We consider a general case of this type (Figure 11) with the dislocations having both a translational/misfit character and a misorientation character.

First, if the ledges are moving under diffusional control, any local drag force that is selectively imposed on a unit ledge leads to a pileup of ledges[51–55] that is stable against redissociation into unit ledges.[56] A local drag force could be produced by intersection with an extrinsic dislocation or an orthogonal misfit dislocation, by impurity adsorption at the ledge, or by a local concentration fluctuation. Once a sharp pileup formed, for the case of the transformation dislocations in Figure 10(c), the local dislocation interactions would tend to snap the array into a tilt configuration (Figure 11(b)). This would cause a long-range, concentrated stress field to emerge[57] and would strongly favor the addition of accommodation dislocations of type B in Figure 11(b). Even with accommodation dislocations present, there would be no constraint additional to drag forces hindering the motion of this multiple-height ledge. Dislocation B could combine with a tilt dislocation to form an alternate shuffle arrangement in Figure 11(b) as well.

The other situation would arise if the transformation dislocations had a weak misorientation dislocation component, as in Figure 11(a). Dislocations of this type have been discussed by Olson.[58] Then, perhaps initiated by the previous mechanism, the configuration of Figure 11(c) could form with a dislocation C reducing strain energy by accommodating the accumulated strain associated with the misorientation component of the unit defects. The

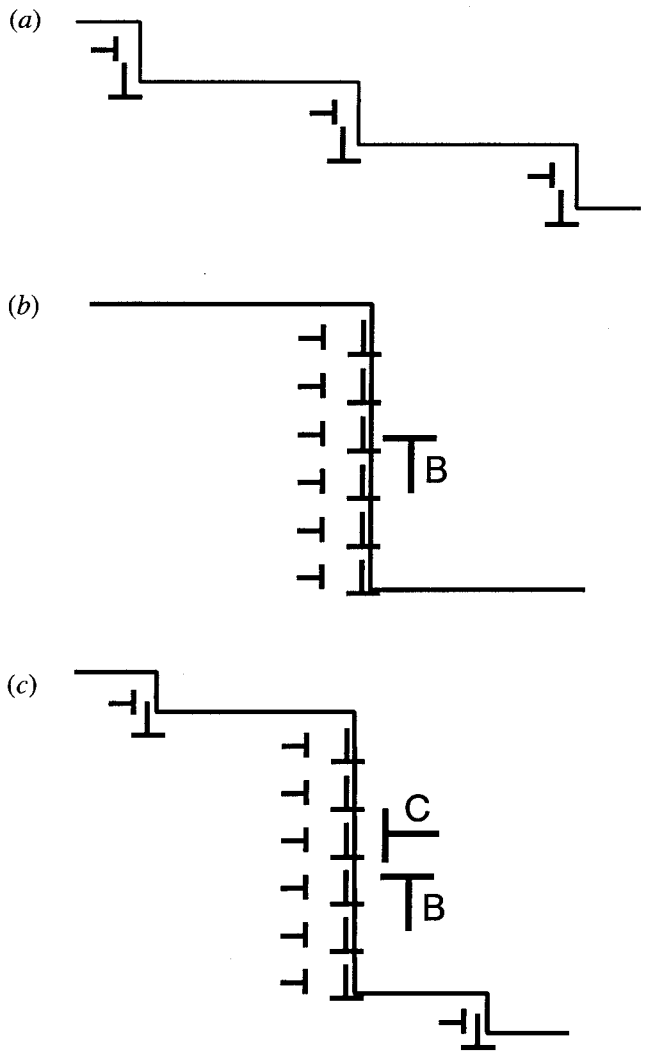


Fig. 11—(*a*) Unit ledges with associated dislocations, (*b*) multiple-height ledge, and (*c*) multiple-height ledge.

presence of dislocation C would provide an added constraint to the motion of the ledge, since C could only move by climb, with vacancy production or annihilation, as the multiple-height ledge moved. If the absence of long-range stress were an important energetic requirement for Figure 11(c), the number of ledges that must condense to balance the Burgers vector of dislocation B or C would provide a prediction of the height of the multiple ledge. The array in Figure 11(c) could also relax, as mentioned previously.

For the cases of Figure 11(a), (b) and/or (c), it appears possible for both unit and multiple height ledges to act as growth ledges. If both operated, the unit ledges would move faster and would move in and out of the multiple ledge, as suggested by the view in Figure 11(c). Since the unit ledges have dislocations with both transformational and misfit character, there is no problem with the misfit dislocations ascending the multiple ledge. If the ledge were to encounter a pure misfit dislocation, *i.e.*, one with no ledge character, it would require added accommodation.[36,59] As unit ledges joined the ledge of Figure 11(b), the dislocation B (or as many of such dislocations as were present) would be displaced from a

symmetric position on the ledge. Then climb would be required to restore the symmetric configuration and an added drag resistance would be connected to this process. Unit ledges could similarly join the ledge in Figure 11(c). Again, climb would be required for dislocation B, but dislocation C could glide to maintain a symmetric position.

The models in Figures 10 and 11 are 2-D in nature, as in Reference 21, implying no misfit and perfect accommodation (or their neglect) in the direction normal to the page. If there were misfit in the latter direction, misfit dislocations (or structural ledges) with Burgers vectors normal to the page would also be required. These too would require added accommodation at unit or multiple height ledges.[11]

As for single ledges, there is the issue of the mobility of multiple height ledges. If the risers are faceted, they too could move by unit or small multiple ledges sweeping the surface by kink motion.[47] Alternatively, the orientation of the riser may differ from the terraces, so that for configurations such as those in Figures 9(b) and 11, the riser could be considered as a set of overlapping dislocation cores. A nonlinear array of this type could facilitate kink nucleation compared to the unit ledge case or could even favor surface roughening of the riser. Either case would provide an expectation for easier motion of multiple height ledges than unit ledges.

V. FAULTED INTERFACES

As has been discussed earlier,[10,60] ledges and translational dislocations can separate interfacial regions of differing structure, analogous to the misfit case of Figure 10(b). The ledges in such cases are partial ledges and bound domain boundaries, either in the interface or in the bulk.[34] In such a case, **t** in **W**(**R**, **t**) is not a perfect lattice vector of the coset of the dichromatic complex.[34] In circumstances where the energies of the two different interfacial structures do not differ markedly, as for intrinsic/extrinsic faults and certain related interfacial faults in fcc crystals,[10] other equilibrium arrays, also without long-range strain fields, are possible. As an example, we consider the simple cubic β to complex cubic α transformation in Figure 12, with an (001) habit plane and a Σ 5 coincident site lattice. A perfect unit ledge is four planes high in crystal α and five planes high in crystal β (Figure 12(a)). In this instance, the ledge height $H = t$ and **t** is a vector of the dichromatic complex. Figure 12(b) shows the ledge split into two partial ledges. In the latter case, **t** is not a perfect lattice vector of the complex and the atomic structure differs on the two sides of the ledge. When such differing configurations are possible, further ledge structures involving partial ledges and any accompanying dislocations would be possible. Their descriptions would follow those of the preceding examples, except that the differing interfacial energies could lead to local relaxation and to slightly differing ledge spacings in the array.

VI. ANISOTROPIC ELASTIC EFFECTS

Anisotropic elasticity, through an influence on the fields of dislocations, can introduce some subtle effects in the

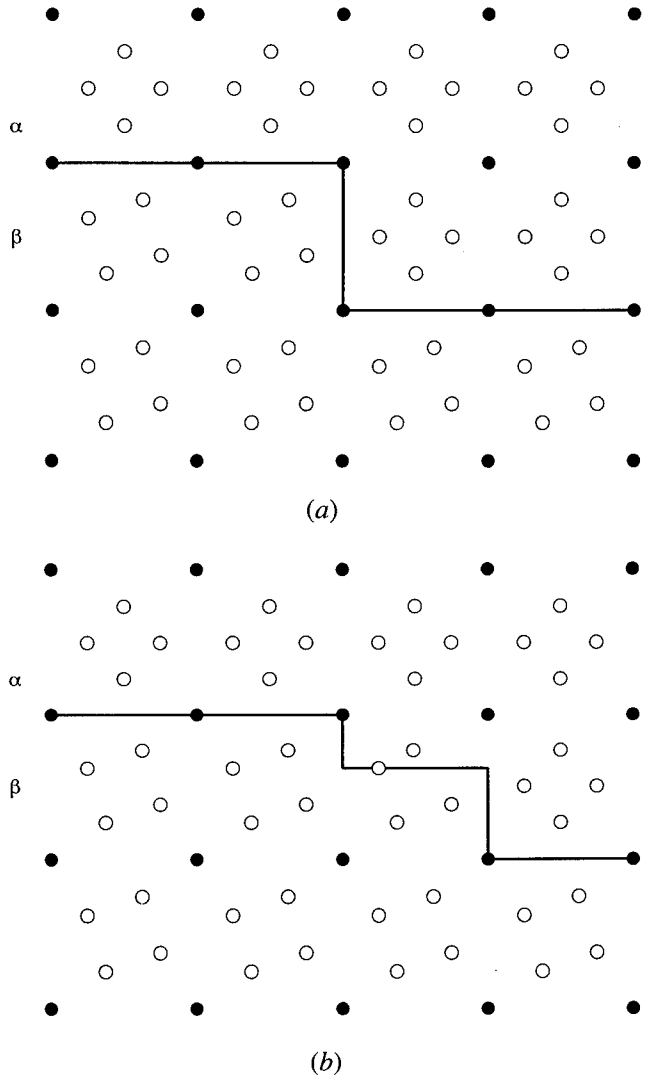


Fig. 12—(*a*) Ledge at interface and (*b*) partial ledges at interface.

interface structure. For example, we suppose that two crystals have a habit plane of the type illustrated in Figure 5, with a twist dislocation array of screw dislocations and an added screw dislocation array to accommodate a pure shear misfit in the interface plane. In the isotropic elastic case, the two superposed dislocation arrays would produce an equilibrium interface with no long-range stresses. In the anisotropic elastic case, the relative rotations of the two crystals reduce the symmetry (to that of the dichromatic complex) and rotate the dislocation line directions relative to the crystal symmetry axes. Both effects lead to added normal stresses in the boundary region, which give rise to long-range coherency stresses in both crystals (a specific example has been presented for Σ 5 and Σ 13 twist boundaries).[61] Thus, a long-period added misfit array of edge dislocations would be needed to restore equilibrium. Without such an added array, the interface could exist in a state of partial relaxation (of in-plane shear stresses) and partial coherency (with associated normal stresses). Observations of such partial relaxation are common in artificial interphase interfaces for thin films and multilayers.[62]

Indeed, rotations of the type illustrated in Figure 11(a) have been observed[63] and analyzed by symmetry theory.[64]

VII. CONCLUSIONS

1. Unit ledges can act as structural ledges if specific symmetry requirements are satisfied, provided that translational and misfit arrays of dislocations are present to remove long-range stress fields; there is no symmetry constraint to these ledges acting as pattern shifting growth ledges as well.
2. When the translational dislocations have a tilt component, deviations, usually small, from low-index habit planes and orientation relationships are expected.
3. Multiple height ledge arrays generally reduce the short-range elastic fields of the arrays, a factor that may favor the presence of such arrays, both during growth and at equilibrium.
4. Faulted partial ledge configurations are possible for low-symmetry crystals and crystals with large unit cells.

ACKNOWLEDGMENT

The author is grateful for the support of this research by the National Science Foundation under Grant No. DMR 9119342 and for helpful comments by H.I. Aaronson, J.W. Christian, and C.M. Wayman.

REFERENCES

1. J.W. Gibbs: in *Scientific Papers*, Dover, New York, NY, 1961, vol. 1, p. 325.
2. W. Kossel: *Nach. Ges. Wiss. Gottingen*, 1927, pp. 135-49.
3. I.N. Stranski: *Z. Phys. Chem.*, 1928, vol. 136, pp. 259-78.
4. W.K. Burton, N. Cabrera, and F.C. Frank: *Phil. Trans. R. Soc. London*, 1950, vol. A243, pp. 299-358.
5. M.G. Hall, H.I. Aaronson, and K.R. Kinsman: *Surf. Sci.*, 1972, vol. 31, pp. 257-74.
6. D.H. Warrington and W. Bollmann: *Phil. Mag.*, 1972, vol. 25, pp. 1195-99.
7. J.P. Hirth and R.W. Balluffi: *Acta Metall.*, 1973, vol. 21, pp. 929-42.
8. R. Bonnet and F. Durand: *Phil. Mag.*, 1975, vol. 32, pp. 997-1006.
9. R.W. Balluffi, A. Brokman, and A.H. King: *Acta Metall.*, 1982, vol. 30, pp. 1453-70.
10. R.W. Balluffi and G.B. Olson: *Metall. Trans. A*, 1985, vol. 16A, pp. 529-41.
11. F.J.J. van Loo, B. Pierragi, and R.A. Rapp: *Acta Metall.*, 1990, vol. 38, pp. 1769-80.
12. J.P. Hirth: *Metall. Trans. A*, 1991, vol. 22A, pp. 1331-38.
13. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. TMS-AIME*, 1953, vol. 197, pp. 1503-14.
14. J.S. Bowles and J.K. MacKenzie: *Acta Metall.*, 1954, vol. 2, pp. 129-47.
15. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-74.
16. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, pp. 1029-34.
17. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, pp. 1963-72.
18. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 351-76.
19. T. Furuhara, J.M. Howe, and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2873-86.
20. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer-Verlag, Berlin, 1970.
21. D.A. Smith and R.C. Pond: *Int. Met. Rev.*, 1976, vol. 205, pp. 61-74.
22. R.C. Pond: in *Dislocations in Solids*, F.R.N. Nabarro, ed., Elsevier, Amsterdam, 1989, vol. 8, pp. 1-66.
23. D.A. Smith and G.J. Shiflet: *Mater. Sci. Eng.*, 1987, vol. 86, pp. 67-92.
24. J.M. Howe and D.A. Smith: *Acta Metall. Mater.*, 1992, vol. 40, pp. 2343-50.
25. F.C. Frank: in *Rep. Symp. on the Plastic Deformation of Crystalline Solids*, Carnegie Institute of Technology, Pittsburgh, PA, 1950, pp. 150-54.
26. B.A. Bilby, R. Bullough, and E. Smith: *Proc. R. Soc. London*, 1955, vol. A231, pp. 263-73.
27. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., Krieger, Melbourne, FL, 1992, chs. 8 and 19.
28. R. Bullough: *Phil. Mag.*, 1965, vol. 12, pp. 1139-41.
29. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., Krieger, Melbourne, FL, 1992, p. 22.
30. W.T. Read and W. Shockley: in *Imperfections in Nearly Perfect Crystals*, Wiley, New York, NY, 1952, pp. 352-76.
31. J.W. Matthews and A.E. Blakeslee: *J. Cryst. Growth*, 1974, vol. 27, pp. 118-25.
32. S.A. Dregia and J.P. Hirth: *J. Appl. Phys.*, 1991, vol. 69, pp. 2169-75.
33. S. Amelinckx and W. Dekeyser: *Solid State Phys.*, 1959, vol. 8, pp. 325-499.
34. R.C. Pond and J.P. Hirth: *Solid State Phys.*, 1993, in press.
35. J.H. van der Merwe, G.J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, pp. 1165-75.
36. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1988, vol. 22, pp. 567-72 and 575-76.
37. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 649-58.
38. R.W. Fonda, W.A. Cassada, and G.J. Shiflet: *Acta Metall. Mater.*, 1992, vol. 40, pp. 2539-46.
39. A. Garg, Y.C. Chang, and J.M. Howe: *Acta Metall. Mater.*, 1993, vol. 41, pp. 235-51.
40. J.M. Howe and N. Prabhu: *Acta Metall. Mater.*, 1990, vol. 38, pp. 881-88 and 889-96.
41. C.M. Wayman: *Introduction to the Crystallography of Martensite Transformations*, Macmillan, New York, NY, 1964.
42. G.R. Srinivasan and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, pp. 621-36.
43. J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 799-803.
44. R. Bullough and B.A. Bilby: *Proc. Phys. Soc.*, 1956, vol. 69B, pp. 1276-86.
45. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., Krieger, Melbourne, FL, 1992, p. 317.
46. C. Laird and H.I. Aaronson: *Acta Metall.*, 1969, vol. 17, pp. 505-19.
47. J.M. Howe, H.I. Aaronson, and R. Gronsky: *Acta Metall.*, 1985, vol. 33, pp. 639-58.
48. B.A. Bilby: *Phil. Mag.*, 1953, vol. 44, pp. 782-85.
49. D.A. Smith: *Scripta Metall.*, 1987, vol. 21, pp. 1009-10.
50. A.W. Sleeswyk: *Acta Metall.*, 1962, vol. 10, pp. 705-25.
51. F.C. Frank and M.B. Ives: *J. Appl. Phys.*, 1960, vol. 31, pp. 1996-2004.
52. N. Cabrera and D.A. Vermilyea: in *Growth and Perfection of Crystals*, R.H. Doremus, B.W. Roberts, and D. Turnbull, eds., Wiley, New York, NY, 1958, pp. 393-99.
53. R. Trivedi: *Metall. Trans.*, 1970, vol. 1, pp. 921-27.
54. C. Atkinson: *Proc. R. Soc. London*, 1981, vol. A378, pp. 351-68.
55. M. Enomoto: *Acta Metall.*, 1987, vol. 35, pp. 935-45 and 947-56.
56. W.W. Mullins and J.P. Hirth: *J. Phys. Chem. Solids*, 1963, vol. 24, pp. 1391-404.
57. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., Krieger, Melbourne, FL, 1992, pp. 785-86.
58. G.B. Olson: *Scripta Metall.*, 1987, vol. 21, pp. 1023-28.
59. J.W. Christian and D.V. Edmonds: *Scripta Metall.*, 1988, vol. 22, pp. 573-74 and 577-80.
60. J.P. Hirth: *Acta Metall.*, 1974, vol. 22, pp. 1023-31.
61. J.P. Hirth and B. Carnahan: *Acta Metall. Mater.*, 1992, vol. 40, pp. 1237-42.
62. B.W. Dodson and J.Y. Tsao: *Appl. Phys. Lett.*, 1987, vol. 51, pp. 1325-27.
63. R. Du and C.P. Flynn: *J. Phys. Condensed Matter.*, 1990, vol. 2, pp. 1335-41.
64. P.J. Dugdale, R.C. Pond, and R. Beanland: *J. Phys.*, in press.

The Role of Structural Ledges as Misfit-Compensating Defects: fcc-bcc Interphase Boundaries

G.J. SHIFLET and J.H. VAN DER MERWE

An energetic justification is given for structural ledges as interfacial misfit compensating defects, based on elastic considerations. We detail the NW-x configuration in which the orientation is imposed by close matching along the ⟨211⟩fcc and ⟨110⟩bcc directions and {111}fcc and {110}bcc planes, respectively. Special prominence is given to the terrace between structural ledges, where it is shown that an elastic relaxation obtains at specific terrace lengths based on atomic matching. It is this relaxation that makes a stepped interface more stable than a planar one at all meaningful values of misfit.

I. INTRODUCTION

THIS article will review, extend, and elucidate our previous work on fcc:bcc interfaces containing linear defects that accommodate misfit.[1–6] Emphasis in our analysis will be placed on consideration of the relative energy savings by inclusion of structural ledges[7] as a misfit-compensating defect. However, rather than duplicate extensive mathematics that has already been published for the fcc:bcc interface in our previous efforts, we will discuss certain aspects of the theory, to make clear the advantages of such an approach toward understanding the energetics of interfacial structures. A large portion of this article will be concerned with the terraces between structural ledges. We will consider a rigid lattice calculation as well as allow the interface to relax elastically. As the interfacial atomic planes of the most studied interfaces have rhombohedral symmetry this will be dealt with here; rectangular meshes have been addressed in earlier papers[1–4] and differ only in the symmetry of the interaction potential.

A. *The {111}fcc-{110}bcc Interface*

The {111}fcc closed packed surface interacting with the {110}bcc surface at azimuthal orientation of current interest is shown in Figure 1, which is a "map" of the relative energetics of the primary fcc:bcc orientation relationships.[8] In this figure, the azimuthal rotations vary from $\theta = 0$ to 18 deg, with the rotation relative to the Nishiyama–Wasserman (NW) orientation placed at 0 deg; r is the ratio of nearest neighbor atoms, b/a, see (Figure 2(a)). Although a larger angular range would illustrate other energy minima, this is sufficient to highlight the two most commonly observed fcc:bcc orientation relationships. The deep minima corresponding to NW[9,10] and Kurdjumov–Sachs (KS)[11] ($r = r_d = 1.09$, $\theta = 5.26$ deg) orientation relationships are indicated. The NW is further divided into NW-x ($r = r_x = 0.94$, $\theta = 0$) and NW-y ($r = r_y = 1.15$, $\theta = 0$), where atomic matching is in the x or y direction, respectively.

A plot of the overlapping rhombuses is shown in Figure 2(a). Here, the various parameters used to construct Figure 1 are shown geometrically. By varying r and θ, the geometric patterns associated with the NW and KS relationships can be mapped out (Figures 2(b) through (d)). This can be understood by comparing Figure 1 with the atomic sketches in Figures 2(b) and (c) for the NW-x and -y, respectively, and Figure 2(d) for the KS orientation relationship. The relative lattice rotations and alignments with respect to the x and y axes make the correlations clear. The importance of Figure 1 is that it reveals a method of ascertaining relative energies of the orientations. From an analysis of Figures 1 and 2, the following important conclusion can be made, *viz.*, ideality is equivalent to row matching. This important statement was first made by Dahmen[12] and, subsequently, for epitaxial interfaces, by van der Merwe.[8] A close examination of Figures 2(b) through (d) reveals that each corresponds to alignment of rows of atoms of crystals A and B. Specifically, Figure 2(b) corresponds to the minimum $r = r_x$, $\theta = 0$ in Figure 1, while Figure 2(c) corresponds to $r = r_y$, $\theta = 0$ and for KS (Figure 2(d)), $r = r_d$, $\theta = 5.26$ deg.

Depending on the interfacial misfit, the plot clearly shows the continuum between the orientation relationships. It is expected that each orientation relationship will dominate over a range of r. Previously, we chose $0.95 < r < 1$ for NW-x and $1 < r < 1.1$ for KS. In most instances, these ranges cover common fcc-bcc interfaces.

II. GEOMETRY

An understanding of the interface geometry that leads to the concept of pattern advance and hence structural ledges necessitates that relevant geometric concepts be reviewed. This is useful because it lends itself to a simple introduction of the theory. A Cartesian coordinate system is employed (Figure 2(a)). The z-axis is normal to the interface, and the x and y axes are parallel to the interface plane. A reference lattice is used to describe the macroscopic interface.[13] For the x direction, the reference has a lattice spacing c_x, where c_x, the vernier period p_x, and the misfit f_x can be expressed in terms of

G.J. SHIFLET, William G. Reynolds Professor, is with the Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA 22903. J.H. VAN DER MERWE, Professor, is with the Department of Physics, University of South Africa, Pretoria 0001, Republic of South Africa.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

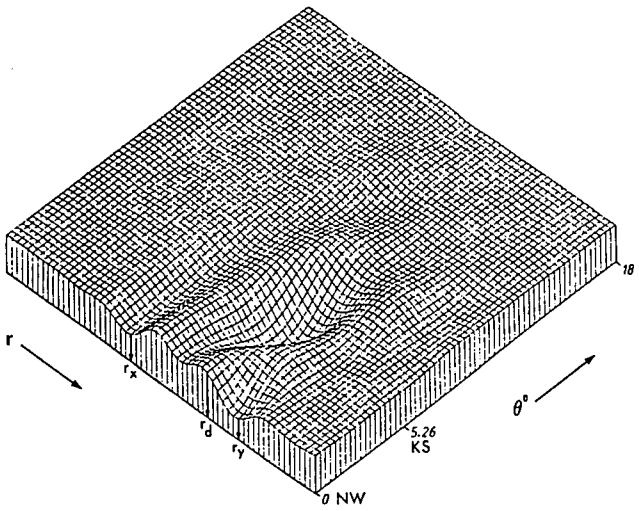


Fig. 1—Surface map of the mean energy per atom for a rectangular island containing rhombohedral meshes of atoms. The interaction is between the {111}fcc and {110} lattices. The minima correspond to row matching for the NW-x, NW-y, and KS orientation relationships.[8]

the spacing a_x (bcc crystal A) and b_x (fcc crystal B) of the real lattices as

$$p_x = P_x b_x = (P_x + 1)a_x = \left(P_x + \frac{1}{2}\right)c_x \qquad [1a]$$

so that

$$c_x = \frac{b_x a_x}{1/2\,(b_x + a_x)}, \qquad p_x = \frac{a_x b_x}{b_x - a_x},$$

$$f_x = \frac{c_x}{p_x} = \frac{b_x - a_x}{1/2\,(b_x + a_x)} \qquad [1b]$$

Here, we have $b_x > a_x$, which gives $p_x > 0$. For the other case, $b_x < a_x$, $p_x = a_x b_x/(a_x - b_x)$. In the following, we will let $b_y > a_y$ so that $b_i > a_i$, where $i = x, y, z$. Further, in Eq. [1a], which really defines the reference lattice, P_x is meant to be an integer. This implies that the calculations based on Eqs. [1a] and [1b] are valid for discrete values of the misfit only. In the present calculations, P_x will be varied continuously. In the harmonic model of the crystal, the curve of the interfacial energy *vs* misfit has cusped minima at the discrete points. The error involved in the continuum approximation (through the discrete points) only becomes significant at large misfits.[14,15]

Figure 3 illustrates a cross-grid of ledges enclosing a checkerboard of rectangular terrace patches. The results for any special case, for example, the case in which there is only one parallel sequence of ledges, can then be simply obtained by assigning appropriate values to the relevant parameters. It is convenient to refer to ledges as x ledges and y ledges when they face, respectively, in the x and y directions. Figure 4 then, is a standard patch pair taken in profile from points Q to R in Figure 3. Here, an atom at the center of the A terrace coincides with the potential minimum in the opposing B terrace. Let there be $(2M + 1)$ A atomic rows and, correspondingly, $(2M + 1)$ B rows of potential minima in the x direction of the standard patch pair.

The basic underpinning of the theory is that the structural ledge is responsible for a pattern advance at the interface relative to the two lattices, *viz.*,

$$\delta_x = \delta_x^a - \delta_x^b \qquad [2]$$

i.e., that the step effects a forward displacement of the interfacial atomic pattern of A relative to that of B, where δ_x^a and δ_x^b are the individual forward pattern advances of crystals A and B in the x direction (Figure 4). When the misfit is positive ($b_x > a_x$), the mismatch which builds up along the interface is summarily canceled by this relative shifting in atomic positions. To then consecutively cancel interfacial misfit, the structural ledges must be periodic with spacing l_x, which accounts for the terrace width plus pattern advance. Also, this ensures that the atomic patterns do not get out of synchronization. The relationships describing this can be written as follows:

$$l_x = L_x a_x + \delta_x^a = L_x b_x + \delta_x^b \qquad [3]$$

$$= \frac{a_x b_x}{b_x - a_x}\left(\frac{\delta_x^a}{a_x} - \frac{\delta_x^b}{b_x}\right) \qquad [4]$$

In this equation, L_x interfacial atoms of one lattice match L_x potential troughs of the other lattice.

Because there is misfit in a direction perpendicular from the interface plane, this must also be compensated. For example, if the atomic plane spacing in B is smaller than in A, $\delta_z = b_z - a_z < 0$, and an extra atomic plane from crystal B must terminate at the boundary for every P_z atomic plane for both crystals (Figure 4, inset, for $b_z - a_z > 0$). The term P_z is defined by Eq. [1a], but with z instead of x. This construction constitutes the tilt misfit dislocation (TMD) (Figure 3). The spacing, $\bar{p}_x$ of the TMDs can be written as

$$\bar{p}_x = \frac{a_x b_x}{b_x - a_x}\left(\frac{\delta_x^a}{a_x} - \frac{\delta_x^b}{b_x}\right) P_z - \frac{a_x b_x}{b_x - a_x}\cdot\frac{\delta_x^b}{b_x} \qquad [5]$$

where the bar signifies an average spacing.

Shear pattern advances must also be considered, *i.e.*, δ_{xy} and δ_{yx}. This would effect a relative displacement parallel to the ledge of patches of good fit and would accordingly cause an increase of energy as compared to the case where δ_{ij} vanishes.

As noted, greatly simplifying the analysis is the result that the misfit in the x and y directions can be represented by a single misfit parameter, $r = b/a$, where b and a are the nearest neighbor distances in lattices B and A, respectively (Figure 2(a)) (rather than lattice parameters as in Reference 16. Making use of the nearest neighbor ratio, r, useful parameters can be written. Misfits (Eq. [16]) can then be defined as

$$f_x = \frac{r/r_x - 1}{(r/r_x + 1)/2} \quad \text{and} \quad f_y = \frac{r/r_y - 1}{(r/r_y + 1)/2} \qquad [6a]$$

for the NW orientation and as

$$f_d = \frac{d_b - d_a}{(d_b + d_a)/2} = \frac{r/r_d - 1}{(r/r_d + 1)/2} \qquad [6b]$$

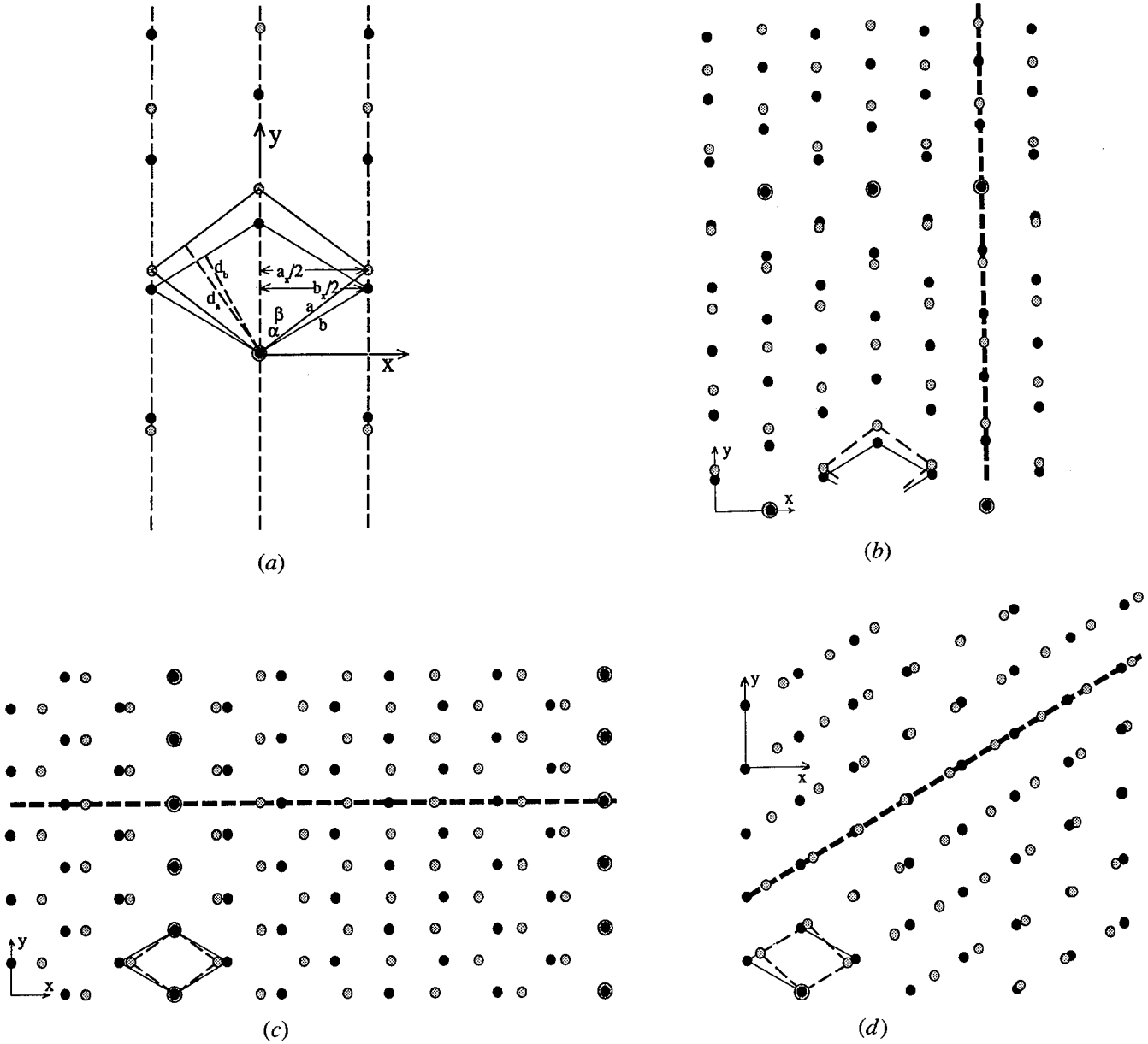


Fig. 2—(*a*) A schematic diagram representing the rhombic surface unit cell structures and related crystal directions of an fcc(111) (β = 60 deg) (filled circles) and bcc(110) (α = 54.7 deg) (open circles). Other lengths and directions are described in the text. (*b*) By altering the rhombuses so that the fcc and bcc atoms form rows in the *x* direction, the NW-*x* orientation relationship is obtained. This occurs for $r = b/a = 0.94$. (*c*) When $r = b/a = 1.15$, the rows of atoms align in the *y* direction, and this is referred to as NW-*y*. (*d*) Alignment of the rhombus edge at $r = b/a = 1.01$ yields the KS orientation relationship. The heavy dashed line in (b) through (d) indicates the row matching described in the text.

for the KS orientation.

In these equations, f_x and f_y are the misfits as measured along the long and short diagonals of the rhombic unit cell, respectively. Specifying the geometric parameters for the fcc{111}-bcc{110} interface, the spacings a_i and b_i ($i = x, y$) are the diagonals of the interfacial unit cell rhombuses in Figure 2. For matching ($b_i = a_i$) in the *x* and *y* directions, respectively, from Fig. 2, we can then also write

$$r_z = \frac{\sin\alpha}{\sin\beta} = \sqrt{\frac{8}{9}} \approx 0.943 \equiv r_{\text{NW}-x};$$

$$r_y = \frac{\cos\alpha}{\cos\beta} = \frac{2}{\sqrt{3}} \approx 1.155 \equiv r_{\text{NW}-y}$$

$$r_d = \frac{\sin\alpha\cos\alpha}{\sin\beta\cos\beta} = \sqrt{\frac{32}{27}} \approx 1.089 = r_{KS} \quad [7]$$

where $\alpha = \arctan\sqrt{2}$ and $\beta = 60$ deg. When the ideal NW-*x* occurs, $f_x = 0$ and $r = r_x$, the misfit in the *y* direction is about 20 pct. This fact allows consideration of ledges in one direction, in this case *x*, whereas in the perpendicular direction, misfit accommodation is by misfit vernier[8] in which the oscillatory strains associated with (misfit) dislocations are absent.

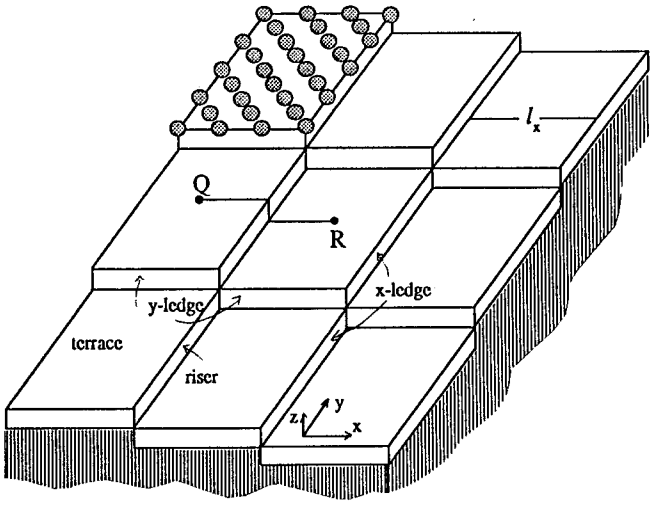
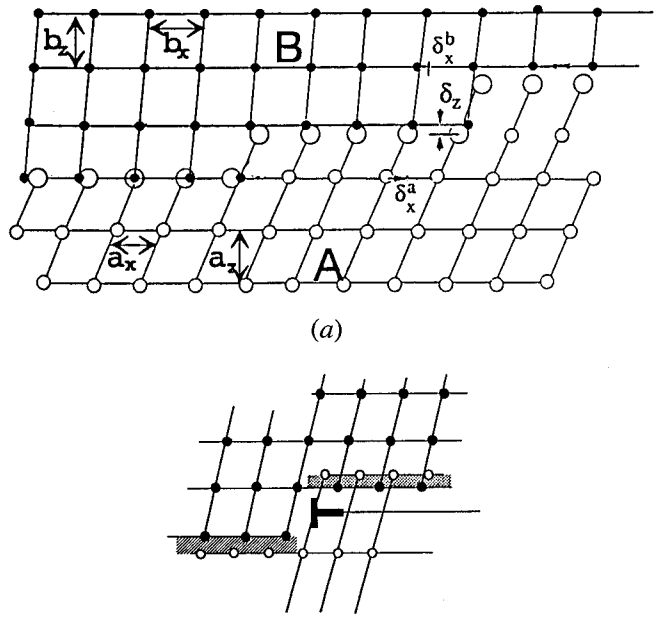


Fig. 3—A three-dimensional perspective drawing showing a series of structural ledges in the x and y directions.



(*a*)

(*b*)

Fig. 4—(*a*) Atomic model illustrating a perspective of the line QR in Fig. 3. At the interface, the big open circles represent the potential troughs for the interfacial atoms (solid circles) of crystal B in the potential field emanating from atoms of crystal A. (*b*) An enlargement of the model where now $\delta_z = b_x - a_z > 0$, which is compensated for by a misfit dislocation.[20]

$$r = \frac{b}{a}, \quad a_x = \sqrt{\frac{8}{3}}a, \quad a_y = \frac{2a}{\sqrt{3}}, \quad a_z = \sqrt{\frac{2}{3}}a$$

$$b_x = \frac{\sqrt{3}}{2}ra, \quad b_y = b = ra, \quad b_z = \sqrt{\frac{2}{3}}b = \sqrt{\frac{2}{3}}ra$$

$$\delta^a_x = \frac{1}{2}a_x, \quad \delta^b_x = \frac{1}{3}b_x \qquad [8]$$

The usefulness of the parameter r can be illustrated by expressing these values in terms of r and a, for example, the forward pattern advance for the NW-x,y ledges is

$$\delta_x = \sqrt{\frac{2}{3}}(1 - r/\sqrt{2})a$$

$$\delta_y = \frac{1}{\sqrt{3}}(1 - r\sqrt{3}/2)a \qquad [9a]$$

with a corresponding shear displacement of

$$\delta_{yx} = 0, \quad \delta_{xy} = -\frac{ra}{2\sqrt{3}} \qquad [9b]$$

for x and y ledges, respectively. It is the shear displacement of terraces that makes the NW-y orientation relationship less energetically favorable.[5] The fact that the shear pattern advances vanish for the NW-x ledges—ledges parallel to the bcc $\langle\bar{1}10\rangle$—but not for the NW-y ledges suggests that the former is more energetically favored for structural ledge misfit accommodation than the latter; accordingly, we focus on interfaces with x ledges only.

The corresponding quantities for the KS orientation (applicable when r is near r_{KS}) is obtained from the NW orientation by a relative rigid azimuthal rotation of about 5 deg. Its occurrence is related to matching of atomic rows along parallel sides of the corresponding unit cell rhombuses (Figures 2(a) through (d)). When there is only a sequence of ledges parallel to the matching closest packed rows,

$$a_x = d_a = \sqrt{8}\frac{a}{3}, \quad b_x = d_b = \sqrt{3}\frac{ra}{2}$$

$$\delta^a_x = \frac{1}{2}d_a, \quad \delta^b_x = \frac{1}{3}d_b, \quad \delta^a_{yx} = \frac{1}{3}a, \quad \delta^b_{yx} = \frac{1}{2}b \qquad [10]$$

where d_a and d_b are the atomic row spacings for the KS orientation (Figure 2). Note that the parameters in Eqs. [5] and [6] for the NW and KS, respectively, can be substituted into Eq. [4] to obtain the corresponding l_x parameters.

III. ENERGY CALCULATIONS

A. *Introduction*

So far, we have summarized the geometrical formalism for the description of misfit and misfit accommodation for planar and stepped fcc:bcc interfaces. For the latter, we have introduced the concept of pattern advance, the structural ledge, and the necessity for the presence of tilt-type misfit dislocations. What remains is to analyze the energetics of such an interface with emphasis on an extendible mathematical formalism. The

derivations are published elsewhere,[4,5,6] and this section will review and detail these ideas specifically for fcc:bcc interfaces.

B. *Terraces: Rigid-Body Considerations*

In order to evaluate the significance of the energy gain due to elastic relaxation, it is also helpful to explore the rigid-body interaction energy. This is calculated assuming that the A terrace patch interacts with a B atom (in the opposing B terrace) according to a potential which has the periodicity and symmetry of the atomic arrangement in the A terrace. The rigid lattice calculations are useful in that they provide a general understanding of the range of energies involved and information concerning the misfit between the fcc and bcc lattices. Figure 1 comes from this approach. Also, we will demonstrate the maximum terrace length l that will be energetically profitable. That is, the formalism confirms energetically and extends what was arrived at from geometrical considerations only.

First consider an entire checkerboard of $(2J + 1)^2$ terrace patches, the extent of which is one superperiod (Figure 3) in each direction. The checkerboard pattern consists of identical rectangular terraces and that have no azimuthal misorientation. In the present analysis of fcc:bcc interfaces, the orientation is NW-x. The goal is a summation of the potential energy of all the atoms in the fcc crystal in the periodic field of the bcc lattice for the entire checkerboard. This is followed by a summation over the interfacial atoms of the fcc crystal in the periodic field of the bcc lattice.

A truncated Fourier representation in the field emanating from the bcc crystal is given by[5,17]

$$V_A(x, y) = V_0^A \left[1 + A_1 \left\{ \cos 2\pi \left\langle \frac{x}{a_x} - \frac{y}{a_y} \right\rangle + \cos 2\pi \left\langle \frac{x}{a_x} + \frac{y}{a_y} \right\rangle \right\} + A_3 \cos \left\langle \frac{4\pi x}{a_x} \right\rangle + A_4 \cos \left\langle \frac{4\pi y}{a_y} \right\rangle \right] \quad [11a]$$

The authors have estimated the scale factor V_0^A and the normalized coefficients A_i as

$$A_1 = -0.573, \quad A_3 = 0.088, \quad A_4 = -0.015; \quad V_0 = 1.183 \text{ eV} \quad [11b]$$

for a W adatom on $W\{110\}$. Whereas the A_i's are dominated by the substrate symmetry, V_0^A depends strongly on the materials combination and environment, for example, whether the adatom is part of a monolayer or is an isolated adatom.

The {111}fcc surface, similarly truncated to KS and NW orientations, yields an expression[18] equivalent to Eq. [11a] as

$$V_B(x, y) = V_0^B \left\{ V_{00} + V_{10} \left[\cos 2\pi \left\langle \frac{x}{b_x} + \frac{y}{b_y} + \frac{1}{3} \right\rangle + \cos 2\pi \left\langle \frac{x}{b_x} - \frac{y}{b_y} - \frac{2}{3} \right\rangle + \cos 2\pi \left\langle -\frac{2x}{b_x} + \frac{1}{3} \right\rangle \right] + V_{11} \left[\cos 2\pi \left\langle \frac{2y}{b_y} \right\rangle + \cos 2\pi \left\langle -\frac{3x}{b_x} - \frac{y}{b_y} \right\rangle + \cos 2\pi \left\langle \frac{3x}{b_x} - \frac{y}{b_y} \right\rangle \right] \right\} \quad [12a]$$

$$V_{00} = \frac{3}{2}(V_{10} - 2V_{11}), \qquad V_{mm} = V(0, 0) = 0,$$

$$V_{\max} = V\left[\frac{b_x}{6}, -\frac{b_y}{2}\right] = V_{10}V_0^B/K, \qquad K = \frac{2}{9} \quad [12b]$$

The potential surface generated by Eq. [12] has six-fold symmetry about the minimum position $x = 0$, $y = 0$. This is unphysical, and stacking fault terms would normally need to be added. These stacking fault terms may be neglected in the present case,[5] as they emanate from atomic layers lower down. Embedded atom methods calculations[19] predict that these stacking fault terms are small and contribute insignificantly to the present estimate of energy.

A single terrace is shown in Figure 5, which represents one of the rectangles from Figure 3. The orientation relationship is NW-x, and the bcc potential trough positions are shown as open circles. The ledges are aligned along the y direction, which is [001] bcc. The direction of pattern advance, δ^a, is [$\bar{1}$10] bcc (or the x direction). For this calculation, the atoms are considered to be located on two interpenetrating rectangular lattices i = 1, 2. In general, the terrace patch contains $4M \pm 1$ atomic rows in the x direction and $4N \pm 1$ rows in the y direction, the x direction being perpendicular to the structural ledges. The "±" signs indicate that counting starts at the

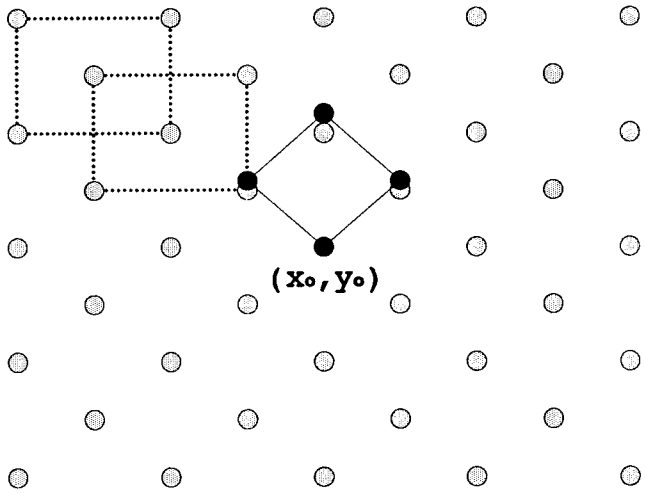


Fig. 5—Single terrace from Fig. 3, where in this case, $M = N$. Open circles represent bcc potential troughs and solid circles represent positions of fcc atoms. x_0, y_0 is the position of the central fcc atom.

centers of patches, the "center" being an atom in A and a trough in B. For the symmetric ($M = N$) patch shown in Figure 5, the positions of the fcc terrace atoms (solid circles) with respect to the coordinate origin, which is fixed at the center of the bcc terraces, can be given as

$$x_m = x_0 + \frac{1}{2} b_x m^{(i)}; \qquad y_n = y_0 + \frac{1}{2} b_y n^{(i)},$$

where

$$m^{(1)} = 0, \pm 2, \ldots, \pm 2(2M - 1), \pm 2M$$
$$\text{and } n^{(1)} = 0, \pm 2, \ldots, \pm 2(N - 1), \pm 2N$$
$$m^{(2)} = \pm 1, \pm 3, \ldots, \pm (2M - 1)$$
$$\text{and } n^{(2)} = \pm 1, \pm 3, \ldots, \pm (2N - 1) \qquad [13]$$

(x_0, y_0) being the position of the central fcc atom.

The number of atoms within the {111}fcc terrace interface is $G = \{(2M + 1)(2N + 1) + 4MN\}$. Nonsymmetric cases can also be detailed in a similar fashion.[8] We can then substitute the relations of Eq. [13], which give all of the interfacial atom positions of crystal B, into Eq. [11a] and perform the summation, yielding the average energy per atom

$$\bar{v}_A = V_0^A \left[1 + \frac{2A_1}{G} \cos\left(\frac{2\pi x_0}{a_x}\right) \cos\left(\frac{2\pi y_0}{a_y}\right) \cdot F_{MN}\left(\frac{b_x}{a_x}, \frac{b_y}{a_y}\right) + \frac{A_3}{G} \cos\left(\frac{4\pi x_0}{a_x}\right) F_{MN}\left(2\frac{b_x}{a_x}, 0\right) + \frac{A_4}{G} \cos\left(\frac{4\pi y_0}{a_x}\right) F_{MN}\left(0, 2\frac{b_y}{a_y}\right) \right] \qquad [14a]$$

$$F_{MN}(k_x, k_y) = \frac{\sin\{(2M + 1)\pi k_x\} \sin\{(2N + 1)\pi k_y\} + \sin(2M\pi k_x) \sin(2N\pi k_y)}{\sin(\pi k_x) \sin(\pi k_y)} \qquad [14b]$$

$$F_{MN}(1, 1) = F_{MN}(2, 0) = F_{MN}(0, 2) = G \qquad [14c]$$

and $F(k_x, k_y)/G \approx 0$ when either k_x or k_y is significantly different from an integer or zero. This is the type of relation generalized to account for varying azimuthal orientation, which is plotted to give Figure 1.

For this interface, when f_x is small ($b_x \approx a_x$), the f_y will be large. The ideal case can be examined, which is $b_x/a_x = 1$. The energetics of NW-x ledges are determined by the region $b_x/a_x = 1$ close to the minimum of the term in A_3. For $b_x/a_x = 1$, $F_{MN}(2, 0) = G$, *i.e.*, large, while the F_{MN} factors of A_1 and A_4 are then very much less than G. Also, the primary objective is to compare the energetics of misfit-compensating ledges in stepped interfaces with those of misfit-accommodating dislocations, if the interface is planar. We may accordingly write approximately

$$\bar{v}_A \approx V_0^A \left[1 + A_3 \cos\left(\frac{4\pi x_0}{a_x}\right) \right] \qquad [15a]$$

Because A_3 is positive (≈ 0.088), a rigid translation of $x_0 = a_x/4$ is required to obtain a stable minimum energy configuration. This conclusion relies on the fact that small f_x ($b_x \cong a_x$) implies large f_y and hence little strain relaxation[8] in the y direction. The depth of the energy well—referred to as the average V_0^A when both $|f_x|$ and $|f_y|$ are significantly different from zero—can then be written as

$$\Delta\bar{v}_A = V_0^A A_3 \qquad [15b]$$

A similar approach can be applied to the y ledges when $b_y \approx a_y$. The result would be similar to Eq. [15a] with the minimum at $y_0 = 0$ and A_d instead of A_3.

The terrace patch size can now be considered energetically. The energy savings that are made by a bcc/fcc lattice orienting into the NW-x relation ($b_x \approx a_x$, $x_0 = a_x/4$) is determined by

$$\bar{v}_A = V_0^A \cdot \left[1 - \frac{A_3}{G} \frac{(2N + 1) \sin\{2\pi(2M + 1)b_x/a_x\} + 2N \sin(4\pi M b_x/a_x)}{G \sin(2\pi b_x/a_x)} \right] \approx V_0^A \left[1 - \frac{A_3}{G} \frac{(2N + 1) \sin\{2\pi(2M + 1)f_x\} + 2N \sin(4\pi M f_x)}{\sin(2\pi f_x)} \right] \qquad [16]$$

The approximation was secured by using the relation $b_x/a_x \approx (1 + f_x)$. The error involved is very small, being <1 pct when $f_x < 20$ pct.

The relation, Eq. [16], provides support for the introduction of the cancellation criterion and facilitates demonstration of the physical significance of the criterion. When regarded as a function of f_x, the quantity $\bar{v}_A$ has a global minimum of $V_0^A(1 - A_3)$ at $f_x = 0$. As $|f_x|$ increases from zero, the oscillatory terms increase from negative values upward—the first vanishes at $2(2M + 1)|f_x| = 1$ and the second at $4M|f_x| = 1$. We may take half of their average as a certain upper limit of the region in which the energy is still well below the average V_0^A, *i.e.*, the terrace size $(4M + 1)$ is certainly energetically favorable below an upper limit defined by

$$(4M + 1)|f_x| = 2 \qquad [17a]$$

for a limited interval of r around r_x

$$|f_x| = \frac{2|r - r_x|}{r + r_x} \approx \frac{|r - r_x|}{r} \qquad [17b]$$

so that Eq. [17a] may be written in the form

$$(4M + 1) = \frac{2r}{|r - r_x|} \qquad [17c]$$

In terms of the number $2M + 1$ of atomic rows and the row spacing $a_x/2$, the terrace width l_x is $(2M + 1)a_x/2$. It accordingly follows from Eq. [4] and [8] that the terrace width that satisfies the cancellation criterion is given by

$$l_x = \frac{1}{3} \frac{2r}{|r - r_x|} \frac{a_x}{2} \equiv (4M + 1) \frac{a_x}{2} \qquad [18a]$$

i.e., the number of atomic rows in a NW-x terrace that satisfies the cancellation criterion is given by

$$(4M + 1) \approx \frac{r}{3|r - r_x|} \qquad [18b]$$

It is seen that this terrace width more than satisfies the energetically favorable one in Eq. [17b].

In order to quantify the energetics, let us investigate a simplification of Eq. [16], simplified along the lines that have led to [17a] by reducing Eq. [16] to a single term and substituting from Eqs. [17b] and [18b] to obtain the approximation

$$\bar{v}_A = V_0^A \left\{ 1 - \frac{2A_3(2N + 1/2) \sin [\pi(4M + 1)|f_x|]}{G \sin (2\pi|f_x|)} \right\}$$

$$= V_0^A \left\{ 1 - \frac{2A_3 \sin (\pi/3)}{\frac{r}{|r - r_x|} \sin \left[2\pi \frac{|r - r_x|}{r} \right]} \right\};$$

$$G = (4M + 1)(2N + 1/2) \qquad [19]$$

Alternatively, one may also write Eq. [16] itself in a more appropriate form by noting that the number periodicity $2N + 1$ in the y direction is indeed the number of rows spanning the misfit dislocation spacing in that direction, *i.e.*, that

$$2N + 1 \approx \frac{2}{|f_y|} = \frac{r + r_y}{|r - r_y|} \qquad [20]$$

Although this is more accurate, it does not change the essentials related to Eq. [17c]. The relation, Eq. [19], also has the merit of simplicity. The value of $\bar{v}_A$ varies little
(<7 pct) in the interval $0.85 \leq r \leq 1.01$ and has a value of

$$\bar{v}_A = V_0^A \left(1 - \frac{\sqrt{3}}{2\pi} A_3 \right) \text{ at } r = r_x$$

We note that this is well below the value $\bar{v}_A = V_0^A (1 - A_3)$ that obtains from Eqs. [16] and [14c] for perfect matching ($r = r_x$; $b_x = a_x$) on a planar interface. The discrepancy is due to "rigidity" and the cancellation criterion which relates terrace width and misfit.

The relation corresponding to Eq. [19], but for KS ledges, is

$$\bar{v}_A = V_0^A \left[1 - |A_1| \sin \left\langle \frac{\pi}{3} \frac{3r_{KS} - 2r}{r_{KS} + r} \right\rangle \right.$$
$$\left. \Big/ \left\{ \frac{3r_{KS} - 2r}{6|r_{KS} - r|} \sin \left\langle \frac{2\pi|r - r_{KS}|}{r + r_{KS}} \right\rangle \right\} \right] \qquad [21a]$$

In this case, the value of $\bar{v}_A$ again varies little (<2 pct) over the range $1.01 < r < 1.14$ from the value

$$\bar{v}_A = V_0^A \left(1 - \frac{3}{\pi} |A_1| \right) \qquad [21b]$$

at $r = r_{KS}$ (perfect matching). While the reasons are the same as those for NW-x ledges, we also note (a) that the value is significantly closer to the ideal value with one instead of $3/\pi$, the reason being traced to the fact that NW-x ledges have second-order harmonic terms, while KS ledges have first-order harmonic terms; and (b) that the average energy gain with respect to the reference level V_o^A would be significantly greater ($|A_1|/A_3 \approx 6$) than for NW-x ledges.

The next step in the rigid lattice calculations is to obtain expressions when the bcc{110} surface atoms interact with the fcc{111} surface potential. The analysis is virtually the same, yielding for this interaction the average energy per interfacial bcc atom:

$$\bar{v}_B = \frac{V_0^B}{G} \left[V_{00} G + 2V_{10} \cos \left(2\pi \frac{x_0}{b_x} + \frac{1}{3} \right) \right.$$
$$\cdot \cos \left(2\pi \frac{y_0}{b_y} \right) F_{MN} \left(\frac{a_x}{b_x}, \frac{a_y}{b_y} \right) \Big] + V_{10}$$
$$\cdot \cos \left\{ 2\pi \left(\frac{2x_0}{b_x} + \frac{1}{3} \right) \right\} F_{MN} \left(\frac{2a_x}{b_x}, 0 \right)$$
$$+ V_{11} \cos \left(4\pi \frac{y_0}{b_y} \right) F_{MN} \left(0, \frac{2a_y}{b_y} \right)$$
$$\left. + 2V_{11} \cos \left(6\pi \frac{x_0}{b_x} \right) \cos \left(2\pi \frac{y_0}{b_y} \right) F_{MN} \left(\frac{3a_x}{b_x}, \frac{a_y}{b_y} \right) \right] \qquad [22]$$

where $F_{MN}(k_x,k_y)$ have been given in Eq. [14c]. An analysis[5] shows that the depth of the energy well can be given as

$$\Delta\bar{v}_B = V_0^B V_{10} \qquad [23]$$

An analysis concerning the cancellation criterion can be made similar to that above. The effective depth of the well for the combination of crystals A on B and B on A can be taken as the average of the quantities, *viz.*,

$$\Delta\bar{v}_{AB} = \frac{1}{2} (V_0^A A_3 + V_0^B V_{10}) \approx V_0^A A_3 \qquad [24]$$

where the approximation follows from Eq. (A2,10) of Reference 2 and the fact that $V_0^B K$ of (2) is the equivalent of V_0^B here.

The important result from this analysis is that the terrace width needed for cancellation of misfit is small enough to ensure adequate energy minimization. The rigid model calculations predicted that the maximum average energy per atom that can be gained by atoms of the B interface from interaction with crystal A, when measured from the level of energy defined by a misfitting interface of large width, is $V_0^A A_3$. This energy is obtained for perfect matching in the NW-x orientation. The rigid model analysis shows that 70 to 80 pct of $V_0^A A_3$ can still be gained below an upper bound of nonzero misfit inversely proportional to the terrace width. It has also been shown that such terrace widths also satisfy the cancellation criterion. Analogous conclusions apply to the interaction emanating from the fcc crystal. The next step is to allow the terrace patch pair to relax elastically. The average terrace energy calculated for the NW-x orientation is then measured from the minimum energy level $\Delta\bar{v}_{AB}$ (Eq. [12]) defined by exact matching ($f_x = 0$), *i.e.*,

at a depth $\Delta\bar{v}_{AB}$ (Eq. [24]) below the global average V_0^A.

C. *Terraces: Elastic Relaxation of the Terrace Patch Pair*

Now that a base line has been established with the rigid model, our next step is to go beyond the rigid lattice by applying the harmonic approximation. One of the strengths of this approach is that it allows for the energy of the terrace patch pair to be further reduced by elastic relaxation. The total mean energy per terrace patch of the combined system will decrease below the rigid model value of Eq. [19]. Howe and Smith[20] incorrectly state that our approach is incomplete, because it does not allow elastic relaxation at the interface. Our approach does allow elastic relaxation of the atoms forming the terrace interface, and in fact, we have shown that it is this energy "gain" that assists to make a ledged structure viable.

Calculation of the energy due to elastic relaxation presents two problems. The first concerns the terrace patch interface and the second deals with the TMDs. In both cases, the crystals on either side of the interface are approximated by isotropic elastic continua A and B, but their discrete atomic nature is taken into account at the interface by introducing interfacial forces that have the periodicity of the crystal lattices in the relative displacements of opposing atoms on either side of the interface;[21] c_x or c_y for the reference lattice parameters patch and $\bar{c}_z$ for the MD interface.[13,22] Also, in both cases, we decompose the problem of two-dimensional (x, y) interfacial misfit into two problems of plane strain in the "vertical" plane (xz plane), each with one-dimensional misfit, and make the simplifying approximation that the mean energies $\bar{E}_x$ and $\bar{E}_y$ are additive:[13]

$$\bar{E} = \bar{E}_x + \bar{E}_y \qquad [25]$$

Previous calculations have shown[19] this energy to be

$$\bar{E}_x^T = \frac{\mu_x U_0^2}{\pi^2 d} \sum_1^\infty \frac{1}{n(n + n_0)}$$

$$n_0 = \frac{\mu l_x}{2\pi d} \cdot \frac{1}{\lambda},$$

$$\frac{1}{\lambda} = \frac{1 - \nu_A}{\mu_A} + \frac{1 - \nu_B}{\mu_B}, \qquad U_0 = \frac{1}{2}\delta_x \qquad [26]$$

where μ, is the shear modulus and v is Poisson's ratio. The mean energy $\bar{E}_y^T$, due to misfit in the y direction, may be written down simply by replacing x by y. This is for the NW-x orientation. The {111}fcc and {110}bcc surfaces are a distance d apart. For the KS orientation μ_x is replaced by μ_x^{KS}, where $\mu_x^{KS}/\mu_x^{NW} \cong 2.5$.[6] This follows from the fact that the shear modulus is proportional to the curvature of the relevant terms in the equations [11], the terms A_1 and A_3 representing the KS and the NW-x orientations, respectively. The average total energy per atom after relaxation is

$$\bar{\varepsilon}_x^T = \bar{E}_x^T c_x c_y/2 \qquad [27]$$

where $c_x c_y/2$ is the area per atom in the reference lattice. We convert to average energy ε per atom for consistency; in the preceding sections, we have expressed energies as averages per atom.

D. *Energy of Stepped Interface vs Planar Interface*

To compare the relative energies of an fcc:bcc interface that contains structural ledges with one that is planar and contains only misfit-compensating dislocations, we need to take the difference between the two:

$$\Delta\bar{\varepsilon} = \bar{\varepsilon}_{step} - \bar{\varepsilon}_{planar} = \bar{\varepsilon}^x_{step} - \bar{\varepsilon}^x_{planar} \qquad [28]$$

In order to assess the relative stability of planar and stepped interfaces, it suffices to consider a single sequence of ledges. In general, there would also be misfit parallel to the ledges, which may be taken up by misfit dislocations, or a near vernier when this misfit is large. Whatever, the simplifying approximation is usually made that the effects of the misfits in the two directions are independent and accordingly additive. An important consequence of this is that the energies associated with the direction normal to the steps would be the same for both the planar and stepped interface and would cancel when we subtract the energies.

The average energies per unit area have been calculated previously[1–6] and are given as follows. For the terrace patch, using the surface stresses, we use Eq. [26]. The second energy for structural ledges is that attributed to TMDs. As noted, misfit gradually builds up in the z direction because $a_z \neq b_z$. For NW-x,

$$\bar{E}_x^z = \frac{\mu_x c_x^2}{2\pi^2(1 - 2\nu)d} F(\bar{\beta}_x) \qquad [29a]$$

$$F(\beta) = 1 + \beta - \sqrt{1 + \beta^2} - \beta \ln [2\beta\sqrt{1 + \beta^2} - 2\beta^2] \qquad [29b]$$

$$\bar{\beta}_x = \frac{\pi d(1 - 2\nu)\lambda}{\bar{p}_x \mu_x} \qquad [29c]$$

For the KS in Eqs. [29a] and [c], μ_x is replaced by μ_x^{KS} and c_x is replaced by c_z. For the misfitting riser interface, the structural ledge contributes directly to the energy of a stepped interface because of misfit, *i.e.*, the misfit along the riser of the ledges of the atomic planes that terminate from either side at the ledge and the misfit due to the difference of the separation of corresponding atomic rows in the two crystals. For the line energy of a ledge, it is the misfitting edges of the atomic planes terminating at a ledge that is approximated by the energy MDs needed to accommodated f_y by assuming that γ is determined by the energy of jogs, which are created by the step in the relevant MDs. This jog energy is equated to the energy per atom length of the MD; *i.e.*,

$$\gamma = \frac{\mu_y c_y^2}{8\pi^2 d} F(\beta_y) \frac{c_x^2 c_y}{l_x} \qquad [30a]$$

$$\beta_y = \frac{2\pi d}{\mu_y p_y} \frac{1}{\lambda} \qquad [30b]$$

for the NW-x ledges and

$$\gamma = \frac{\mu_y^{KS} c^2 c_d}{4\pi^2 d l_x} \qquad [31a]$$

$$c = \frac{2ba}{b+a} = \frac{ra}{r+1}, \qquad \mu_y^{KS} \approx \frac{2}{3}\mu_x^{KS} \qquad [31b]$$

for the KS ledges, where c is a reference lattice parameter along the dense-packed rows and c_d normal thereto.

Finally, for the planar interface, we have[17]

$$\bar{E}_y^P = \frac{\mu_x^{KS} c_d^2}{4\pi^2 d} F(\beta_x) \qquad [32a]$$

$$\beta_x = \frac{2\pi d\lambda}{p_d \mu_x^{KS}}, \qquad p_d = \frac{c_d}{f_{KS}} \qquad [32b]$$

Employing these equations to calculate the relative energies from Eq. [28] shows that, energetically, the ranges of f_x that are geometrically meaningful for stable stepped interfaces are as follows:
for NW-x,

$$-0.1035 < f_x < 0.0589$$

and for KS,

$$-0.0852 < f_{KS} < 0.0281$$

III. CONCLUSIONS

In the consideration of the fcc{111}-bcc{110} interface with rhombic symmetry, it is again demonstrated that all geometrical quantities are interrelated through the misfit ratio r of atomic diameters. The role of row matching in the NW-x and NW-y is shown to be a sufficient condition for energy minimization. Here, the Cartesian coordinate system is employed in which the x and y axes are parallel to the long ⟨110⟩ and the short ⟨001⟩ diagonals, respectively, and the two energy minima are referred to as NW-x and NW-y ledge configurations. For the KS orientation, row matching is again sufficient for energy minimization; here, the close-packed directions of the fcc and bcc rhombuses within the interface align for row matching. This corresponds to the independent findings of Dahmen.[12]

Emphasis was placed on the energetic information to be gained concerning the terraces. The two-phase system was first analyzed by a rigid model approach. This analysis, employing a truncated Fourier series, demonstrated an energetic justification of the optimum terrace widths, which were originally selected on geometrical grounds. The approach further revealed upper and lower bounds of average energies that result in energy-minimizing lattice configurations such as row matching. Elastic relaxation of the terrace patch pair was accomplished by employing the harmonic approximation. This approach revealed further energy gains. A calculation of excess energy based on meaningful values of r (or misfit f_x, Eq. [6]) suggests that stepped interfaces in the NW-x and KS configurations are energetically more favorable as a misfit-accommodating mode than planar interfaces.

ACKNOWLEDGMENTS

One of the authors (GJS) acknowledges the United States National Science Foundation, Division of Materials Research, under Grant No. DMR-9102966. He also acknowledges the hospitality of the Department of Physics, University of South Africa, while he worked on the manuscript.

REFERENCES

1. J.H. van der Merwe: *S. Afr. J. Phys.*, 1985, vol. 9, p. 55.
2. G.J. Shiflet, M.W.H. Braun, and J.H. van der Merwe: *S. Afr. J. Sci.*, 1988, vol. 84, p. 653.
3. J.H. van der Merwe, G.J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, p. 1165-75.
4. J.H. van der Merwe and G.J. Shiflet: *Acta Metall. Mater.*, 1994, vol. 42, p. 1173.
5. G.J. Shiflet and J.H. van der Merwe: *Acta Metall. Mater.*, 1994, vol. 42, p. 1189.
6. J.H. van der Merwe and G.J. Shiflet: *Acta Metall. Mater.*, 1994, vol. 42, p. 1199.
7. M.G. Hall, H.I. Aaronson, and K.R. Kinsman: *Surf. Sci.*, 1972, vol. 31, p. 257.
8. J.H. van der Merwe: *Phil. Mag.*, 1982, vol. A45, pp. 127, 145, and 159.
9. Z. Nishiyama: *Sci. Rep. Tohoku Univ.*, 1934, vol. 23, p. 637.
10. G. Wassermann: *Arch. Eisenhuttwes.*, 1933, vol. 16, p. 647.
11. G. Kurdjumov and G. Sachs: *Z. Phys.*, 1930, vol. 64, p. 325.
12. U. Dahmen: *Acta Metall.*, 1982, vol. 30, p. 63.
13. J.H. van der Merwe: *Proc. R. Soc. London*, 1950, vol. A63, p. 616.
14. J.C. du Plessis and J.H. van der Merwe: *Phil. Mag.*, 1965, vol. 11, p. 43.
15. J.A. Snyman and van der Merwe: *Surf. Sci.*, 1974, vol. 45, pp. 619-639.
16. R.C. Ecob: *J. Microsc.*, 1985, vol. 137, p. 313.
17. P.M. Stoop and J.A. Snyman: *Thin Solid Films*, 1988, vol. 1258, p. 151.
18. J.H. van der Merwe: *J. Appl. Phys.*, 1963, vol. 34, p. 117.
19. J.H. van der Merwe and D.L. Tonsing: University of South Africa, Pretoria, unpublished research, 1994.
20. J. Howe and D.A. Smith: *Acta. Metall. Mater.*, 1992, vol. 40, p. 2343.
21. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. London*, 1949, vol. 198, pp. 205 and 117.
22. J.H. van der Merwe: in *Treatise on Materials Science and Technology*, H. Herman, ed., Academic Press, New York, NY, 1973, vol. 2, p. 1.

Intrinsic Ledges at Interphase Boundaries and the Crystallography of Precipitate Plates

YIWEN MOU

The structure of intrinsic ledges at interphase boundaries has been interpreted with extended O-lattice/DSC-lattice approaches. The distribution of structural ledges can be predicted if the spacing difference between parallel matrix and product planes is treated as a measure of the relaxed coincidence condition. A small rotation away from the low-index planar parallelism introduces a series of interfacial dislocations that cancels the spacing difference, resulting in a lattice invariant line. Misfit-compensating ledges at bcc:hcp interfaces are produced as a ledged interface intersects additional O-points that are recognized with the incorporation of previously omitted bcc atom positions into the O-lattice construction. Energetic consideration suggests that structural interfacial energy may decrease when a flat interface becomes ledged with misfit-compensating ledges. Burgers vectors associated with structural ledges and misfit-compensating ledges are displacement shift complete (DSC) lattice vectors. Precipitate and martensite crystallography may both include a lattice invariant line, but they are involved in different interphase boundary characteristics. Assumptions and implications in precipitate and martensite crystallography are discussed in the framework of the O-lattice theory and phenomenological theory of martensite crystallography.

I. INTRODUCTION

STRUCTURAL ledges[1,2,3] and misfit-compensating ledges,[3–6] just like well-known misfit dislocations,[7,8,9] are intrinsic parts of partially coherent interphase boundaries, as they are related with somewhat low interfacial energy. In contrast, growth ledges[10,11] are extrinsic to interphase boundaries, as their existence would increase the interfacial energy.[12] This non-equilibrium character permits growth ledges to add atoms to the precipitate without the need to surmount a high-energy barrier.[12] It is now widely accepted that in the nucleation process, nuclei of a precipitate phase take up a lattice orientation relationship that allows them to be enclosed largely by low-energy coherent interfaces.[13] Such interfaces are usually present during growth as partially coherent boundaries that can migrate only through the agency of the ledge mechanism[13] when the crystal structures of the product and matrix phases are significantly different.[12] The structure of partially coherent interphase boundaries is thus an important aspect in the knowledge of solid-solid phase transformations.

Hall *et al.*[1] first suggested that interfacial matching can be improved by the periodic insertion of structural ledges as a result of the difference in the stacking sequence in the face-centered cubic (fcc) and body-centered cubic (bcc) structures. Rigsbee and Aaronson[2] used a graphical model to predict the distribution of structural ledges at fcc:bcc interfaces. Shiflet *et al.*,[14] van der Merwe *et al.*,[15] and Shiflet and van der Merwe[16] explained the geometry of structural ledges with the concept of pattern advance and analyzed their energetics by rigid-body and harmonic approaches. With a computer-aided graphic technique, Furuhara and Aaronson[17] found that a set of structural ledges can effectively replace misfit dislocations at Burgers-related[18] and Potter-related[19] bcc:hcp interfaces. Structural ledges were observed at a fcc:bcc broad face of a ferrite sideplate in an Fe-C-Si alloy[20] by transmission electron microscopy (TEM) and at a Burgers-related bcc:hcp interface in a Ti-Cr alloy[3] by high-resolution TEM.

Structural ledges are recognized as interfacial defects in terms of their misfit-compensating ability, usually expressed as Burgers vectors.[21,22] Although a network of misfit dislocations can in general accommodate misfit across a partially coherent interface,[7–9,17,23] the replacement of misfit dislocations by structural ledges may in some cases lower the structural interfacial energy.[15] The analysis of van der Merwe *et al.*[15] indicates that for small misfits and large Burgers vectors, ledged interfaces possess lower interfacial energies than do the corresponding misfit dislocation interfaces.

Misfit-compensating ledges were first found by Furuhara and Aaronson[4,5] and Furuhara *et al.*[3] at $\alpha:\beta$ (bcc:hcp) interfaces associated with proeutectoid α plates and grain-boundary allotriomorphs in a hypoeutectoid Ti-Cr alloy. These ledges are associated with Burgers vectors parallel to the terrace planes and can compensate misfit by means of extra half-atomic planes perpendicular to the terraces.[12,22] These Burgers vectors were usually $(1/2)[0001]_\alpha$, determined by comparing the measured spacing between parallel ledges with the spacing modeled for $[0001]_\alpha$ misfit dislocations.[17] Recently, Mou and Aaronson[6] observed misfit-compensating ledges at $\beta:\zeta$ (bcc:hcp) interfaces formed during a massive transformation in a Ag-26 at. pct Al alloy.

Structural ledges and misfit-compensating ledges are different mainly in their modes of misfit compensating.[5,12,22,24] Structural ledges can compensate misfit

YIWEN MOU, Research Associate, is with the Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA 22903.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

only within the terrace plane,[1,5,17,24] while misfit-compensating ledges can compensate misfit both within and perpendicular to the terrace plane.[17,24] However, this difference is probably not as essential as was previously proposed,[17,24] especially when the associated Burgers vectors are described with the displacement shift complete (DSC) lattice.[25,26,27] Nonetheless, these ledges may still be treated separately with emphasis on their characteristics in interfacial misfit compensation.

The concept of structural ledges is closely related with that of the lattice invariant line,[21,22,28,29] as these ledges permit the matching across an already well-matched interface to be improved by the more frequent repetition of coherent areas.[1,2] These concepts have long been recognized as approaches to the interfacial structure and crystallography of precipitate plates.[2,17,30–32] Dahmen[21] suggested that structural ledges stepping down along the lattice invariant line can eliminate the periodic insertion of edge dislocations in the riser planes. This suggestion was confirmed by Furuhara and Aaronson[17] for partially coherent bcc:hcp interfaces by comparing graphical and O-lattice results with invariant line predictions. Howe and Smith[33] proposed that the structural ledge and invariant line models are significantly different in their fundamental approaches and that the two models can lead to the similar or different predictions for habit plane, orientation relationship, and dislocation structure,[34] depending on the type of transformation strain. However, Kato *et al.*[35] recently discussed the strain type that led to different habit planes in Howe and Smith's analysis and concluded that there are no true differences in habit plane orientation predicted with the two models. Aaronson *et al.*[22] proposed that the kinetics of growth ledge generation at the interphase boundary orientation represented by the invariant line direction are a pronounced minimum because of the zero misfit along the invariant line. This proposal provides an important linkage between the invariant line concept and diffusional growth by means of the ledge mechanism.[13]

The O-lattice[25,36] and DSC-lattice[25–27,36] approaches, or so-called "fit-misfit" models,[27] have long been applied to describing interfacial ledges.[27,37–41] Mou and Aaronson[24] recently extended the O-lattice approach[25] for developing a general analytical approach to predicting structural ledge distribution. They suggested[24] that misfit-compensating ledges result from intersection of a ledged interface with additional O-points, which are best-matching positions but not recognized by previous investigations. The present article is aimed mainly at further developing these concepts and approaches.

In the present work, structural ledges and misfit-compensating ledges are analyzed with the O-lattice theory[25,36] and related concepts, such as the surface dislocation[42] and DSC lattice.[25,26] The dislocation description and energetics of these ledges are then considered with emphasis on low-energy partially coherent interfacial structure. Based on its connection to interfacial structure, the crystallography of platelike precipitates is finally discussed in comparison with martensite crystallography described with the phenomenological theory.[43,44]

II. STRUCTURAL LEDGES

A. *Geometric Distribution*

A new approach somewhat different from that in Reference 24 will now be introduced to facilitate the misfit and dislocation description of structural ledges. As shown in Figure 1, $(h_1k_1l_1)$ and $(h_2k_2l_2)$ are two sets of crystallographic planes, usually but not necessarily of the low-index type, in lattices (1) and (2), respectively. Suppose that an orientation relationship allows the two sets of planes to be parallel to the x-y coordinate plane. Let $\mathbf{b}_z$ be the separation between the pair of $(h_1k_1l_1)$ and $(h_2k_2l_2)$ planes actually shown in Figure 1; $\mathbf{b}_z$ is thus a vector parallel to the z axis and equal in magnitude to the difference between the $(h_1k_1l_1)$ and $(h_2k_2l_2)$ spacings. If the position vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, of lattices (1) and (2), respectively, satisfy the conditions

$$\mathbf{x}^{(2)} - \mathbf{x}^{(1)} = \mathbf{b}_z \qquad [1]$$

and

$$\mathbf{x}^{(2)} = \mathbf{A}\mathbf{x}^{(1)} \qquad [2]$$

their endpoints P and Q will be "equivalent" positions (from Eq. [2]) and "coincident" with respect to the x and y dimensions, because $\mathbf{b}_z$ has zero components along both the x and y axes (from Eq. [1]). The transformation matrix $\mathbf{A}$ may be expressed as[25]

$$\mathbf{A} = \mathbf{R}\mathbf{S}^{(2)}(\mathbf{S}^{(1)})^{-1} \qquad [3]$$

where $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ are the structure matrices[25] of lattices (1) and (2), respectively, and $\mathbf{R}$ is the rotation matrix defined by the orientation relationship.[24] Substituting Eq. [2] into Eq. [1] gives

$$(\mathbf{I} - \mathbf{A}^{-1})\mathbf{x}^{(2)} = \mathbf{b}_z \qquad [4]$$

where $\mathbf{I}$ is the identity matrix.

Equation [4] has almost the same form as the O-lattice basic equation[25] except the vector $\mathbf{b}_z$ is not a lattice vector (hence without the superscript L). As a result, $\mathbf{x}^{(2)}$ is not an O-lattice vector since the endpoints of $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(1)}$ have different internal coordinates in the z direction. It is obvious that, however, the two lattices are somewhat

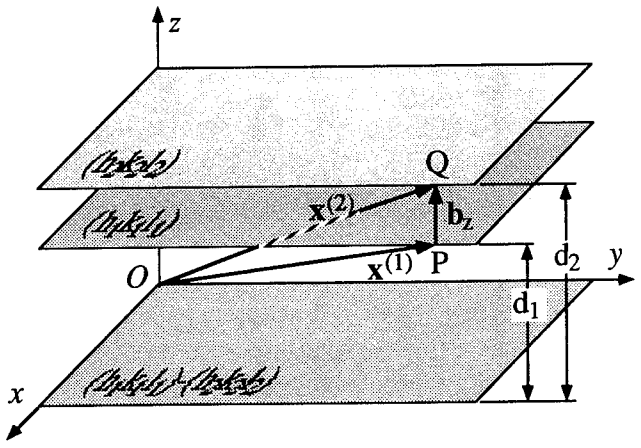


Fig. 1—Lattice matching across consecutive terraces of structural ledges, analyzed with a tolerance or relaxation vector normal to the terraces.

best matched at the endpoint of $\mathbf{x}^{(2)}$ if b_z (the magnitude of $\mathbf{b}_z$) is small enough. Indeed, $\mathbf{x}^{(2)}$ represents an O-point[25] when b_z approaches 0.

On the other hand, Eq. [4] is apparently equivalent to the basic equation of the surface dislocation model[42] if lattice (1) is taken as the reference lattice as in the previous derivation of Eq. [4]. However, instead of having arbitrary components in the surface dislocation model, $\mathbf{b}_z$ in Eq. [4] has zero components in the two dimensions defining the $(h_1k_1l_1)$ or $(h_2k_2l_2)$ plane. The mathematical formulation here is thus equivalent to Rigsbee and Aaronson's graphical approach:[2] the two planes separated by a small distance are superimposed for considering two-dimensional matching.

The geometric distribution of structural ledges on Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces, predicted with Eq. [4], is shown in Figure 2. In this example, $(h_1k_1l_1) = (2\bar{1}\bar{1})_{bcc}$, $(h_2k_2l_2) = (1\bar{1}00)_{hcp}$, and $b_z = |d_1 - d_2|$, where d_1 and d_2 are the $(2\bar{1}\bar{1})_{bcc}$ and $(1\bar{1}00)_{hcp}$ interplanar spacings, respectively. The lattice parameter ratio a_{bcc}/a_{hcp} is chosen as 1.11 for clarity as this ratio gives a relatively large misfit across the terraces and allows the $\mathbf{x}^{(2)}$ vector to end at an atom position. The matrices $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ are formed with the dimension vectors $\mathbf{s}_i^{(1)}$ and $\mathbf{s}_i^{(2)}$,[24] respectively, as shown in Figure 3. The rotation matrix $\mathbf{R}$ is taken as $\mathbf{I}$ for the Burgers orientation relationship treated here. As the two lattices meet at the $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ interface, only one atom of either structure would occupy each pair of atom positions (open and solid circles in Figure 2). This may be called a relaxation process. Suppose relaxation always occurs at atom pairs with relatively small misfit. The interface would become ledged with a structural ledge located midway between the best-matching positions O and P, as outlined by thick lines in Figure 2.

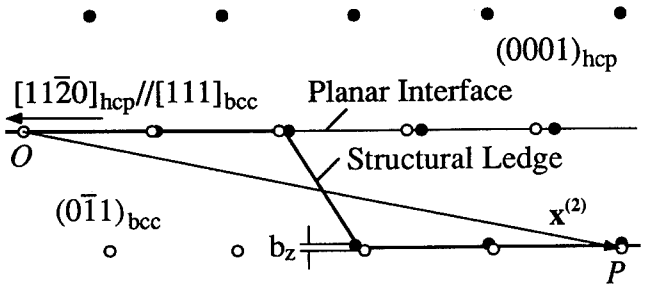


Fig. 2—Distribution of structural ledges on Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces, predicted with Eq. [4] and a lattice parameter ratio of 1.11.

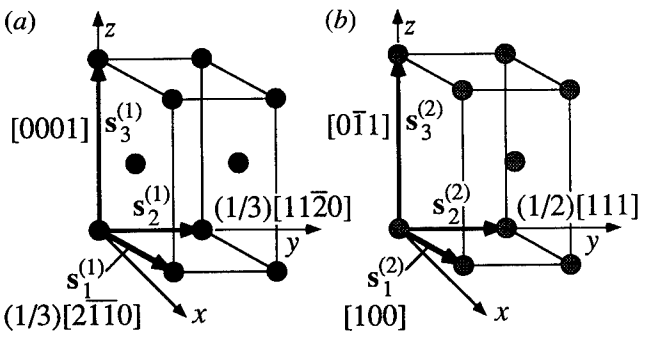
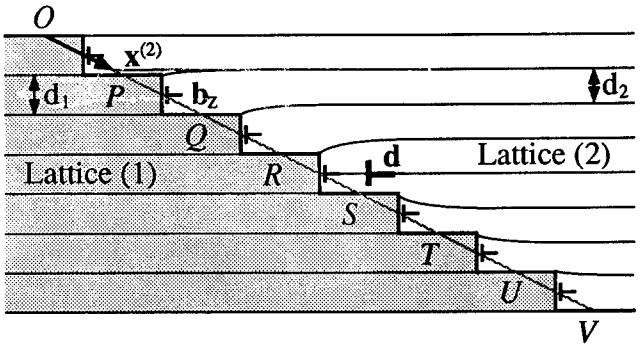


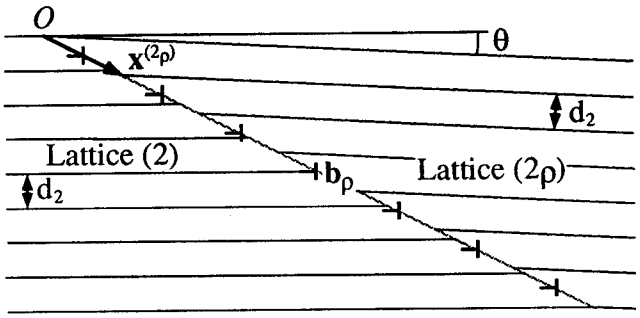
Fig. 3—(*a*) The bcc and (*b*) hcp unit cells chosen for the O-lattice analyses. Notice the base-central positions in the bcc unit cell.

B. *Misfit Description and the Invariant Line*

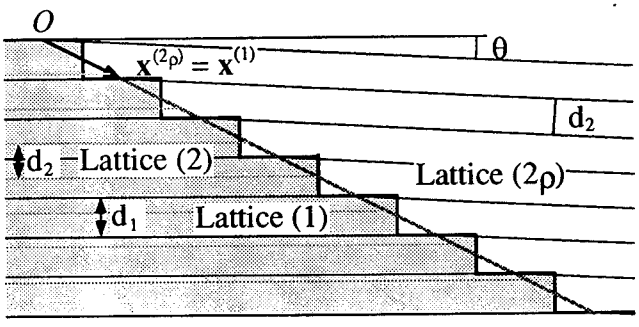
Macroscopically, the ledged interface in Figure 2 is recognized as an inclined plane defined by the $\mathbf{x}^{(2)}$ vector and $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ direction, as shown in Figure 4(a). The misfit between the two lattices, accumulated from O to P, is $\mathbf{b}_z$ according to Eq. [4]. Since Eq. [4] is linear, the two lattices will build up a misfit of $2\mathbf{b}_z$ from O to Q with $OQ = 2\mathbf{x}^{(2)}$, $3\mathbf{b}_z$ from O to R with $OR = 3\mathbf{x}^{(2)}$, *etc.* In other words, a series of coherency dislocations[37] each with the Burgers vector $\mathbf{b}_z$ will be distributed with a spacing of $|\mathbf{x}^{(2)}|$ at each ledge riser.



(*a*)



(*b*)



(*c*)

Fig. 4—(*a*) Misfit normal to structural-ledge terraces, distributed as coherency dislocations along an inclined boundary; (*b*) a set of dislocations created with a small rotation of part of lattice (2) with respect to the rest; and (*c*) misfit-free coherent boundary produced by cancellation of the dislocations in (a) and (b).

In view of the surface dislocation concept,[42] a surface dislocation with a Burgers vector density of $\mathbf{b}_z/|\mathbf{x}^{(2)}|$ will be present in the inclined interface. As misfit builds up to the extent nearest to the $(h_1k_1l_1)$ spacing (lattice (1) taken as the reference), an edge dislocation or an extra half-atomic plane is required to insert into the more densely packed lattice.[1] These edge dislocations, with a Burgers vector of $\mathbf{d}$ ($|\mathbf{d}| = d_1$) and a spacing of $(d_1/b_z)\mathbf{x}^{(2)}$, are so-called anticoherency dislocations.[37]

In most practical cases,[1,2,17,24] $\mathbf{b}_z$ is very small in magnitude; $\mathbf{x}^{(2)}$ (or $\mathbf{x}^{(1)}$) is an approximate lattice invariant line (Eq. [1]). When $\mathbf{b}_z$ approaches zero, $\mathbf{x}^{(2)}$ (or $\mathbf{x}^{(1)}$) becomes a lattice invariant line. In the example shown in Figure 2, therefore, the two lattices are matched better along the $\mathbf{x}^{(2)}$ direction than along $[11\bar{2}0]_{hcp}//[111]_{bcc}$. This argument provides a simple energetic basis for the existence of structural ledges.

An additional rotation of lattice (2) away from the low-index orientation relationship ($\mathbf{R}$ in Eq. [3]), designated as the matrix ρ, may improve and at best cancel the misfit by introducing a set of dislocations with an equal and opposite Burgers vector, shown in Figure 4(b). A position vector $\mathbf{x}^{(2\rho)}$ in lattice (2ρ) (rotated lattice (2)) is related to $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(1)}$ as follows:

$$\mathbf{x}^{(2\rho)} = \rho\mathbf{x}^{(2)} = \rho\mathbf{A}\mathbf{x}^{(1)} \qquad [5]$$

The misfit between lattices (2ρ) and (2), accumulated along the boundary segment $\mathbf{x}^{(2)}$, can be expressed as a Burgers vector $\mathbf{b}_\rho$:

$$\mathbf{x}^{(2\rho)} - \mathbf{x}^{(2)} = (\rho - \mathbf{I})\mathbf{x}^{(2)} = \mathbf{b}_\rho \qquad [6]$$

On the other hand, if $\mathbf{b}$ is the Burgers vector or misfit along the same segment $\mathbf{x}^{(2)}$ but between lattices (2) and (1), Eq. [4] will have the form

$$(\mathbf{I} - \mathbf{A}^{-1})\mathbf{x}^{(2)} = \mathbf{b} \qquad [4a]$$

Suppose that $\mathbf{b}_\rho$ cancels $\mathbf{b}$ in Eq. [4a], *i.e.*, $\mathbf{b}_\rho = -\mathbf{b}$. The addition of Eqs. [4a] and [6] gives

$$(\rho - \mathbf{A}^{-1})\mathbf{x}^{(2)} = \mathbf{0} \qquad [7]$$

where $\mathbf{0} = [0\ 0\ 0]^T$ (T for transpose). Equation [7] may also be written in form of $\mathbf{x}^{(1)}$:

$$(\rho\mathbf{A} - \mathbf{I})\mathbf{x}^{(1)} = \mathbf{0} \qquad [8]$$

or

$$\mathbf{x}^{(2\rho)} = \mathbf{x}^{(1)} \qquad [9]$$

Equation [9] indicates that $\mathbf{x}^{(1)}$ is a lattice invariant line[21,23,25,28] with respect to the transformation $\rho\mathbf{A}$ or the adjusted orientation relationship $\rho\mathbf{R}$ between lattices (1) and (2). The rotation matrix ρ can be determined from[23,30]

$$|\rho\mathbf{A} - \mathbf{I}| = 0 \qquad [10]$$

Since a pure rotational transformation is isometric (no changes in vector length), notice from Eqs. [5] and [9] that

$$|\mathbf{x}^{(2)}| = |\mathbf{x}^{(2\rho)}| = |\mathbf{x}^{(1)}| \qquad [11]$$

This result shows that $\mathbf{x}^{(1)}$ is an unstretched line[29] with respect to the transformation $\mathbf{A}$ or the initial low-index orientation relationship $\mathbf{R}$. Interfacial dislocations resulting from the difference between the final ($\mathbf{x}^{(2)}$) and initial ($\mathbf{x}^{(1)}$) positions of an unstretched vector can be exactly canceled by a set of dislocations produced by a rigid-body rotation of lattice (2). This formulation addresses the concepts of lattice matching and dislocation configuration that are not explicit in the invariant line model,[21,28–32] suggesting that structural ledges are generally related to the lattice invariant line. Figure 4(c) shows that with Burgers vector densities $\mathbf{b}/|\mathbf{x}^{(2)}|$ and $\mathbf{b}_\rho/|\mathbf{x}^{(2)}|$ canceling each other, $\mathbf{x}^{(2)}$ becomes a misfit-free coherent boundary.

In general, the invariant line interface $\mathbf{x}_I^{(1)}$ secured as such differs from both $\mathbf{x}_S^{(1)}$ and $\mathbf{x}_S^{(2)}$, which are the structural ledge interfaces calculated from Eq. [4]. However, as $\mathbf{b}_z$ in Eq. [4] is assumed to have a very small magnitude in the structural ledge model,[1,2] either $\mathbf{x}_S^{(1)}$ or $\mathbf{x}_S^{(2)}$ is a very good approximation to $\mathbf{x}_I^{(1)}$. The rotation ρ, indicating a slight difference between the predicted orientation relationships, approaches $\mathbf{I}$ when $|\mathbf{b}_z|$ approaches 0. In most practical cases, ρ is a rotation of less than 0.5 deg.[23,28,29] The structural ledge and invariant line models hence predict very close results that become identical when $|\mathbf{b}_z|$ approaches 0.

C. *DSC Dislocation Description*

The dislocation characteristic of structural ledges can be better understood with the DSC model.[27] Figure 5(a) shows the relative atom positions of the bcc and hexagonal close-packed (hcp) phases in the $(0001)_{hcp}//(0\bar{1}1)_{bcc}$ planes aligned as $[11\bar{2}0]_{hcp}//[111]_{bcc}$. A diagonal transformation $\mathbf{D}$ can correlate the atom positions with emphasis on matching between parallel $(1\bar{1}00)_{hcp}$ and $(2\bar{1}\bar{1})_{bcc}$ planes:

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad [12]$$

Where λ_1 and λ_2 are the length changes involved in making $(1\bar{1}00)_{hcp}$ coincident with $(2\bar{1}\bar{1})_{bcc}$ and the hcp atom positions along $[11\bar{2}0]_{hcp}$ coincident with the bcc atom positions along $[111]_{bcc}$, respectively. In general, λ_1 and λ_2 are referred to relatively close-packed planes and directions of interest. As a convention, these parameters are evaluated as the ratios of corresponding lengths in lattice (2) to those in lattice (1).

If lattice (2) is stressed to fit with lattice (1), the stressed lattice (2), or (2′), will be related to lattice (2) as follows:

$$\mathbf{x}^{(2')} = \mathbf{D}^{-1}\,\mathbf{x}^{(2)} \qquad [13]$$

As Figure 5(b) shows, lattices (2′) and (1) have coincident sites other than the origin. Figure 5(c) indicates the coincident site lattice that consists of all the coincident sites (open circles) of lattices (1) and (2′). A DSC lattice composed of all the difference vectors of lattices (1) and (2′) has a unit cell (shaded area) defined by the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$. In the present example, $\mathbf{b}_1 = (1/3)[2\bar{1}\bar{1}]_{bcc}$ and $\mathbf{b}_2 = (1/12)[111]_{bcc}$.

Since $\mathbf{b}_2$ is 1/6 the lattice translation vector

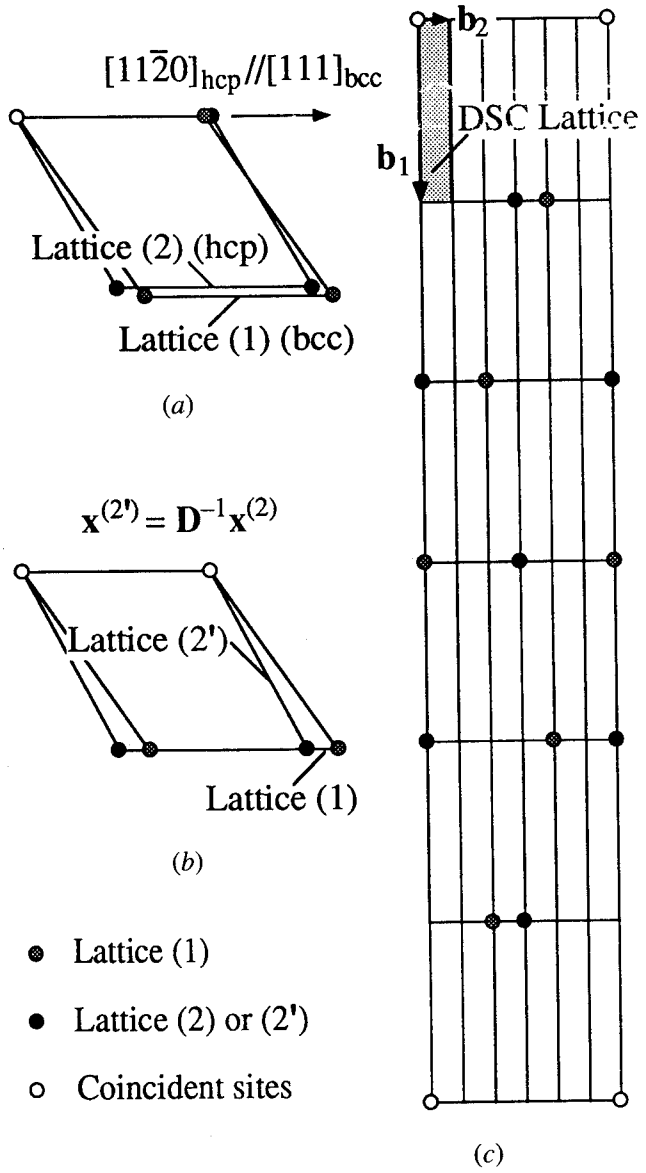


Fig. 5—(*a*) Atom positions of the bcc and hcp structures in the $(0001)_{hcp}//(0\bar{1}1)_{bcc}$ planes; (*b*) coincident sites between lattice (1) and stressed lattice (2) or (2′); and (*c*) DSC lattice composed of all the difference vectors of lattices (1) and (2′).

$(1/2)[111]_{bcc}$, it is recognized as a local perturbation accumulated every 1/6 the spacing between misfit dislocations distributed along the $[11\bar{2}0]_{hcp}//[111]_{bcc}$ direction.[24] On the other hand, the stacking displacements of lattices (1) and (2′) differ by $-\mathbf{b}_2$ (Figure 5(c)). If the interface steps down an atom layer at a position where a local perturbation with the Burgers vector $\mathbf{b}_2$ builds up, the perturbation will be canceled by the stacking displacement difference $-\mathbf{b}_2$. As a result, a somewhat best-matching position between lattices (1) and (2) is realized in the next layer, and a structural ledge forms midway between the two adjacent best-matching positions.

Figure 6 shows schematically the matching between the bcc and hcp phases across consecutive terraces of structural ledges, using lattice parameters selected for the example in Figure 2. It is clearly seen that there is

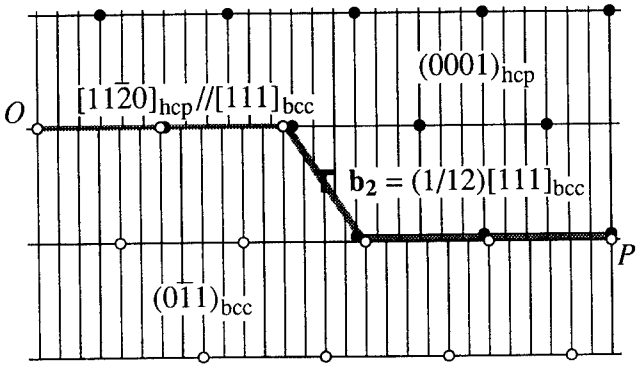


Fig. 6—Matching between the bcc and hcp phases across consecutive terraces of structural ledges, expressed as DSC Burgers vectors.

a DSC or interphase boundary dislocation[37] with the Burgers vector $\mathbf{b}_2$ between the best-matching positions *O* and *P*. The DSC dislocations associated with structural ledges may be interpreted as extra half-atomic planes. As for the present example, these extra half planes are $(12, 12, 12)_{bcc}$. This is a revised view of the previous description that there are no extra half-atomic planes associated with structural ledge terraces.[24]

The Burgers vector $\mathbf{b}_1$, equivalent to the Burgers vector **d** discussed in Section II–B, reflects accumulated differences between the $(0001)_{hcp}$ and $(0\bar{1}1)_{bcc}$ spacings, or the $\mathbf{b}_z$ vector discussed in Sections II–B and II–C. It is clear now, however, that dislocations with the Burgers vector $\mathbf{b}_1$ are also of the DSC or interphase boundary type. These dislocations cannot be recognized with a regular O-lattice analysis that includes only lattice translation vectors.[25] A $\mathbf{b}_1$ vector $[(1/3)[2\bar{1}\bar{1}]_{bcc}]$ corresponds to an extra half $(6\bar{3}\bar{3})_{bcc}$ plane, present only at bcc:hcp interphase boundaries.

It is seen from the DSC lattice approach that $\mathbf{b}_2 = 0$ if there are no stacking sequence changes in the two lattices or if the stacking displacements are equivalent. Under these circumstances, a structural ledge would not be able to stabilize the local perturbation and improve the matching between the two lattices.

The rectangular coincident site lattice in Figure 5(c) indicates that the matrix **D** is the lattice correspondence used in the above DSC lattice analysis. The use of two lattice correspondences is fairly reasonable since the correspondence **A** used before is aimed at lattice matching along arbitrary planes and directions, while **D** emphasizes matching along certain parallel low-index planes and along their normal direction.

III. MISFIT-COMPENSATING LEDGES

A. *Additional Best-Matching Positions*

When a structure base is involved in at least one of the two structures in an O-lattice construction, some atom positions must be omitted in specifying a lattice correspondence.[45] In the hcp structure, for instance, a two-atom base is associated with each lattice point in the simple hexagonal lattice. One of the two atoms in an hcp base was traditionally ignored because of the difficulty in expressing *both* the atoms in O-lattice constructions.

In bcc/hcp O-lattice analyses,[17,46–48] accordingly, half the atom positions of the bcc structure had to be omitted in specifying a unit cell corresponding to the hcp unit cell.

It was recently recognized,[24] however, that the atom positions omitted in many previous O-lattice analyses may be important for the interpretation of diversified interphase boundary structures. Notice that the phenomenological theory of martensite transformations[43,44,45] ignores these atom positions with the argument that the omitted atoms in the matrix and product phases are related by shuffle movements.[45] This is a reasonable assumption for a theory like the martensite theory based on optimal strain accommodation. However, when interfacial matching or atomic bonding across an interphase boundary is the main concern as in O-lattice problems, an approach will be physically sounder if it can consider these atom positions.

Figure 7 shows schematically that best-matching positions or O-points constructed from square lattices (1) (Figure 7(a)) and (2) (Figure 7(b)) at the identity orientation relationship form a square O-lattice (Figure 7(d)). If one of the square lattices, say lattice (1), includes two-atom bases (Figure 7(c)), the resultant O-lattice will remain square but have two-point bases, as shown in Figure 7(e). It is obvious that omission of the atoms inside the unit cells of lattice (1) makes it impossible to recognize these additional best-matching positions.

A special case is that the atoms inside the unit cells of lattice (1) are centered with respect to the unit cells. Lattice (1) is now indeed a simple square lattice with a smaller unit cell. However, if the smaller unit cell is chosen for O-lattice analyses, it is impossible to make its corner atom positions be nearest neighbors to those of lattice (2).[25] The best approach for this situation would still choose the larger unit cell and treat the central atom positions as if they were involved in a base structure.

B. *Search for Additional O Points*

Similarly to the case of structural ledges treated in Section II–A, the following is a generalized and simplified version of Mou and Aaronson's approach. As shown in Figure 8, the atom position A in a two-atom base of structure (1) is set coincident with an atom position in simple lattice (2). Structure (1) can then be divided into two subsets (1A) and (1B), composed of atom positions equivalent to A and B, respectively. A lattice correspondence $\mathbf{A}$ is specified between subset (1A) and lattice (2) with the coincident atom position being the origin O. Notice that by this specification subset (1A) has been taken as the reference lattice.[25] A best-matching position between subset (1B) and lattice (2) will be recognized at the endpoint of the position vector $\mathbf{x}^{(2)}$ in lattice (2) if the difference between $\mathbf{x}^{(2)}$ and the position vector $\mathbf{x}^{(1)}$ in structure (1) equals $\mathbf{b}_B$, the vector starting from $O(A)$ and ending at B:

$$\mathbf{x}^{(2)} - \mathbf{x}^{(1)} = \mathbf{b}_B \qquad [1a]$$

Substituting Eq. [2] gives

$$(\mathbf{I} - \mathbf{A}^{-1})\mathbf{x}_B^{(0)} = \mathbf{b}_B \qquad [14]$$

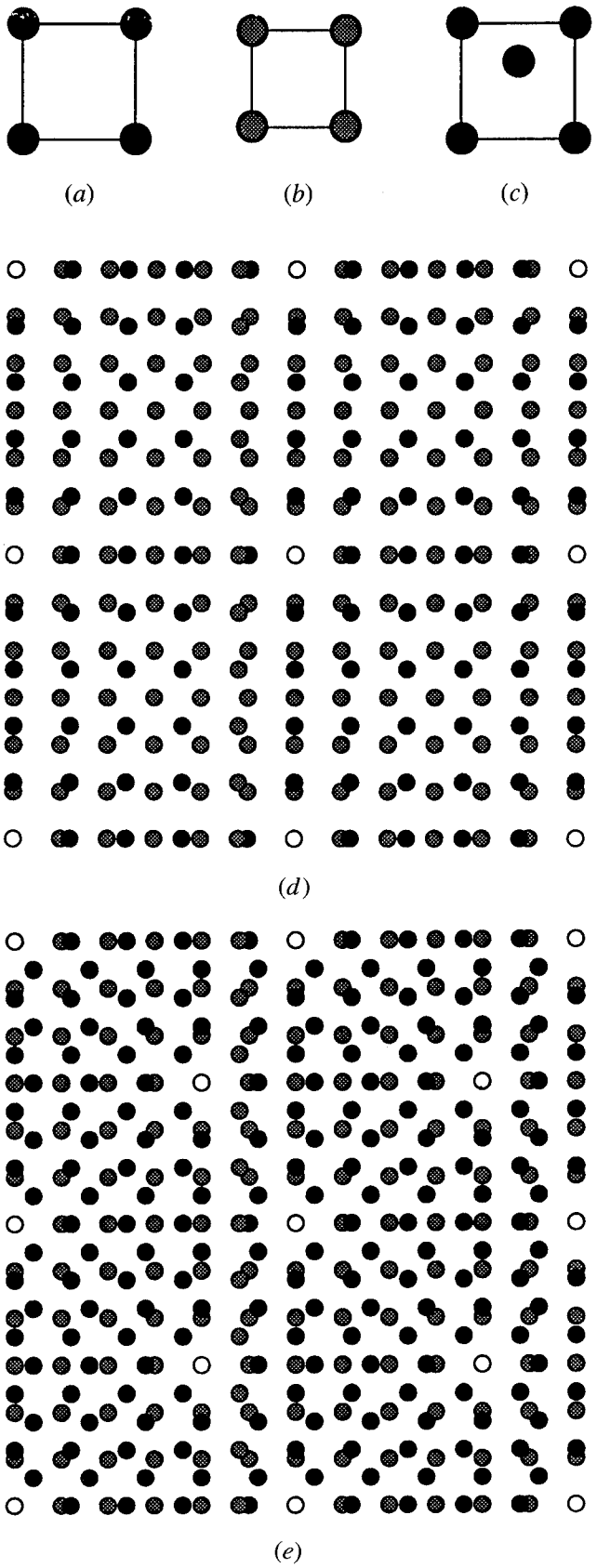


Fig. 7—O-points of two square lattices: (*a*) simple structure (1); (*b*) lattice (2); (*c*) structure (2) with two-atom bases; (*d*) O-points of the structures in (a) and (b); and (*e*) O-points of the structures in (c) and (a).

In the bcc/hcp case, the transformation matrix **A** can be evaluated from the unit cells given in Figure 3 with a specified orientation relationship.[17,24]

Equation [14] is different from Eq. [4] in that the endpoint of $\mathbf{x}_B^{(0)}$ (equal to $\mathbf{x}^{(2)}$) is an O point. Although $\mathbf{b}_B$ is generally not a lattice translation vector, the endpoint $\mathbf{x}_B^{(0)}$ is a coincident position between subset (1B) and lattice (2) because the translation relationship expressed by $\mathbf{b}_B$ exists between subsets (1A) and (1B) of structure (1). Notice that as for best matching, the O-points secured as such have no difference from the O-points between subset (1A) and lattice (2), which may be obtained with a regular O-lattice analysis.

In three dimensions, $\mathbf{b}_B$ can be written as follows (Figure 8(a)).

$$\mathbf{b}_B = k_1\mathbf{s}_1^{(1)} + k_2\mathbf{s}_2^{(1)} + k_3\mathbf{s}_3^{(1)} \qquad [15]$$

where the $\mathbf{s}_i^{(1)}$'s are the basis vectors of subset (1A), and the k_i's ($0 \le k_i < 1$) are the internal coordinates of B in the unit cell of subset (1A). Specially, when two or all three k_i's are $1/2$, B is a face- or body-central lattice position. Substituting Eq. [15] into Eq. [14] yields

$$\mathbf{x}_B^{(0)} = k_1\mathbf{x}_1^{(0)} + k_2\mathbf{x}_2^{(0)} + k_3\mathbf{x}_3^{(0)} \qquad [16]$$

where the $\mathbf{x}_i^{(0)}$'s are the basis vectors of the O-lattice of subset (1A) and lattice (2): *i.e.*

$$(\mathbf{I} - \mathbf{A}^{-1})\mathbf{x}_i^{(0)} = \mathbf{s}_i^{(1)} \qquad [17]$$

Equation [16] indicates that the O-point represented by $\mathbf{x}_B^{(0)}$ is located "inside" the O-lattice unit cell regularly constructed from subset (1A) and lattice (2) with the same internal coordinates as the atom position B is located "inside" the (1A) unit cell, as shown in Figure 8(c). When lattice (2) is also involved in a two-atom base as in the bcc/hcp O-lattice analysis,[24] the inside atom positions in its unit cell may first be expressed as positions in subset (1A) through the transformation **A**. This is an important result that makes the search of additional O-points very intuitive and convenient and has potential applications to the most general case in which both crystal structures involve multiple-atom bases. However, a complex O-point distribution such obtained might be unnecessary or difficult to interpret in many cases.

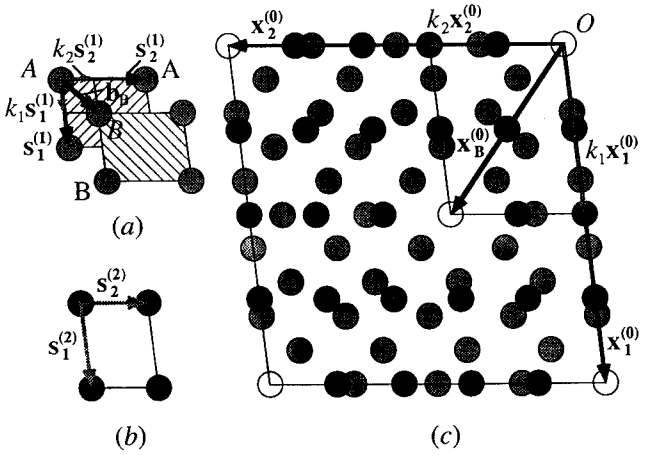


Fig. 8—Search for additional O points: (*a*) structure (1) with two-atom bases; (*b*) lattice (2); and (*c*) additional O points located "inside" the O-lattice unit cell regularly constructed.

C. *Geometric Distribution of Misfit-Compensating Ledges*

Figure 9 shows the O-lattice obtained from the bcc and hcp unit cells shown in Figure 3 with the Burgers orientation relationship or $\mathbf{R} = \mathbf{I}$. The lattice parameters are taken as $a_\beta = 0.324$ nm, $a_\zeta = 0.2865$ nm, and $c_\zeta = 0.4653$ nm, appropriate to the β-to-ζ_m massive transformation in a Ag-26 at. pct Al alloy.[6] Note that the inside atom position B in the bcc unit cell shown in Figure 3 has internal coordinates of $(0.5, 0, 0.5)$ with respect to the crystal axes $\mathbf{s}_i^{(1)}$. An additional O point is thus located inside the O-lattice unit cell, having the same internal coordinates, $(0.5, 0, 0.5)$, but with respect to the O-lattice axes $\mathbf{x}_i^{(0)}$, as shown in Figure 9.

The geometry of misfit-compensation ledges can be predicted according to the O-point distribution (Figure 10). The additional O-points found with the above approach are designated as B_{ij} and others as A_{ij}. A ledged interface stepping down to intersect alternate A_{ij} and B_{ij} will include more best-matching positions than

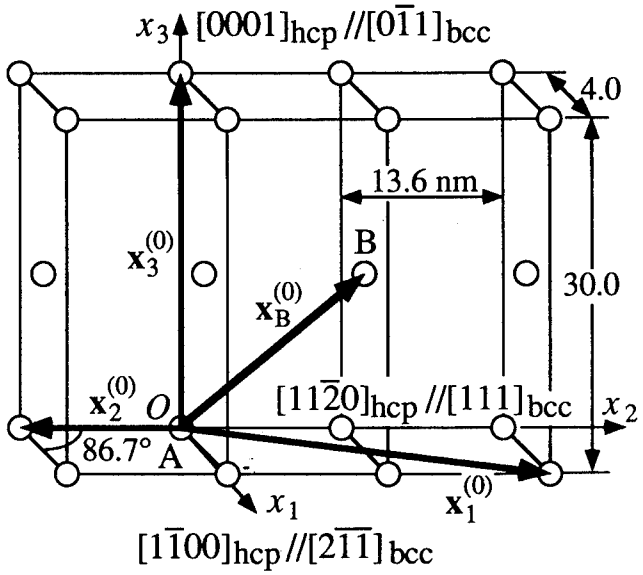


Fig. 9—O-lattice of the bcc and hcp structures shown in Fig. 3 with the Burgers orientation relationship.

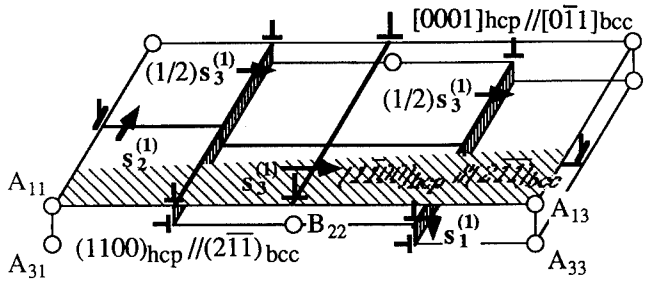


Fig. 10—Geometric distribution and associated Burgers vectors of misfit-compensating ledges with Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces.

does the flat interface crossing only A_{ij}. For the purpose of misfit accommodation, a misfit dislocation is required to be located midway between every two adjacent A_{ij}'s in the flat interface, while two ledges are needed between every adjacent A_{ij}'s in the ledged interface. This model also explains the findings[3–6] that the spacing between misfit-compensating ledges at a ledged interface is about half the spacing between misfit dislocations at the corresponding flat interface.

Figure 10 shows the distribution of misfit-compensating ledges with Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces.[24] For $\beta:\zeta_m$ interfaces associated with the Burgers orientation relationship in a Ag-26 at. pct Al alloy, the observed distributions of misfit-compensating ledges were found in good agreement with the predicted geometry including mainly interledge spacing and Burgers vector as well as ledge height in some cases.[6,24]

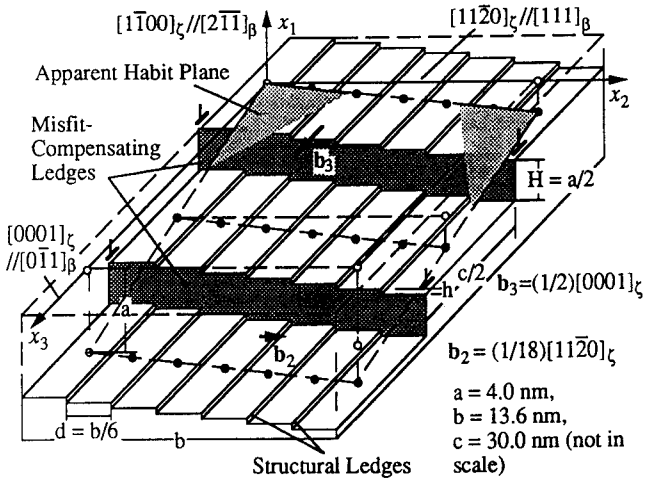


Fig. 11—Predicted distribution of structural ledges and misfit-compensating ledges with $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces at a $\beta:\zeta_m$ massive interface.[24]

D. *Dislocation Description and Energetics*

Referred to the bcc unit cell in Figure 3, the misfit built up between adjacent A_{11} and A_{13} along the $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ direction is known as $\mathbf{s}_3^{(1)}$, and that between A_{11} and its lower-right adjacent B_{22} as $\mathbf{b}_B$ or $(1/2)(\mathbf{s}_1^{(1)} + \mathbf{s}_3^{(1)})$ (Figure 10). With the inclined segment $A_{11}B_{22}$ resolving into a misfit-compensating ledge, the Burgers vector $\mathbf{b}_B$ decomposes into $(1/2)\mathbf{s}_1^{(1)}$ and $(1/2)\mathbf{s}_3^{(1)}$, or $(1/2)[100]_{bcc}$ with the risers and $(1/2)[0\bar{1}1]_{bcc}$ with the terraces.

Structural interfacial energy may now be assumed as dislocation strain energy evaluated with associated Burgers vectors. Suppose that the elastic strain energy of a dislocation is proportional to the square of its Burgers vector. The interfacial energy per unit interface segment along $A_{11}A_{13}$ (length: m) is proportional to the following:

$$(\mathbf{s}_3^{(1)})^2/m = 2a^2/m \qquad [18]$$

where a is the bcc lattice parameter, whereas the interfacial energy along $A_{11}B_{22}$ (length: n) is proportional to the following:

$$[(\mathbf{s}_1^{(1)}/2)^2 + (\mathbf{s}_3^{(1)}/2)^2]/n = (3/4)a^2/n < (3/2)a^2/m \qquad [19]$$

where the inequality arises from $n > m/2$. Comparison of Eqs. [18] and [19] indicates that the interfacial energy of the ledged interface with misfit-compensating ledges is less than that of the flat interface with misfit dislocations.

Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ interfaces have been suggested as probable conjugate habit planes in view of the good lattice matching recognized with a computer-aided graphic technique and the O-lattice theory.[17] As discussed here and earlier in Section II, misfit-compensating ledges and structural ledges may both develop with $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ terraces but stepping down along $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ and $[11\bar{2}0]_{hcp}//[111]_{bcc}$, respectively. Figure 11 shows the distribution of structural ledges and misfit-compensating ledges predicted for $\beta:\zeta_m$ massive interfaces[24] using the formulations discussed in Sections II–A and III–C. The apparent habit plane determined from this geometry was observed by TEM.[6]

IV. CRYSTALLOGRAPHY

A. *Precipitate Crystallography*

It has been widely recognized that crystallography, product morphology, and interphase boundary structure involved in solid-solid phase transformations are interrelated.[12,13,22,28,29,31,45] The O-lattice theory,[25] including the surface dislocation[42] and DSC lattice[25–27,36] models, is among the approaches to interlinking these important aspects.

The O-lattice theory can predict interfacial structure, such as misfit dislocations[17,25,26,46–48] and structural ledges and misfit-compensating ledges[24] as treated in Sections II and III. Since an orientation relationship is an input to the O-lattice framework, the theory has hardly been used in predicting optimal orientations between two crystals, except the only case treated by Bollmann and Nissen.[51] However, as structural ledges are related to such best-matching elements as parallel low-index planes and the lattice invariant line or O lines, the search of optimal orientation relationships may begin with the possible existence of structural ledges. This relation suggests that an O-line lattice needs to be constructed by rotating one crystal lattice with respect to the other just slightly away from certain low-index planar parallelisms.

As shown in Figure 12, the bcc/hcp case may again be treated as an example. From the Burgers orientation relationship (Figure 12(a)), the hcp structure is rotated around $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ with the originally parallel $(1\bar{1}00)_{hcp}$ and $(2\bar{1}\bar{1})_{bcc}$ planes deviated away from each other (Figure 12(b)). To some extent of rotation, which is dependent on the bcc/hcp lattice parameter ratio, a lattice invariant line can be recognized lying within the $(0001)_{hcp}//(0\bar{1}1)_{bcc}$ plane. In the O-lattice framework, however, the matching between the two lattices is best demonstrated by an O-line lattice (Figure 12(c)) in which

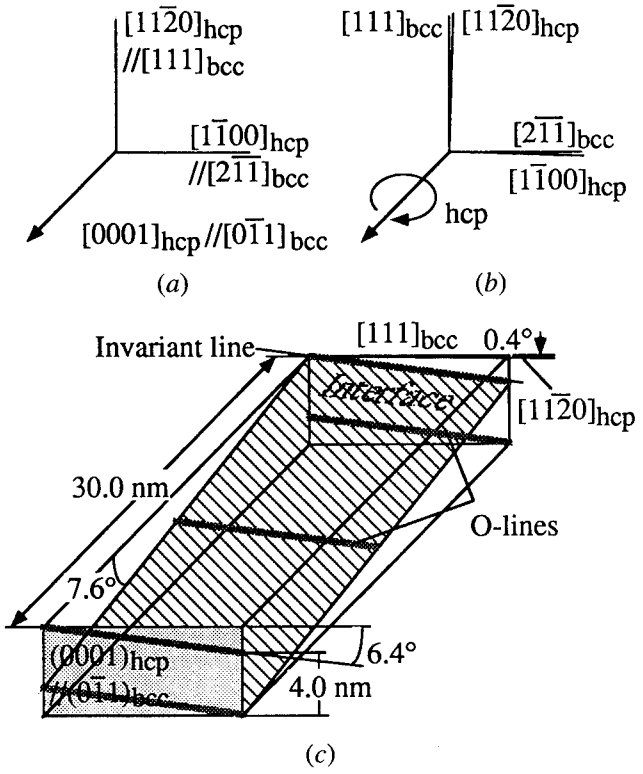


Fig. 12—(*a*) The Burgers orientation relationship; (*b*) rotation of the hcp phase around the $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ axis; and (*c*) O-line lattice obtained with the orientation relationship in (b) and the predicted low-energy interface.

all O-lines are parallel to or coincident with the invariant line.[23] Notice that in most situations a three-dimensional O-line distribution can only be obtained by introducing the concept of near coincidence.[23,25] Intersecting the parallel O-lines, an interface with somewhat low interfacial energy can be located, as shown in Figure 12(c). It should be emphasized that the inside *B* atom positions shown in Figure 3 have been taken into account in this bcc/hcp O-line lattice construction.

Although the two lattices are best matched along the invariant line or O-lines, the matching in the direction perpendicular to the O-lines requires consideration. The hcp lattice may be rotated again but around the invariant line to keep the best matching along it unchanged. In general, a best-matched interface may be finally determined since the O-line distribution is a function of rotation angle. It is found that, however, any rotation taking $[0001]_{hcp}$ away from $[0\bar{1}1]_{bcc}$ makes the matching worse for this specific example. The optimal orientation relationship can now be expressed as $[0001]_{hcp}//[0\bar{1}1]_{bcc}$ and the directional parallelism described with the invariant line. Figure 12(c) shows that this orientation relationship differs from the Burgers orientation by only 0.4 deg if the lattice parameters mentioned in Section III–C are employed. This angular difference in orientation relationship results in a difference of only 0.2 and 0.4 deg in apparent habit plane between the interfaces of Figures 11 and 12(c), with the hcp and bcc lattice taken as the reference in the formulation of Section II–A, respectively.

It is interesting to compare the present approach and the invariant line interpretation of optimal orientation relationships.[29] The invariant line model begins with a very general argument that an optimal orientation relationship between two lattices allows them to be matched with a lattice invariant line.[29] However, there are an infinity of lattice invariant lines that can be obtained with any rotations bringing the final cone into tangent to the initial cone of unextended lines.[45] The possible choices must then be restricted with additional conditions, for example, the requirement that the invariant line be the intersection of the initial cone and a low-index plane.[29,31,32] In contrast, the restriction to low-index planes is considered first in the structural-ledge approach. This difference is immaterial in most practical cases.

B. *Comparison of Martensite and Precipitate Crystallography*

Optimization of strain accommodation has long been suggested as the energetic basis for the phenomenological theory of martensite crystallography (PTMC).[49] The assumption that the martensite habit plane is both unstretched and unrotated during the transformation is energetically reasonable since an invariant plane strain can be related with minimum volume strain energy by Eshelby's approach.[50] As an extraction from the PTMC, the invariant line model is also based on optimal strain accommodation[28,29,31] and has been quite successful in predicting crystallography and morphology involved in diffusional transformations.[25,28–32]

On the other hand, the basic assumption in the O-lattice theory is that two crystals tend to optimize their matching across the interface at which they meet.[25] Best-matched interfaces are then associated with the minima of structural interfacial energy that can be evaluated or at least estimated from predicted interfacial defect structures. It is worth noting, however, that because the argument is purely geometric, the chemical component of interfacial energy has been ignored in the O-lattice theory.[25,36] The O-lattice approach also predicts that product crystals may have facets parallel to the best-matched planes.

An important crystallographic linkage[12,29,31] between diffusional and diffusionless transformations is that many transformations, regardless of their mechanisms, are associated with a lattice invariant line or O-line lattice. The concepts of optimal strain accommodation and of best lattice matching lead to the same mathematical formulation. The physical appearance of an invariant line is probably structural ledges,[21] as suggested by some evidence[2,3] and deduced from the best-matching assumption.[1,2,17,20,24] The success of the invariant line model can be alternatively attributed to its correct linkage to low interfacial energy arising from best lattice matching[23] at least for the case of platelike precipitates.

Although martensite and precipitate interfaces or apparent habit planes both include an invariant line, the situations may differ very much through rotations around their own invariant line. For instance, their interfacial structures may possess quite different characteristics arising from various thermodynamic and kinematic constraints.[52] The assumption of minimum structural interfacial energy in the O-lattice approaches discussed

before considers only the thermodynamic constraint. Optimal strain accommodation in the invariant line model is also based on thermodynamic arguments. It is obvious, however, that kinematic constraints are often the most important for martensite transformations, since misfit dislocations in martensite interfaces must be glissile[22,52] when their conservative motion is required, as in most martensite transformations.

The concept of a second rotation around the invariant line in Section IV–A may be helpful in understanding the dependence of martensite interfacial orientation on the mobility of interfacial misfit dislocations. Consider here only the simplest case where a single simple shear[45] is sufficient to produce an invariant plane strain. The kinematic constraint is that the Burgers vectors of misfit dislocations aligned parallel to the invariant line must be parallel to the product slip plane incorporating the lattice invariant line.[45,53] A martensite interface must then be oriented such that the mismatch across it, treated as a surface dislocation,[42] has a Burgers vector parallel to the given product slip plane.[53]

Volume strain energy and interfacial energy are competitive parameters: either can be dominant under some circumstances or both can be significant under other conditions. Moreover, lattice invariant strains considered in the PTMC may be very complicated (*e.g.*, multiple shears[45,49]) and system dependent, and interfacial energy minimization cannot be equally important in different transformation stages. The crystallography predicted with the PTMC, invariant line model, and best-matching approaches may probably give no distinct characteristics that distinguish between precipitate and martensite reactions under various conditions.

A martensite interface is an invariant plane during the transformation only in the sense that a lattice invariant strain (or a group of strains) cancels the transformation strain distributed in the interface.[45,52] A "*general*" martensite interface is not usually an invariant plane in the sense of lattice matching, given arbitrary lattice parameters. In the O-lattice framework, therefore, although a coherent twin interface is an O-plane,[25] a *general* martensite interface is not usually an O-plane. This is why a simple O-lattice approach does not apply to predicting martensite interfaces.[54] It is interesting enough that a precipitate interface can be an invariant plane in the same sense as is a martensite interface if all the misfit dislocations predicted with O-lattice constructions exist in the precipitate interface. The matrix product of the lattice deformation (the matrix **A** in Eq. [2]) and the lattice invariant deformation (a matrix formed with misfit dislocation Burgers vectors) may exactly give an invariant plane strain deformation.

It is no surprise, therefore, that the PTMC, at least the invariant line model extracted from it, can apply to some diffusional transformations as for predicting orientation relationships and habit planes. Obedience to the PTMC is not an adequate test for suggesting that a phase transformation be martensitic.[22] A more important criterion may probably be interphase boundary mobility that is dependent not only upon the kinematic natures of interfacial defects but also upon the formation and movement mechanisms of growth ledges.[12,22] Although interfacial mobility may be deduced to some extent from the kinematic natures of intrinsic interfacial dislocations and ledges, more reliable experiments are needed to understand the interaction between intrinsic defects and growth ledges.

V. SUMMARY

Structural ledges and misfit-compensating ledges at interphase boundaries were analyzed with an extended O-lattice approach[24,25,36] and the DSC lattice model.[25,26,27] The principle in these approaches remains the same as that in many previous O-lattice analyses:[17,25,46–48] Finding best-matching positions between two crystal lattices and locating interfaces that intersect as many best-matching positions as possible.[25,36] The requirements for best matching, however, are somewhat relaxed or particularly tailored to meet diversified needs.

Structural ledges were suggested to step down parallel low-index planes in the matrix and product phases when the interplanar spacings differ only slightly.[1,2,15] If the difference between two position vectors ending on adjacent parallel planes equals the difference in interplanar spacing, the two lattices will be somewhat best matched, especially when the spacing difference is very small. The distribution of structural ledges can thus be predicted with the spacing difference understood as a measure for the relaxed coincidence condition.[23,25] This approach describes the mismatch along the normal to the terraces of structural ledges as a series of misfit dislocations. These dislocations can be eliminated by a small rotation of one lattice with respect to the other, which generates a set of dislocations with a Burgers vector equal and opposite to the mismatch Burgers vector. Burgers vectors associated with structural ledges are DSC lattice vectors.

Misfit-compensating ledges[3–6] may be predicted at bcc:hcp interfaces as they intersect additional best-matching positions obtained from considering previously omitted bcc atom positions.[24] The O-lattice basic equation[25,36] is tailored to include atom position vectors in a structure base (two or more atoms belong to one lattice point). The additional best-matching positions such secured are O-points even under the strict coincidence condition.[25] Burgers vectors associated with misfit-compensating ledges are DSC lattice vectors. Energetic considerations based on dislocation strain energy suggest that the structural interfacial energy may lower when a flat Burgers-related $(1\bar{1}00)_{hcp}//(2\bar{1}\bar{1})_{bcc}$ interface is replaced by a ledged interface with misfit-compensating ledges.

Although both may include a lattice invariant line, precipitate and martensite interfaces possess different defect characteristics concerning thermodynamic and kinematic constraints.[52] Martensite interfaces are in general not invariant planes with respect to lattice matching, while a precipitate reaction may be associated with an invariant-plane strain if O-lattice-predicted interfacial defects all exist at the interface. A criterion in transformation mechanism, more important than obedience to the phenomenological theory of martensite crystallography, may probably be interphase boundary mobility

that is dependent not only upon the kinematic natures of interfacial defects but also upon the formation and movement mechanisms of growth ledges.[12,22]

ACKNOWLEDGMENTS

The author is deeply indebted to Dr. Hubert I. Aaronson, under whose supervision many concepts in this contribution were developed in the author's Ph.D. thesis investigations. Dr. Aaronson and the Conference Committee are gratefully appreciated for the invitation and financial support for the author's attendance at the conference.

REFERENCES

1. M.G. Hall, H.I. Aaronson, and K.R. Kinsman: *Surf. Sci.*, 1972, vol. 31, pp. 257-74.
2. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 351-76.
3. T. Furuhara, J.M. Howe, and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2873-86.
4. T. Furuhara and H.I. Aaronson: *Scripta Metall.*, 1988, vol. 22, pp. 1635-37.
5. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2887-99.
6. Yiwen Mou and H.I. Aaronson: Carnegie Mellon University, Pittsburgh, PA, unpublished research, 1992.
7. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. A*, 1949, vol. 198, pp. 205-16.
8. F.C. Frank and J.H. van der Merwe: *Proc. R. Soc. A*, 1949, vol. 198, pp. 216-25.
9. D.A. Porbter and K.E. Easterling: *Phase Transformations in Metals and Alloys*, Van Nostrand Reinhold Co., New York, NY, 1981, p. 145.
10. H.I. Aaronson, C. Laird and K.R. Kinsman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 313.
11. H.I. Aaronson: *J. Microsc.*, 1974, vol. 102, p. 275.
12. H.I. Aaronson: *Metall. Trans. A*, 1993, vol. 24A, pp. 241-76.
13. H.I. Aaronson: in *The Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, p. 387.
14. G.J. Shiflet, M.W.H. Braun, and J.H. van der Merwe: *S. Afr. Tydskrif vir Wentenskap*, 1988, vol. 84, p. 653.
15. J.H. van der Merwe, G.J. Shiflet, and P.M. Stoop: *Metall. Trans. A*, 1991, vol. 22A, pp. 1165-75.
16. G.J. Shiflet and J.H. van der Merwe: *J. Electron. Mater.*, 1991, vol. 20, pp. 785-91.
17. T. Furuhara and H.I. Aaronson: *Acta Metall. Mater.*, 1991, vol. 39, pp. 2857-72.
18. W.G. Burgers: *Physica*, 1934, vol. 1, p. 561.
19. D.I. Potter: *J. Less-Common Met.*, 1973, vol. 31, p. 299.
20. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, p. 365.
21. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, pp. 1029-34.
22. H.I. Aaronson, T. Furuhara, J.M. Rigsbee, W.T. Reynolds, Jr., and M.J. Howe: *Metall. Trans. A*, 1990, vol. 21A, pp. 2369-2409.
23. Yiwen Mou and H.I. Aaronson: Carnegie Mellon University, Pittsburgh, PA, unpublished research, 1992.
24. Yiwen Mou and H.I. Aaronson: Carnegie Mellon University, Pittsburgh, PA, unpublished research, 1992.
25. W. Bollmann: *Crystal Defects and Crystalline Interfaces*, Springer Verlag, Berlin, 1970.
26. D.H. Warrington and W. Bollmann: *Phil. Mag.*, 1972, vol. 25, pp. 1195-99.
27. R.W. Balluffi, A. Brokman, and A.H. King: *Acta Metall.*, 1982, vol. 30, pp. 1453-70.
28. U. Dahmen: *Scripta Metall.*, 1981, vol. 15, pp. 77-81.
29. U. Dahmen: *Acta Metall.*, 1982, vol. 30, pp. 63-74.
30. J.M. Lang, U. Dahmen, and K.H. Westmacott: *Phys. Status Solidi A*, 1983, vol. 75, p. 409.
31. U. Dahmen and K.H. Westmacott: *Acta Metall.*, 1986, vol. 34, pp. 475-82.
32. C.P. Luo and G.C. Weatherly: *Acta Metall.*, 1987, vol. 35, pp. 1963-72.
33. J.M. Howe and D.A. Smith: *Acta Metall. Mater.*, 1992, vol. 40, pp. 2343-50.
34. J.M. Howe and D.A. Smith: *Scripts Metall. Mater.*, 1993, vol. 28, pp. 1771-72.
35. M. Kato, T. Fujii and T. Mori: *Scripta Metall. Mater.*, 1993, vol. 28, pp. 1767-70.
36. D.A. Smith and R.C. Pond: *Int. Met. Rev.*, 1976, vol. 205, pp. 61-74.
37. R.W. Balluffi and G.B. Olson: *Metall. Trans. A*, 1985, vol. 16A, pp. 529-41.
38. A.P. Sutton and V. Vitec: *Scripta Metall.*, 1980, vol. 14, pp. 129-32.
39. A.P. Sutton, P.W. Balluffi, and V. Vitec: *Scripta Metall.*, 1981, vol. 15, pp. 989-94.
40. R.C. Pond: in *Dislocations in Solids*, F.R.N. Nabarro, ed., Elsevier, Amsterdam, 1989, vol. 8, pp. 1-66.
41. J.P. Hirth: *Acta Metall.*, 1974, vol. 22, pp. 1023-31.
42. B.A. Bilby, R. Bullough, and E. Smith: *Proc. R. Soc. A*, 1955, vol. 231, pp. 263-73.
43. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. TMS-AIME*, 1953, vol. 179, pp. 1503-14.
44. J.S. Bowles and J.K. MacKenzie: *Acta Metall.*, 1954, vol. 2, pp. 129-47.
45. C.M. Wayman: *Introduction to the Crystallography of Martensite Transformations*, Macmillan, New York, NY, 1964.
46. M.R. Plichta and H.I. Aaronson: *Acta Metall.*, 1980, vol. 28, pp. 1041-57.
47. Wang Rong, G.L. Dunlop, and K.H. Kuo: *Acta Metall.*, 1986, vol. 34, pp. 681-90.
48. E.S. Menon and H.I. Aaronson: *Acta Metall.*, 1986, vol. 34, pp. 1975-81.
49. J.W. Christian: *The Theory of Phase Transformations in Metals and Alloys*, Pergamon Press, Oxford, United Kingdom, 1975.
50. J.D. Eshelby: *Prog. Solid Mech.*, 1961, vol. 2, p. 89.
51. W. Bollmann and H.-U. Nissen: *Acta Crystallogr.*, 1968, vol. A24, p. 546.
52. G.B. Olson and M. Cohen: in *Dislocations in Solids*, F.R.N. Nabarro, ed., North Holland, New York, NY, 1986, vol. 7, p. 297.
53. R. Bullough and B.A. Bilby: *Pros. Phys. Soc.*, 1956, vol. 69B, pp. 1276-86.
54. J.W. Christian: in *New Aspects of Martensitic Transformations*, Proc. 1st JIM Int. Symp., Kobe, Japan, Suppl. to *Trans. JIM*, 1976,. pp. 21-33.

Atomic Site Correspondence and Surface Relief in the Formation of Plate-Shaped Transformation Products

JAMES M. HOWE

The concept of an atomic site correspondence between two phases is developed for simple diffusional and martensitic transformations involving precipitate plates. It is shown that a surface relief can be produced in both diffusional and martensitic phase transformations as a result of atomic site correspondence. Extension of this concept to more complicated transformations is discussed.

I. INTRODUCTION

ONE feature that has often been used to define a martensitic (displacive) transformation and, hence, to distinguish martensitic plates from those formed by a diffusional (reconstructive) transformation, is the presence of a surface relief effect.[1] The origin of the surface relief is due to the invariant plane strain associated with the martensitic transformation.[2] In contrast to this, Liu and Aaronson[3] presented experimental evidence that a surface relief occurred with the formation of hexagonal close-packed (hcp) γ (Ag_2Al) precipitate plates in Al-Ag alloys; a transformation which is classically diffusional. Further, in a series of articles concerning the formation of bainite,[4–7] Dahmen[7] presented a crystallographic analysis (using a lattice correspondence) to show that features such as an invariant plane strain and surface relief can be associated with diffusional transformations. In spite of these data, it is still not clear how a diffusional transformation, which proceeds by vacancy diffusion and in which there is no atomic correspondence between the phases, can produce a surface relief effect.[5,8] Hence, the idea that diffusional transformations exhibit a surface relief effect is not widely accepted at present.

The purpose of this article is to show that diffusional (reconstructive) transformations can exhibit a surface relief, not because there is a lattice correspondence (or an atomic correspondence) between the phases as in a martensitic (displacive) transformation, but because there is an *atomic site correspondence* across the interface. That is, in many diffusional transformations, there is continuity of atomic sites across the interface, resulting in only certain sites that atoms can occupy as they cross the interface. Thus, regardless of which atoms occupy particular atomic sites, continuity of atomic sites across an interface can preserve the shape deformation and produce a surface relief effect even in a diffusional transformation, as demonstrated in this article. The concept of an atomic site correspondence across an interface between two phases is directly related to the atomic mechanisms of the transformation, and this point is used to develop the concept of atomic site correspondence in Section II.

JAMES M. HOWE, Associate Professor, is with the Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA 22903.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

II. MECHANISMS OF ATOMIC SITE CORRESPONDENCE

In order to develop the idea of an atomic site correspondence between two phases, the discussion begins with a simple example, namely, the face-centered cubic (fcc) → hexagonal close-packed (hcp) transformation found in both martensitic (Co-Ni alloys for example[9,10]) and diffusional (Al-Ag for example[11,12,13]) alloys. The shape deformation associated with these transformations occurs by the passage of $a/6\langle 112\rangle$ Shockley partial dislocations on alternate $\{111\}$ matrix planes. This discussion is followed by further generalization to transformations exhibiting other types of shape deformation and by discussion of possible limitations and applicability of an atomic site correspondence.

A. *Atomic (Site) Correspondence in an fcc → hcp Martensitic (Displacive) Transformation*

The fcc → hcp martensitic transformation in Co and its alloys occurs by the passage of $a/6\langle 112\rangle$ Shockley partial dislocations on alternate $\{111\}$ matrix planes.[9,10,14,15] For purposes of discussion, assume that movement of the Shockley partial dislocations occurs by the migration of kinks along the dislocation line, as is known to occur for the fcc → hcp diffusional transformation in Al-Ag.[16,17] (It is subsequently shown that the end result of this discussion is the same independent of this assumption.) The plane in which the transformation occurs can be drawn in a "hard sphere" representation as shown in Figure 1, which shows the possibility of both edge and screw dislocation configurations. The plane below the transformation dislocation is shown as gray atoms, which are unaltered in the transformation. In Figure 1(a), the kinked partial dislocation is moving from top to bottom and thus separates a region of transformed hcp crystal (filled circles in B positions) from fcc matrix crystal (open circles in A positions). In Figure 1(b), the hcp crystal would expand by moving upward into the fcc matrix.

The sequence of sketches in Figures 2(a) through (d) then illustrates how the transformation front moves upward as the kink propagates by $\mathbf{b} = a/6[\bar{1}2\bar{1}]$ to the left one atom at a time. The screw dislocation in Figure 1(b)

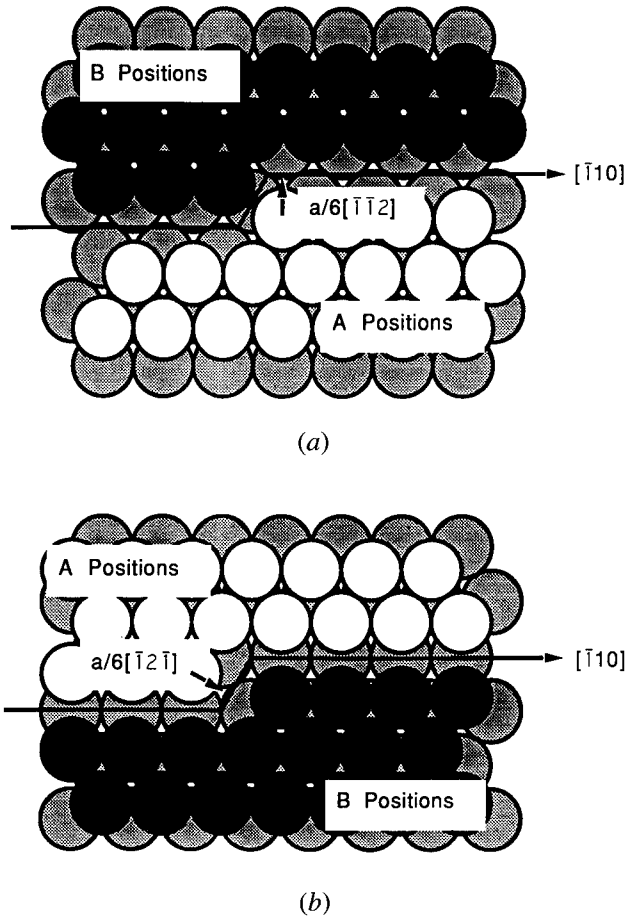


Fig. 1—Hard-sphere models showing the atomic configurations around kinks in $a/6\langle 11\bar{2}\rangle$ Shockley partial dislocations in the close-packed (111) matrix plane when the Burgers vector of the dislocation (*a*) is in an edge orientation with $\mathbf{b} = a/6[\bar{1}\bar{1}2]$ perpendicular to the dislocation line direction $\xi = [\bar{1}10]$ and (*b*) has a substantial screw component since it is inclined 30 deg to the line direction with $\mathbf{b} = a/6[\bar{1}2\bar{1}]$. The positions of the dislocation cores are indicated by solid lines in the figures, and arrows indicate the atomic displacements (Burgers vectors) associated with movement of an atom from an A position (open circle) to a B position (filled circle). The viewing direction is [$\bar{1}\bar{1}\bar{1}$], and the gray circles represent the plane of atoms in C positions beneath the dislocation core that is common to both the fcc and hcp phases.

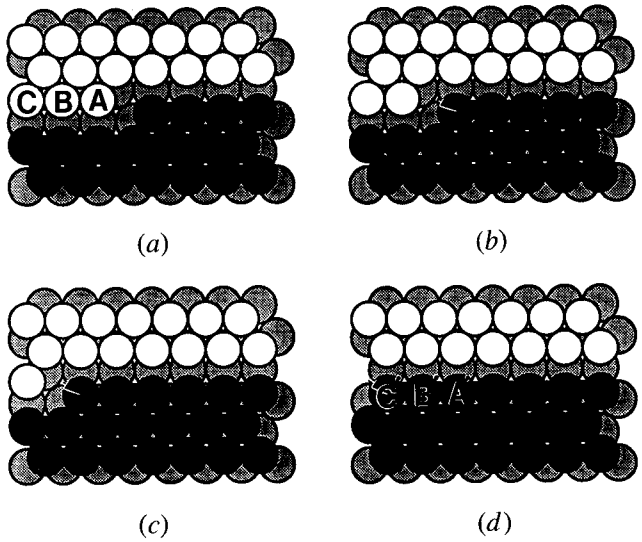


Fig. 2—(*a*) through (*d*) Hard-sphere model illustrating propagation of the kink in a screw-type dislocation to the left one atom at a time. Arrows indicate the atomic displacement associated with each atom. (a) is the same as Figure 1(b), and (d) shows the configuration after the kink has passed along the transformation dislocation line.

was used in Figure 2 for illustration, but a similar mechanism would apply to the edge dislocation in Figure 1(a). In an actual martensitic transformation, the transfer of atoms across the kink probably occurs at a rate similar to the vibrational frequency of the atoms, *i.e.*, about 10^{13} s^{-1},[18] so that the transformation dislocation moves upward at a velocity approaching the speed of sound.[9,19] It is particularly important to note that a definite atomic correspondence is associated with the motion of the martensitic interface in Figure 2. For example, based only on the crystallography of the transformation, *i.e.*, the Burgers vector of the Shockley partial dislocation, it is possible to predict that the atoms labelled A, B, and C in the matrix in Figure 2(a) will be the corresponding atoms labeled A′, B′, and C′ in the product phase in Figure 2(d). This occurs because there is no long-range diffusion or mixing of atoms in the martensite transformation. Note that if the interface in Figure 2 had propagated downward by motion of the entire dislocation line or large segments of it by $a/6\langle 11\bar{2}\rangle$ rather than by discrete motion at kinks, the end result of this argument would be the same. In this case, the mechanism of the transformation occurs in a classic military fashion, as discussed by Christian.[20] The same result would also apply to an interface where the kinks are considerably more diffuse.

It is also important to note that as the transformation interface continues to propagate upward, *i.e.*, A atoms are converted to B atoms, it leads to a surface relief effect, as illustrated in Figure 3. In Figure 3(a), the top row of A atoms constitutes the surface of the crystal, which is indicated by a solid line labelled S. As the transformation interface propagates through the crystal and intersects the surface, displacement of the surface occurs to position S′ (Figure 3(b)). Complete transformation leads to the new surface position of the hcp phase at S′ shown in Figure 3(c). The displacement of the surface (S-S′) is equal to the edge component of the Burgers vector of the transformation dislocation normal to the surface, in this particular case $a/12[11\bar{2}]$, and this is the surface relief associated with the single-Shockley partial dislocation. This surface relief would be doubled for the edge partial dislocation in Figure 1(a).

B. *Atomic Site Correspondence in an fcc → hcp Diffusional (Reconstructive) Transformation*

Now, consider the same interface as in Figures 1(b) and 2(a), but let it move upward by the substitutional diffusion of solute atoms to the kink in the Shockley partial dislocation. For simplicity, assume that the hcp phase is composed entirely of element *X* (represented by filled circles) and that the fcc matrix phase is a dilute solution of *X* in *Y* (represented by open circles). Further, assume that diffusion only occurs in the plane containing the transformation dislocation. (In reality, diffusion of solute would likely involve jumps above and

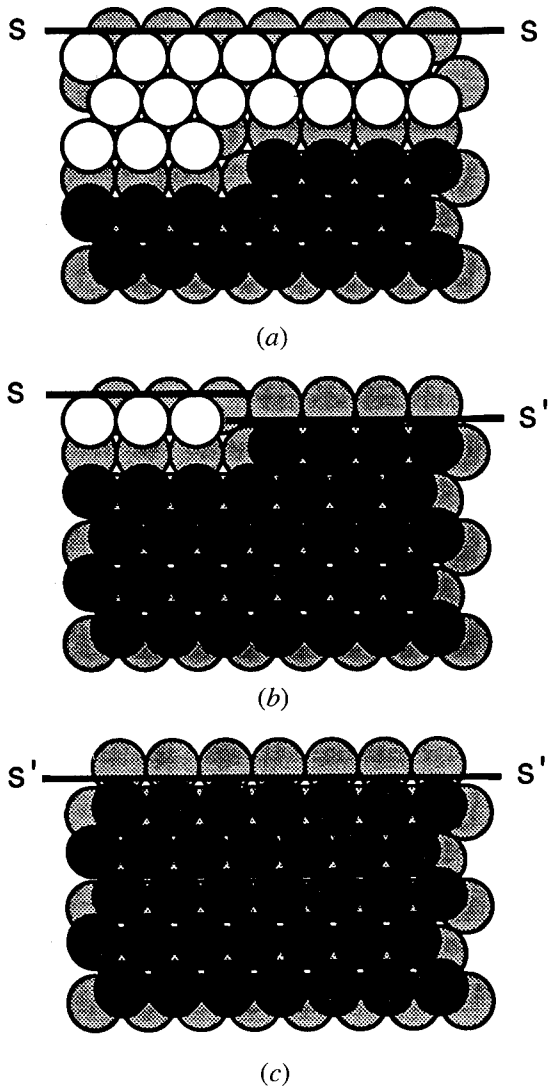


Fig. 3—Hard-sphere model illustrating the motion of a transformation dislocation out to the surface of a crystal: (*a*) the same starting configuration as in Figure 1(b), (*b*) after advancement of a portion of the transformation dislocation to the surface; and (*c*) after complete intersection of the dislocation with the surface. The surface step associated with the intersection is S-S′ = $a/12[11\bar{2}]$.

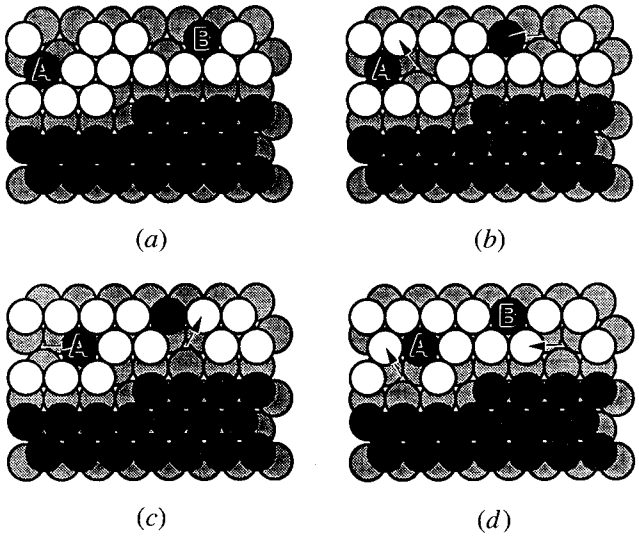


Fig. 4—(*a*) through (*d*) A series of atomic jumps involving fcc matrix atoms Y (open circles), vacancies, and two solute atoms labeled A and B (filled circles). The hcp precipitate is assumed to consist of pure X (filled circles) for simplicity. The atomic jumps are indicated by arrows in each figure. In (d), the atom labeled A is nearly adjacent to the kink and in a position to cross the interface, as illustrated further in Figure 5.

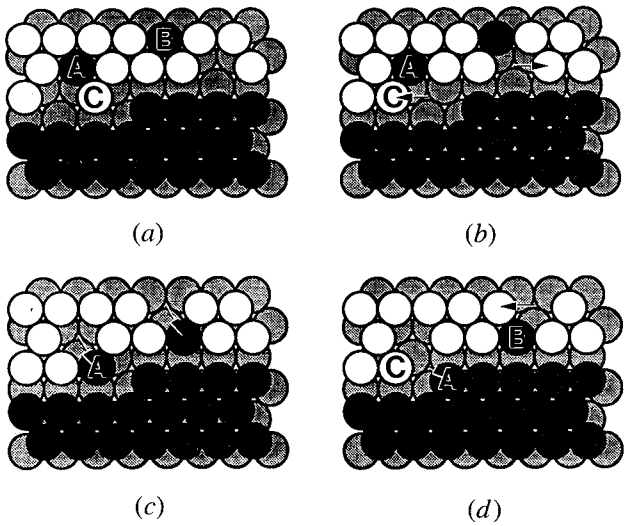


Fig. 5—(*a*) through (*d*) A series of diffusional jumps which leads to the incorporation of atom A into the kink and propagation of the kink one atom to the left. The atomic jumps are indicated by arrows in each figure.

below the plane of the transformation dislocation.) A sequence of possible diffusional jumps that could lead to propagation of the kink one atom to the left is illustrated in Figures 4 and 5.

In Figure 4(a), two solute atoms labeled A and B and two vacancies were included in the matrix to allow for substitutional diffusion. Figures 4(b) through (d) represent a series of atomic exchanges (like time snapshots) between the solute atoms and vacancies that ultimately leads to the situation shown in Figure 4(d), where the solute atom labelled A is in a position to cross the interface at the kink.

The actual mechanism by which the solute atom A makes the jump across the interface at the kink is not known, but several simple processes can be envisioned, one of which is illustrated in Figure 5. In Figure 5(a), the vacancy exchanges with the atom adjacent to the kink labelled C, as shown in Figure 5(b). Atom A then fills the vacant atom position, as shown in Figure 5(c). It subsequently shuffles into the final position to join the hcp phase as the kink propagates one atom to the left, as illustrated in Figure 5(d). Notice that the sequence in Figure 5 assumes that the transformation dislocation core separates the two phases but does not participate in the reaction. In an actual interface where atoms are associated with the transformation dislocation core, the sequence of atomic jumps shown in Figure 5 would occur within the dislocation core rather than to one side of the core, although the basic process would remain the same. In this article, the core atoms were omitted to simplify the figures.

The possibility that atom A joins the kink by filling the vacant atom site in Figure 5(b) is an appropriate point to introduce the concept of atomic site correspondence. Referring back to Figure 4(a) and considering the series of diffusional jumps which followed, it is clear that it was not possible to know whether atom A or B, or any other solute atom that may have been present, would be in a position to cross the kink as shown in Figure 5. Hence, the atomic mixing that occurs in a diffusional transformation clearly negates any atomic correspondence between the two phases. However, when one focuses on the process of atomic attachment to the kink illustrated in Figure 5, it is apparent that once an atom has reached a position where it can cross the interface at the kink (such as atom A in Figure 5(c)), there is only one *site* where it can join the new phase. This site is indicated by the tip of the arrow in Figure 5(d). Hence, while there is no atomic correspondence between the two phases in a diffusional (reconstructive) transformation, *i.e.*, it occurs in a classically civilian fashion,[8,20] there can be an *atomic site correspondence* across the interface. This atomic site correspondence arises from the continuity of atoms that occurs across an interface that is fully or partially coherent and it is determined by the crystallography of the transformation, *i.e.*, by the Burgers vector of the transformation dislocation.

The result of an atomic site correspondence is that the shape deformation that occurs between two phases is preserved across the transformation interface, even though there is no atomic correspondence between the phases. Thus, as the kink in the diffusional interface in Figures 4 and 5 continues to move to the left, the resulting interface is identical to that shown in Figure 2(d). Furthermore, as the interface moves upward by the same diffusional mechanism and intersects the surface, it produces a surface relief which is identical to the martensitic transformation shown in Figure 3(c) because of the atomic site correspondence across the interface. This is not shown, but it can be readily demonstrated by propagating the kink across the dislocation line in Figure 5, just as in Figure 3 for the martensitic transformation.

In summary, the previous discussion has shown that an atomic site correspondence exists across the interface in an fcc → hcp transformation and that the atomic site correspondence is capable of producing a surface relief effect regardless of whether or not there is an atomic correspondence between the two phases. The atomic site correspondence is due to the coherency between the two phases that is provided by the transformation dislocation. The concept of an atomic site correspondence across other types of interface is discussed further in Section C.

C. *Extension to Other Types of Transformation*

In the fcc → hcp transformations mentioned earlier, the shape deformation (Burgers vector) was parallel to the habit plane of the plate-shaped transformation product, and the interface was coherent and glissile. In addition, the unit cell of the hcp phase was only two {111} matrix planes high, and it contained only two atoms. This was convenient because it allowed the concept of an atomic site correspondence to be developed for a well-defined transformation, but it represents a specific situation and thus suffers from a lack of generality. Many transformations have a shape deformation that contains a nonglissile (climb) component, and the unit cells can be quite large and contain a number of different elements. Thus, it is important to consider whether the idea of an atomic site correspondence is valid for such transformations or whether some transformations do not support any type of atomic site correspondence. At present, there are so few diffusional transformations where the atomic correspondence across the interface is known that it is difficult to answer the question of generality of atomic site correspondence in various types of transformation. However, consideration of two additional situations provides some insight into this issue and is presented next.

The fcc → hcp transformation that was accomplished by the passage of Shockley partial dislocations parallel to the interface could have been accomplished by the climb (expansion) of an $a/3\langle 111\rangle$ Frank partial dislocation along the interface. This situation represents the opposite extreme in terms of a shape deformation and transformation dislocation. In this case, the $a/3\langle 111\rangle$ dislocation behaves as a pure expansion (or contraction) deformation and in terms of the edge of the precipitate plate would be a misfit or anticoherency dislocation.[21] The θ' (Al_2Cu) precipitates in Al-Cu alloys display this type of shape deformation, although the atomic correspondence between the matrix and precipitate phase is more complicated[22,23] and γ' (Ag_2Al) precipitates in Al-Ag have been shown to form by this mechanism as well.[11,24] Given this mode of transformation, the Frank partial would climb by atom attachment to jogs in the dislocation.[18,25] Since the atom positions in and around the dislocation core are uniquely defined and many atomic planes around the dislocation core are continuous between the fcc and hcp phases, an atomic site correspondence exists for this mode of transformation as well. Any particular transformation can contain both glide and climb components and thus lie somewhere between the two extremes analyzed. In this case, some type of atomic site correspondence between two phases should exist provided the transformation dislocation maintains some degree of coherency between the two structures.[2,21,26]

As an example of a much more complicated transformation where the growth ledge contains both glide and climb components and the crystal structures are very different across the interface, consider the case of {111} θ precipitate plates that form in Al-Cu-Mg-Ag alloys. The atomic correspondence across a unit-cell ledge between the fcc α-Al and tetragonal θ (Al_2Cu) precipitate structure has been determined and is reproduced in Figure 6.[27] The most important points to note from Figure 6 are that (1) the θ unit-cell contains a residual shear component of $a/6\langle 112\rangle$ and a 9.3 pct contraction along $\langle 111\rangle$ relative to the fcc matrix, (2) formation of one unit-cell of θ is nonconservative with six Al atoms diffusing out of and four Cu atoms diffusing into the structure, and (3) there is preservation of some atomic sites between the two phases in the Al planes labelled ABCAB in the matrix and planes Al(1) through Al(4) in

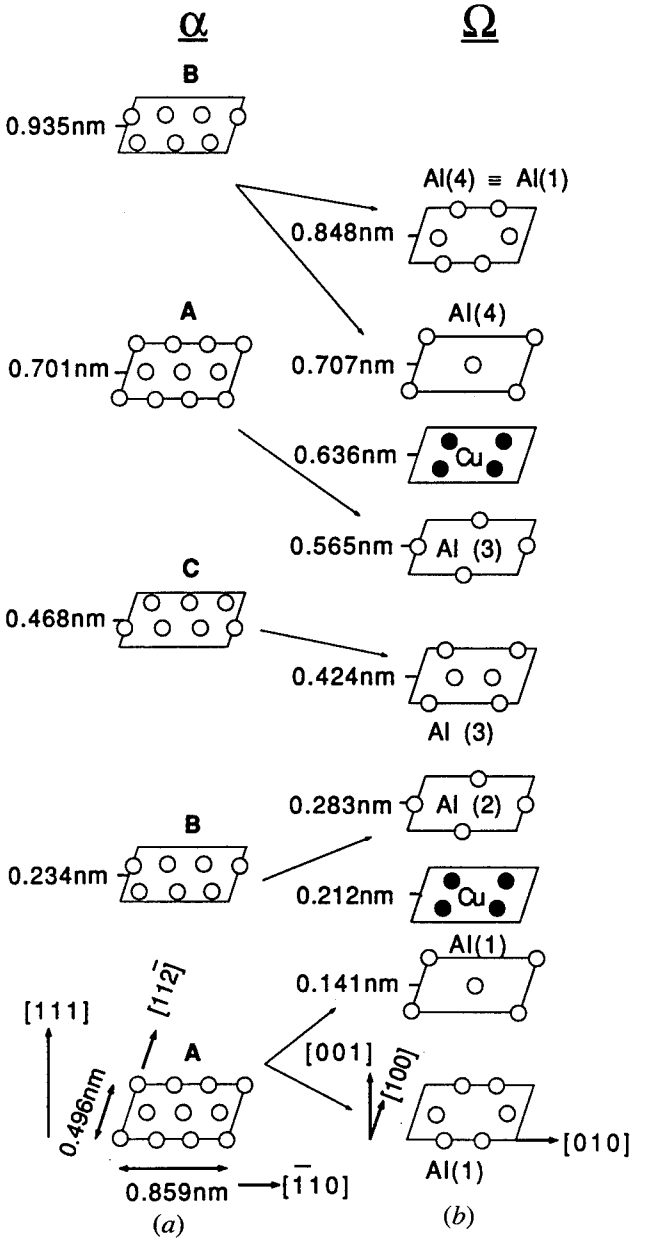


Fig. 6—Illustration of the atomic correspondence between the (*a*) fcc matrix and (*b*) plate-shaped {111} θ precipitate phase in the Al-Cu-Mg-Ag system. A detailed description of the transformation is provided in Ref. 27.

the precipitate. Thus, even in this relatively complicated diffusional transformation, which is nonconservative, where many atoms of both Cu and Al are involved, and where the shape deformation includes both a shear and contraction, there is some atomic site correspondence between the two phases.

D. *Summary*

The previous examples illustrated the concept of atomic site correspondence between two crystalline plate-shaped phases and the effect of this correspondence on properties of the transformation such as a surface relief. They also provided a qualitative assessment of the degree to which atomic site correspondence can be maintained between two crystals and indicated that there is a dependence on the degree of coherency of the phases across the interface. These concepts were shown to apply regardless of whether the interface advances by discrete kinks or continuously along its line. However, this treatment was largely conceptual and it needs to be analyzed rigorously, perhaps using a description along the lines of the coincident site lattice (CSL) theory[28,29,30] or based on the symmetry between the phases.[31] This area will be addressed in the future. It is also important to note that while the discussion in this article focused on plate-shaped transformation products, atomic site correspondence should apply to other precipitate shapes as well.

III. CONCLUSIONS

1. Surface relief can be produced in both martensitic (displacive) and diffusional (reconstructive) transformations, and this is due to an atomic site correspondence between the phases across the interface in both types of transformation.
2. There is no atomic correspondence in diffusional transformations but continuity of atomic planes across the transformation interface can lead to an atomic site correspondence between the two phases.
3. Since some degree of coherency is maintained across the interface in many phase transformations, it is likely that atomic site correspondence is a general phenomenon.
4. The concept of atomic site correspondence needs to be developed within a rigorous analytical framework.

ACKNOWLEDGMENTS

The concept of atomic site correspondence was developed during the course of this symposium and the author is grateful to all of the participants, particularly Professor B.C. Muddle and Dr. U. Dahmen, whose presentations and discussions contributed to this development. The author is also indebted to Professors H.I. Aaronson and C.M. Wayman for providing the opportunity to participate in this symposium. This research was supported by the National Science Foundation under Grant No. DMR-9107550.

REFERENCES

1. H.M. Clark and C.M. Wayman: *Phase Transformations*, ASM, Metals Park, OH, 1970, pp. 59-114.
2. B.A. Bilby and J.W. Christian: *The Mechanisms of Phase Transformations in Metals*, The Institute of Metals, London, 1956, pp. 121-32.
3. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, pp. 845-57.
4. H.I. Aaronson and H.J. Lee: *Scripta Metall.*, 1987, vol. 21, pp. 1011-16.
5. H.K.D.H. Bhadeshia: *Scripta Metall.*, 1987, vol. 21, pp. 1017-22.
6. G.B. Olson: *Scripta Metall.*, 1987, vol. 21, pp. 1023-28.
7. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, pp. 1029-34.
8. J.W. Christian: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience Publishers, New York, NY, 1962, pp. 371-86.

9. Z. Nishiyama: *Martensitic Transformation*, Academic Press, New York, NY, 1978, pp. 48-60 and 232-38.
10. P. Gaunt and J.W. Christian: *Acta Metall.*, 1959, vol. 7, pp. 529-33.
11. R.B. Nicholson and J. Nutting: *Acta Metall.*, 1961, vol. 9, pp. 332-43.
12. C. Laird and H.I. Aaronson: *Acta Metall.*, 1967, vol. 15, pp. 73-103.
13. J.M. Howe, U. Dahmen, and R. Gronsky: *Phil. Mag. A*, 1987, vol. 56, pp. 31-61.
14. C. Hitzenberger and H.P. Karnthaler: *Phil. Mag. A*, 1991, vol. 64, pp. 151-63.
15. G.B. Olson and M. Cohen: *Metall. Trans. A*, 1976, vol. 7A, pp. 1897-1904.
16. J.M. Howe and N. Prabhu: *Acta Metall. Mater.*, 1990, vol. 38, pp. 881-87.
17. N. Prabhu and J.M. Howe: *Acta Metall. Mater.*, 1990, vol. 38, pp. 889-96.
18. J.P. Hirth and J. Lothe: *Theory of Dislocations*, 2nd ed., John Wiley and Sons, New York, NY, 1982, pp. 532-34, 555-59, and 569-73.
19. R.F. Bunshah and R.F. Mehl: *Trans. TMS-AIME*, 1953, vol. 197, pp. 1250-58.
20. J.W. Christian: *The Theory of Transformations in Metals and Alloys*, Pergamon Press, Oxford, 1965, pp. 1-16 and 802-12.
21. G.B. Olson and M. Cohen: *Acta Metall.*, 1979, vol. 27, pp. 1907-18.
22. G.C. Weatherly and R.B. Nicholson: *Phil. Mag.*, 1968, vol. 17, pp. 801-31.
23. U. Dahmen and K.H. Westmacott: *Phys. Status Solidi A*, 1983, vol. 80, 249-62.
24. J.A. Hren and G. Thomas: *Trans. TMS-AIME*, 1963, vol. 227, pp. 308-18.
25. J. Silcox and M.J. Whelan: *Phil. Mag.*, 1960, vol. 5, pp. 1-23.
26. J.W. Christian: *Dislocations and Properties of Real Materials*, The Institute of Metals, London, 1985, pp. 94-124.
27. A. Garg, Y.C. Chang, and J.M. Howe: *Acta Metall. Mater.*, 1993, vol. 41, pp. 235-51.
28. A. Brokman and R.W. Balluffi: *Acta Metall.*, 1981, vol. 29, pp. 1703-19.
29. R.W. Balluffi, A. Brokman, and A.H. King: *Acta Metall.*, 1982, vol. 30, pp. 1453-70.
30. J.P. Hirth and R.W. Balluffi: *Acta Metall.*, 1973, vol. 21, pp. 929-42.
31. R.C. Pond: *Dislocations in Solids*, Elsevier Publishers, Amsterdam, 1989, vol. 8, pp. 9-65.

Formation of Invariant Plane-Strain and Tent-Shaped Surface Reliefs by the Diffusional Ledge Mechanism

M.G. HALL and H.I. AARONSON

Electron channeling contrast and electron backscattering patterns (EBSPs), generated in scanning electron microscopes (SEMs), were used to study the question of whether tent-shaped surface reliefs associated with ferrite plates in Fe-0.07 pct and Fe-0.50 pct C low-alloy steels are single crystals or back-to-back pairs of plates. Although more than 50 ferrite plates were examined in detail, any misorientation contained within them must have been <1 deg. Additionally, no sightings of the low-angle grain boundaries approximately bisecting ferrite plates expected on the back-to-back plates mechanism were ever made. The efforts of previous investigators to explain tent-shaped (and invariant plane-strain) reliefs on the basis of diffusional ledgewise growth were further extended.

I. INTRODUCTION

AN invariant plane-strain (IPS) surface-relief effect formed where a plate-shaped transformation product intercepts a free surface polished flat prior to transformation has long been accepted as a defining characteristic of a shear or martensitic phase transformation.[1–4] Diffusional transformations have also been found, however, to be capable of producing IPS reliefs. Changes in long-range order and/or in composition (the latter particularly in substitutional alloys) relative to the matrix phase taking place continuously throughout the transformation process are usually acceptable criteria for diffusional transformations.[5,6] A prototypical example of a drastic change in long-range order across the transformation interface accompanying an IPS surface relief is represented by the formation of AuCu II plates from a disordered equiatomic Au-Cu matrix.[7] Counterpart examples of a marked change in composition formed in conjunction with an IPS relief are the precipitation of γ' (later γ) $AlAg_2$ plates from α Al-Ag solid solutions[8,9] and of α_1 Cu-Zn plates from ordered β Cu-Zn.[10,6] There do not appear to have been any questions raised about the long-range order change in AuCu II or about the composition change in γ or γ' $AlAg_2$. However, the question as to whether or not α_1 Cu-Zn plates inherit the composition of their ordered β matrix has been under active dispute in the literature for some time. In these proceedings, though, Hamada *et al.*[10] have evidently brought to an end the long controversy about the composition of α_1 plates at early growth times by demonstrating that even when these plates are only 3-nm thick, their composition differs significantly from that of their matrix phase. Also in these proceedings, Muddle[11] confirms that the transformation crystallographies of AuCu II, γ $AlAg_2$, and α_1 Cu-Zn plates are all in good agreement with the predictions of the phenomenological theory of martensite crystallography (PTMC).[1,2] Hence, the presence of an IPS surface-relief effect no longer serves to identify uniquely the operation of a shear transformation mechanism.

A further problem that has arisen in connection with the use of surface reliefs to identify the atomic mechanism of plate formation is the appearance of surface reliefs that are not an IPS but that are nonetheless composed of tilted planar surfaces (rather than the curved surfaces implied by the term "surface rumpling"[5]). Of these, the simplest and best known is the tent-shaped or roof-gable[6] relief. Examples of tent-shaped reliefs associated with plates or laths have been found in the proeutectoid ferrite reaction in Fe-C alloys,[12,13,14] lower bainite plates in an Fe-C-Cr alloy,[15] Cr-rich precipitates in a Cu-Cr alloy,[16] and proeutectoid α plates in a hypoeutectoid Ti-Cr alloy.*[17] This relief morphology cannot be explained in terms of the PTMC unless one simply assumes that such a relief is produced by two plates oriented parallel to the same conjugate habit planes but having different conjugate habit directions. Except during the crystallographically degenerate face-centered cubic/hexagonal close-packed (fcc/hcp) transformations,** one of these plates must thus have been sympathetically nucleated[18] at the broad face of the other. In response to presentation of a micrograph of a ferrite plate that exhibited a pronouncedly tent-shaped surface relief,[12] Srinivasan and Wayman[19] proposed such a back-to-back formation mechanism for the development of this relief. While neither thermionic electron emission microscopy contrast nor nital etching gave any indication that the sideplates observed by Kinsman *et al.*[14] were actually pairs of plates, the discrimination permitted by these techniques was obviously limited. Lee and Aaronson[17] mechanically ground and polished and then ion milled from the side opposite to that on which reliefs had been observed some Ti-Cr specimens partially transformed to hcp proeutectoid α in

*Among the surface reliefs examined in the Ti-Cr alloy, approximately 80 pct were tents, and the balance were of the IPS type.[17]

**Christian has privately pointed out to us that in the presence of a near crystallographic degeneracy, *e.g.*, In-Tl, banded morphologies, equivalent to "tents of indefinite length," tend to be formed.

M.G. HALL, Lecturer, is with the School of Metallurgy and Materials, University of Birmingham, Birmingham, B15 2TT, United Kingdom. H.I. AARONSON is with Geo-Centers, Inc., Ft. Washington, MD, and is stationed in the Physical Metallurgy Branch, Code 6320, the Naval Research Laboratory, Washington, DC 20375-5343.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

a body-centered cubic (bcc) β matrix. Transmission electron microscopy (TEM) observations made on the foils thus prepared showed that all α plates observed, most of which had almost certainly contributed to the reliefs seen on the specimen surface, were single crystals. They thus demonstrated that the tent-shaped reliefs in this alloy could not have been formed by pairs of plates.

The present investigation was undertaken to perform a higher-resolution test of the sympathetic nucleation mechanism for tent-shaped surface-relief formation on proeutectoid ferrite plates. Because these plates are so large relative to the thickness of typical thin foils, different techniques had to be employed. Those chosen were electron channeling contrast and electron backscatter patterns (EBSPs), both generated in scanning electron microscopes (SEMs). Bulk specimens could now be employed instead of thin foils.

II. EXPERIMENTAL PROCEDURES

Two low-alloy steels were employed in this investigation. They contained 0.07 and 0.50 pct C, respectively. Their detailed compositions are given in Table I. Initial experiments were performed on specimens of the 0.50 pct C alloy that had been continuously cooled from the austenite region so as to yield a number of Widmanstätten sideplates. These specimens were used for EBSP studies to determine whether or not the sideplates were single crystals. Surface-relief specimens were prepared from the Fe-0.07 pct C alloy, which was given isothermal heat treatment.

The EBSP specimens were first mechanically polished and then electropolished in a CrO_3-glacial acetic acid solution[20] so that there was a minimum of surface topography. They were then examined with a JEOL* 840A SEM, using the backscattered electron detector. Typical beam conditions were 30 kV, 3 to 7 $\times$ 10^{-10} A, and a working distance of 0.008 m. Images were acquired using a frame-store that controlled the scanning of the SEM,[21] and enabled noise-free images to be produced with lower beam currents than would have been required using conventional scanning electron microscopy. Plates were selected for further study that diffracted strongly so that small orientation changes would be visible using electron channeling contrast. Several such plates were then relocated with the optical microscope of a microhardness tester and identified with impressions in this instrument. The specimen was then transferred to a Hitachi Hu 4000 field-emission SEM equipped with a Dingley EBSP system.[22]

*JEOL is a trademark of Japan Electron Optics Ltd., Tokyo, Japan.

The EBSP technique requires a highly tilted sample. With the limited contrast available, the microhardness indentations proved essential to again locating the selected plates. At least eight, and usually more, EBSPs were taken at successive positions across the width of these plates. The patterns were recorded either on photographic film or on video tape; the latter were then printed with a video printer.

Surface-relief specimens of the 0.07 pct C alloy were mechanically polished and sealed in quartz ampoules under a reduced pressure of purified He. These specimens were then austenitized at 1420 K, reacted at 973 K for times up to 180 seconds, and then quenched or air-cooled to room temperature. Particularly clear examples of tent-shaped surface reliefs were selected optically, marked with microhardness impressions, and then examined with the HU 4000 field-emission SEM using the techniques just described.

III. RESULTS

1. *Resolution of the Electron Backscatter Technique*

A junction between two plates yielding tent-shaped reliefs (Figure 1(a)) was chosen, and EBSPs were taken at 50-nm intervals along a line perpendicular to the impingement boundary. Seven patterns from a sequence of 10 are shown in Figures 1(b) through (h). The boundary between these two plates is straddled in Figure 1(e), where a mixed pattern composed of contributions from crystals on both sides of the boundary is present. Figures 1(f) and (g) are seen to show the same pattern, but one that is abruptly rotated relative to the pattern shown in Figures 1(b) through (d). Hence, the resolution of this method must be better than 100 nm. This resolution is consistent with that reported by Randle,[23] who indicated a resolution of 200 nm for an SEM with a conventional W emitter. The field-emission SEM used here is capable of generating a smaller probe size at the beam currents required to produce an EBSP and therefore should give improved spatial resolution, as was experimentally observed.

2. *Electron Channeling Contrast Studies*

Three sample electron channeling contrast micrographs, taken of two ferrite sideplates but representative of the more than 50 examined during this part of the investigation, are shown in Figures 2(a) through (c). The overall morphological perfection of the Widmanstätten sideplates produced by continuous cooling was apparent in these experiments, in agreement with the more detailed study of Mehl *et al.*[24] Figure 2(a) was taken in the vicinity of the base of a sideplate formed by sympathetic nucleation at a grain boundary allotriomorph. Figure 2(b) shows the region near the tip of the same sideplate. Figure 2(c) displays much of another sideplate

Table I. Alloy Compositions in Weight Percent

Alloy No.	C	Mn	Si	Ni	V	Al	S	P
1	0.07	0.65	0.05	0.96	0.14	0.03	0.04	0.04
2	0.50	0.61	0.5	0.07	0.003	0.12	0.03	0.02

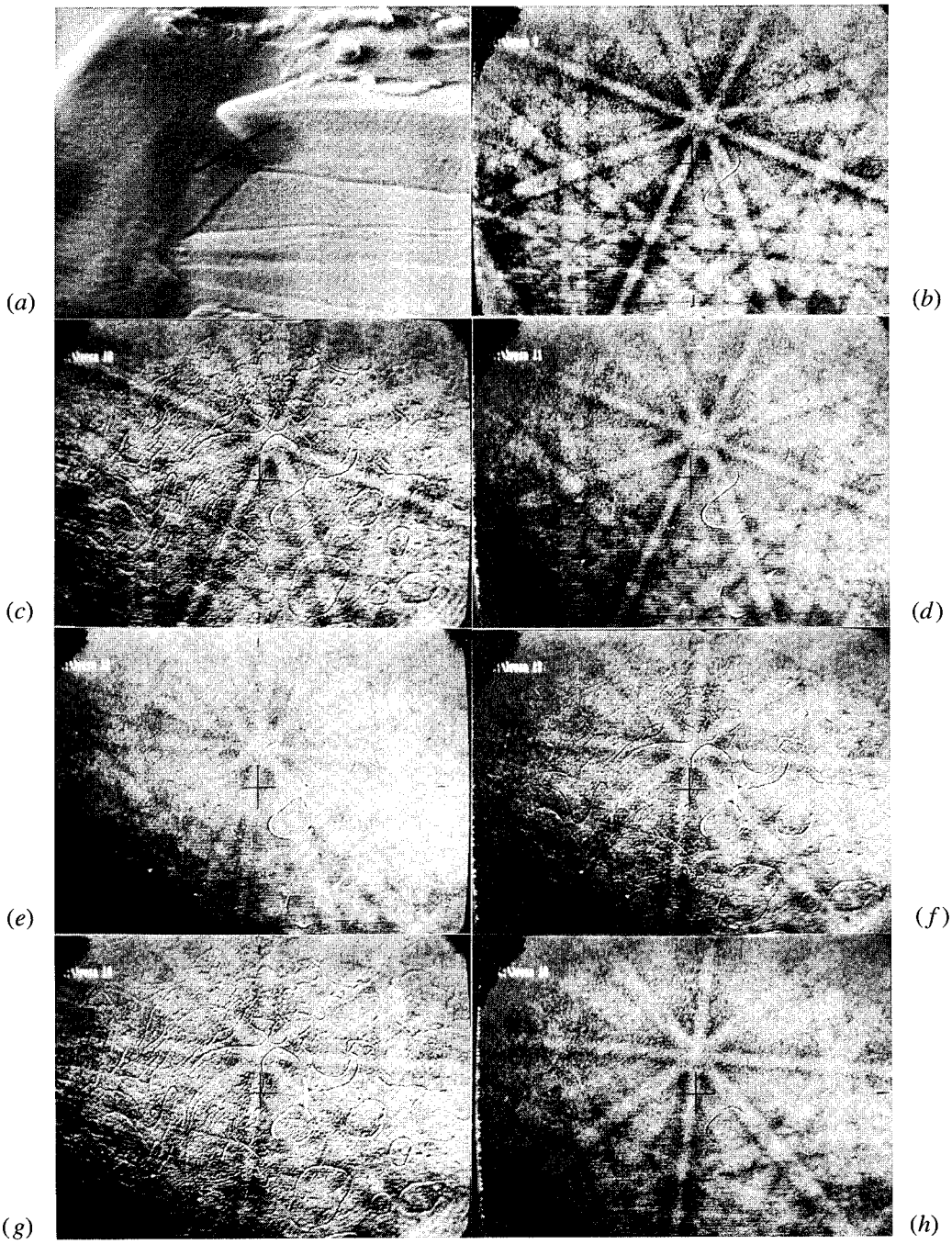


Fig. 1—An illustration of the resolution of the EBSP technique. Patterns (*b*) through (*h*) were taken at 50-nm intervals along the line indicated in (*a*). An abrupt change in the pattern is seen to have occurred between (*d*) and (*f*). Only in (*e*) is there a pattern composed of contributions from ferrite crystals on both sides of the boundary. Note: superimposed on some of the patterns is the outline of the islands of good and poor contact between the phosphor support and the light guide. This aspect of the pattern can be ignored and does not affect the positions of the major poles in the EBSPs.

that appears to have developed directly from a grain boundary allotriomorph and to be part of the same single crystal, thereby confirming a mechanism proposed earlier.[25] Despite the presence of clear ferrite:ferrite boundaries within the impinged grain boundary allotriomorphs, and between the sideplate and the allotriomorph from which it developed (Figure 2(a)) there is no evidence for a boundary within any ferrite sideplate aligned parallel to its long axis. Although it is possible that such a boundary would occasionally be missed because of sectioning effects, it is unlikely that all of these boundaries would be thus omitted, particularly since plates with the smallest aspect ratio in a particular area were always chosen for study.

Electron backscatter patterns were also used to confirm that these plates were single crystals, as illustrated in Figure 3. This figure includes three micrographs from a sequence of eight taken across a similar sideplate. The

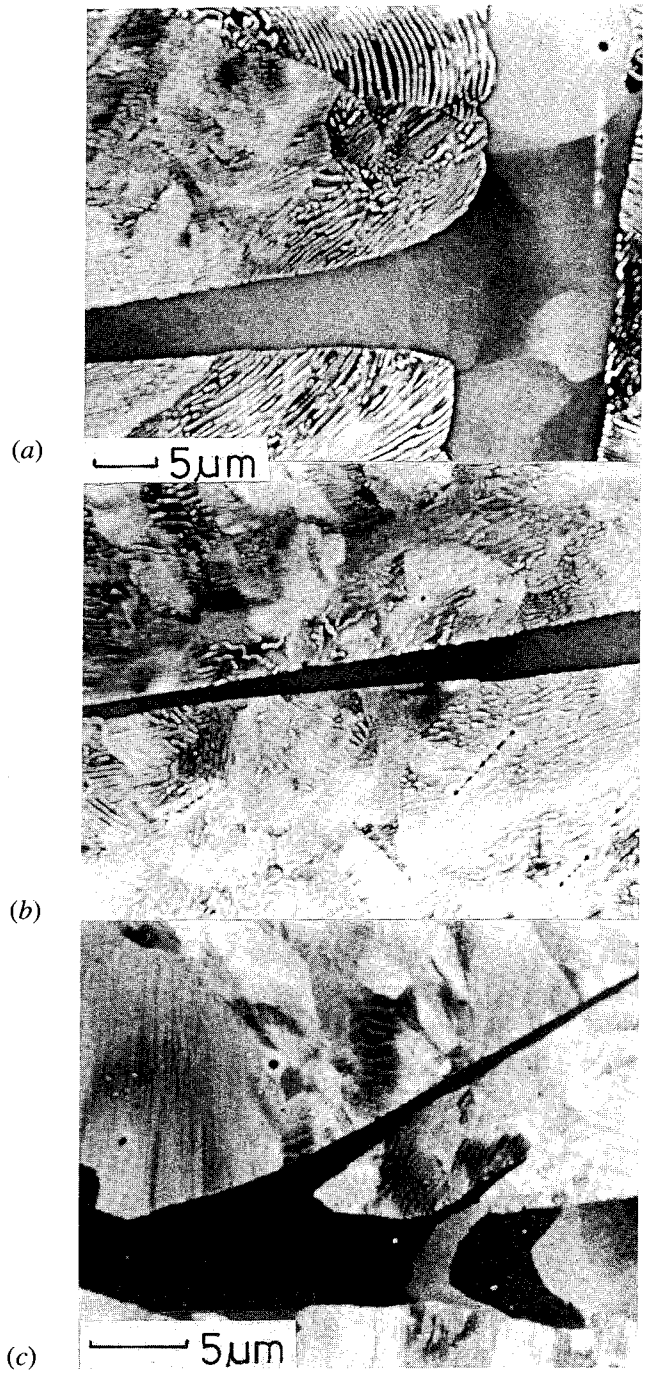


Fig. 2—Widmanstätten ferrite sideplates examined using SEM electron channeling contrast: (*a*) near to the prior austenite grain boundary; (*b*) but further along the same plate; and (*c*) in another plate. Note the strong contrast between different orientations and both the absence of boundaries aligned parallel to the length of the individual sideplates as well as of significant contrast within the sideplates.

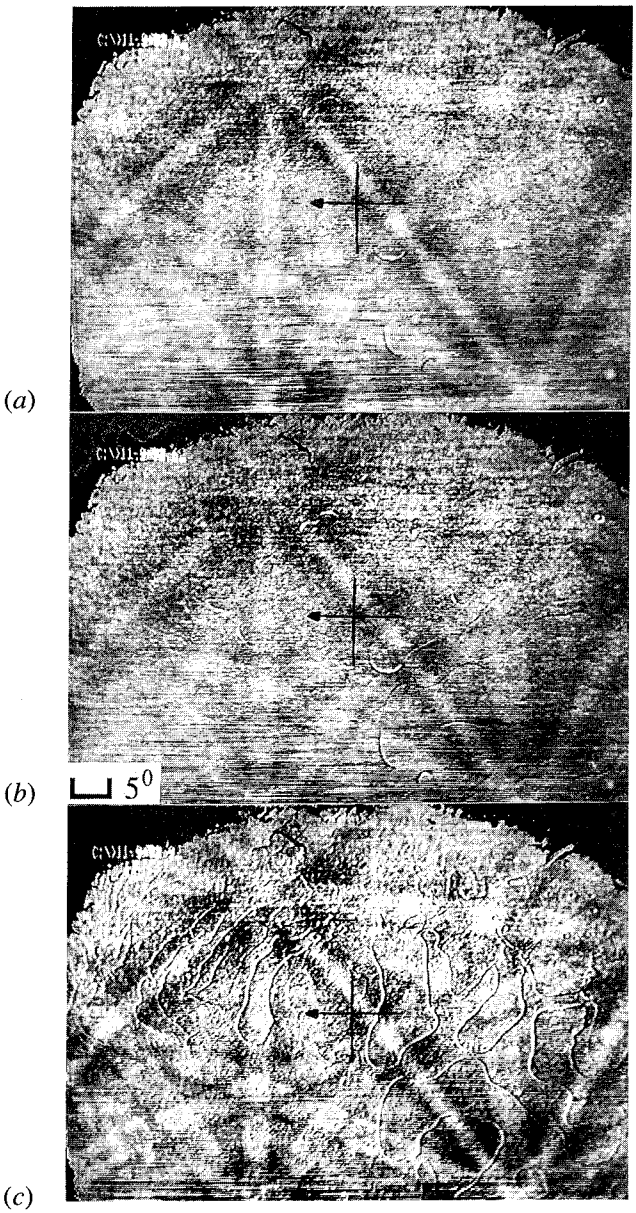


Fig. 3—EBSPs taken across an example plate nucleated in the interior of a bulk sample. Notice the absence of any detectable change in the pole positions.

orientation change within this plate, if any, must have been less than 1 deg.

3. *Surface-Relief Specimen Studies*

Figure 4 shows the range of surface reliefs and ferrite morphologies studied. Measurements across all of the plates shown were made and showed no significant orientation change. A typical example of EBSPs taken across the upper sideplate in Figure 4(b) (indicated by an arrowhead) is shown in Figure 5 to illustrate this point. Only when two or more plates were in intimate contact, with a visible boundary between the plates, as indicated by arrowheads in Figures 4(c) and (e), was EBSP evidence obtained for the presence of more than one plate. A boundary between two ferrite plates was never found near to or at the apex of a "tent" but always appeared in the "trough" between two plates whose individuality was already obvious without EBSP evidence.

A particularly clear example of a single, very well-formed tent-shaped relief is shown in Figure 6(a). The plate can be seen to be symmetrical, with the apex of the tent parallel to the long axis of the plate. An overall micrograph of the plate taken with backscattered electrons is shown in Figure 6(b). The difference in contrast

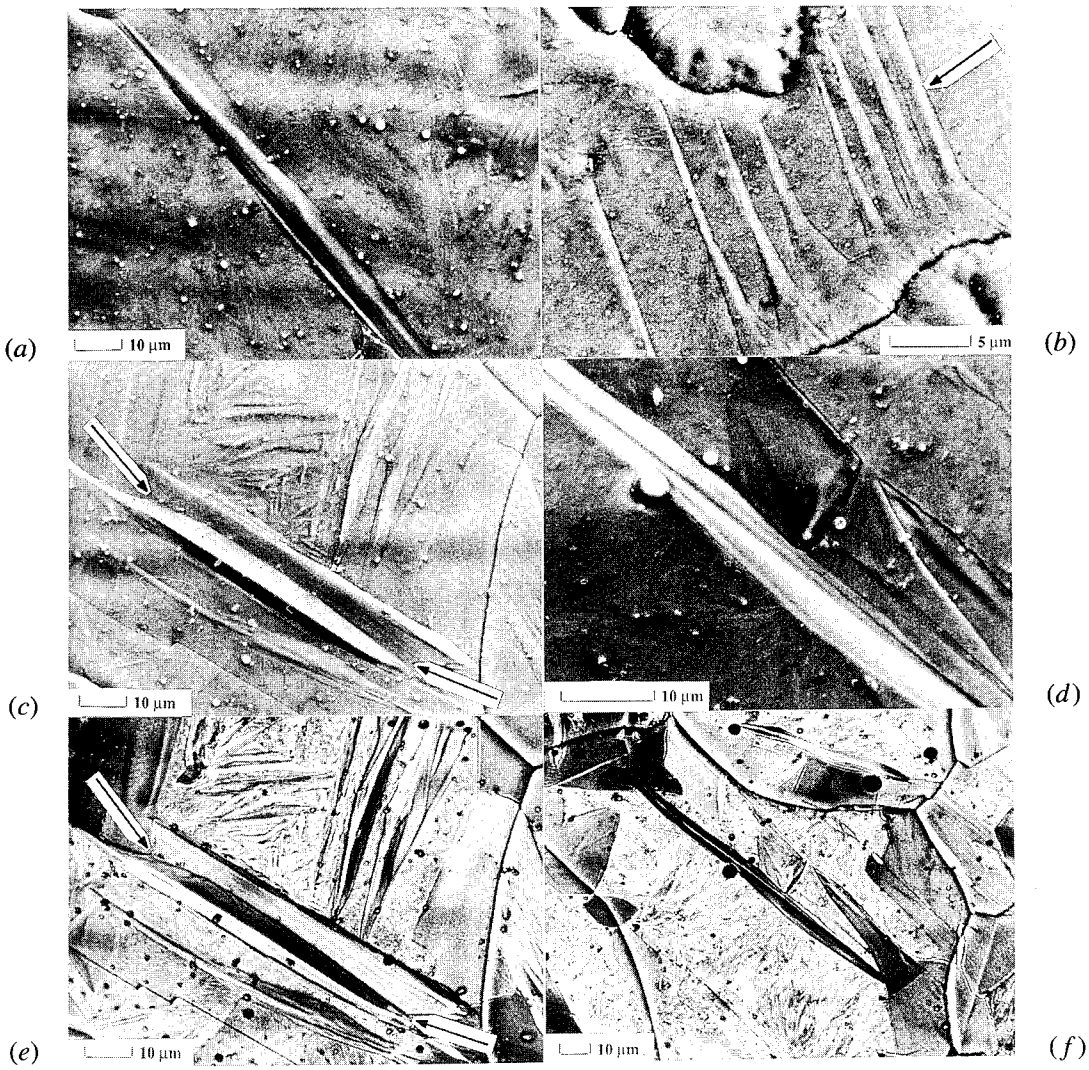


Fig. 4—Examples from the surface-relief samples showing the range of surface-relief effects observed: (*a*) secondary electron image showing isothermally formed ferrite crystals, as well as additional crystals sympathetically nucleated against these during slow cooling; (*b*) a group of eight sideplates; (*c*) two intragranular plates (a sharp change in orientation was observed across the boundary between these plates); (*d*) an isolated intragranular plate; (*e*) same as (*c*) but imaged with backscattered electrons; (*f*) same as (*d*) but imaged with backscattered electrons at slightly lower magnification.

between the two sides of this plate is due to the difference in their surface orientations resulting from the formation of the surface relief. A particular region of the plate was selected, and 12 EBSPs were taken at approximately equal intervals across the plate. Two of these patterns, taken on opposite sides of the tent, are shown in Figures 7(a) and (b). Measurements of the 12 patterns are summarized in Table II. They show the plate to be an almost perfect single crystal containing misorientations <1 deg across the plate.

The tilt angle on either side of the apex of the plate in Figure 6 was also measured. Using normal incidence, the microscope was continuously scanned for several minutes along a line perpendicular to the long axis of the plate to produce a contamination line on the surface of the sample. The sample was next tilted 45 deg about an axis parallel to this line, so that the profile of the surface, depicted by the contamination line, could be clearly seen. Measurements were then made of the projected angle of the "v" in the line of contamination on either side of the apex, from which the true tilt angle of each surface comprising the tent-shaped relief was calculated. Table III summarizes measurements of the tilt angle made at three positions normal to the long axis of the plate. The tilt angle can be seen to be approximately equal on both sides of the apex. This angle was also somewhat larger near the center of the plate, where the plate was widest.

IV. DISCUSSION

1. *The Sympathetic Nucleation Hypothesis for Tent-Shaped Reliefs*

The view that such reliefs result from the formation of two plates back-to-back,[19] presumably by sympathetic nucleation, has not been supported by the results of the present investigation. Within experimental limits of error (<1 deg), ferrite plates yielding tent-shaped reliefs have been shown to be single crystals. During the

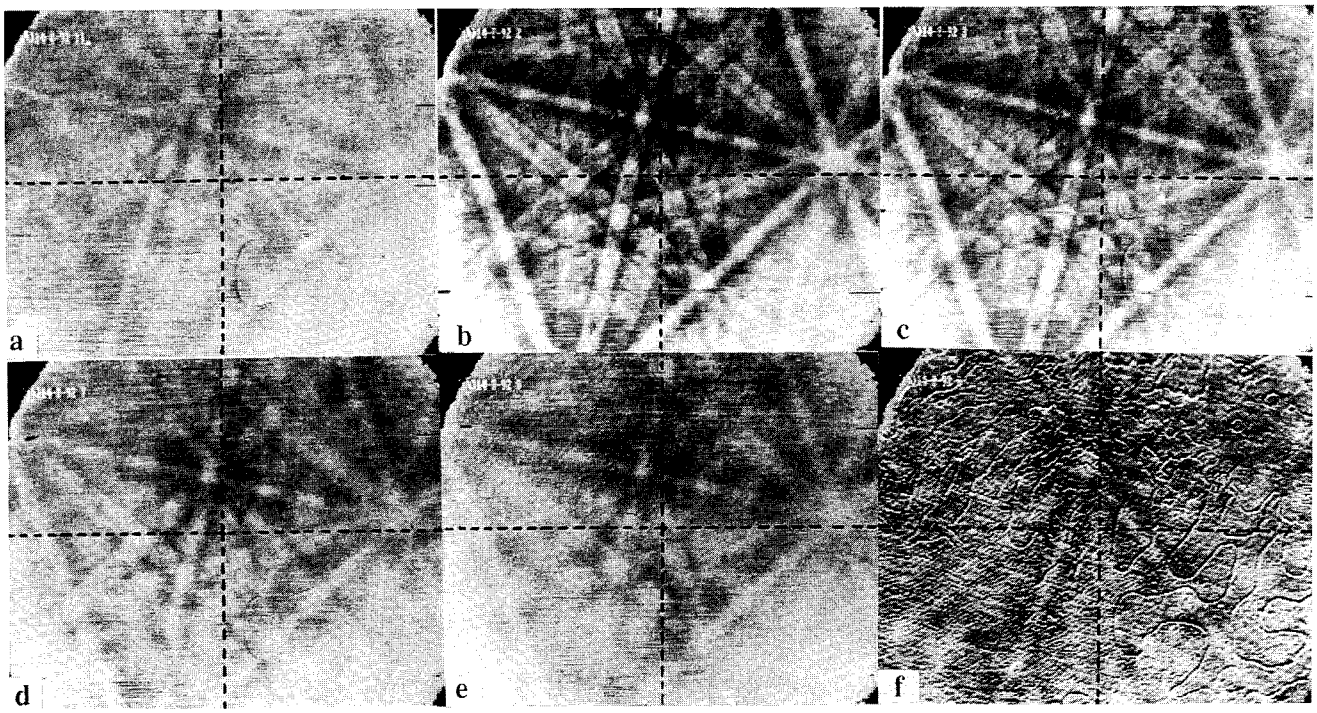


Fig. 5—EBSP measurements taken from the right-hand plate in Figure 4(b). Pattern 5(*a*) was formed in the pearlite just outside of the plate. Patterns (*b*) through (*e*) were taken at regular intervals across the plate. Figure 5(*f*) was taken in the pearlite on the other side of the plate.

proeutectoid α reaction in a hypoeutectoid Ti-Cr alloy, misorientations in this range were found with TEM in microstructures formed by sympathetic nucleation.[26] However, tent-shaped reliefs formed by proeutectoid α plates were later shown, also with TEM, to be associated only with monocrystals. No α:α boundaries more or less bisecting these plates were ever seen.

2. *Two Ferrite Plates with Mutually Accommodating Shears Hypothesis*

This hypothesis was proposed by Bhadeshia,[27] on the basis of some experimental and analytical research. Bhadeshia's specimens of an Fe-04 pct C alloy were isothermally reacted at 873 K and then slowly cooled to room temperature. A lateral boundary was found to run along the length of the ferrite plates in these specimens. This heat-treatment cycle, however, was suggested to have resulted in sympathetic nucleation at a lower temperature of a second plate at a broad face of the plate formed isothermally.[28] Reanalyzing Bhadeshia's data on his pairs of plates, we[29] concluded that the misorientations between the component plates are correctly described as small-angle boundaries, such as are commonly generated during sympathetic nucleation.[26] Dahmen[30] has criticized the theoretical basis of Bhadeshia's analysis. The results of the present investigation demonstrate experimentally that when ferrite plates are formed solely during isothermal reaction, the tent-shaped reliefs they often produce are associated only with monocrystalline plates. Back-to-back pairs of plates were not found during the present study, despite an intensive search. The misorientation required by the concept of mutual accommodation of shears postulated by Bhadeshia[27] is at least 10 deg, in sharp conflict with the present finding that any misorientation present must be <1 deg.

3. *Formation of IPS and Tent-Shaped Surface Reliefs by the Ledge Mechanism*

The proposal has been made repeatedly that an IPS surface can be produced by diffusional ledgewise growth, as long as all ledges have the same shear direction.*[9,12,14,31–34] As long as the phase transformation can be geometrically represented by a homogeneous deformation (and, if necessary, by a lattice-invariant deformation and a rigid-body rotation), this IPS relief will also have an appreciable shear component. Unlike transformation in the interior of a specimen, when a plate grows in contact with a free surface, it is likely that one of the possible shear directions will lead to a larger reduction in strain energy than the others, thereby favoring the repeated formation of that set of ledges. Particularly when both matrix and precipitate phases are substitutional, the conclusion thus has been drawn repeatedly that the presence of an IPS surface relief is proof of the presence of partial (or full) coherency, but when unsupported by other evidence is not proof of the operation of a shear transformation mechanism.[31–34]

*The term "shear" is used here in the crystallographic rather than in the transformation mechanistic sense.

Turning now to tent-shaped surface reliefs, consider first fcc/hcp transformations proceeding by ledgewise diffusional growth. Three different Shockley partial dislocations can serve as transformation dislocation ledges.[32] Hence, a surface relief associated with a single product phase plate can, in principle, have a very complex shape.

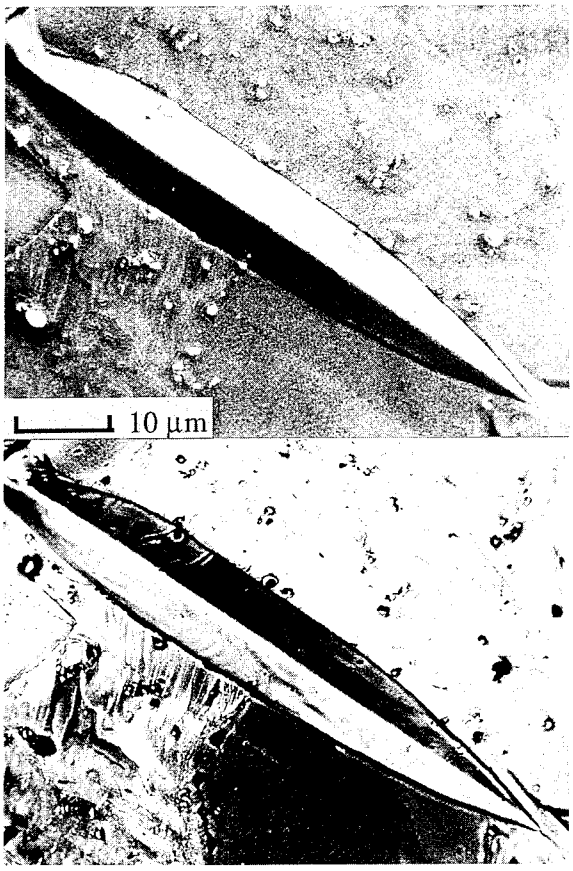


Fig. 6—A nearly perfect example of a Widmanstätten ferrite plate showing a tent-shaped surface relief, imaged with: (*a*) secondary electrons, and (*b*) backscattered electrons. Figure 6(*b*) shows no indication of a boundary, particularly in the region of the apex of the surface relief.

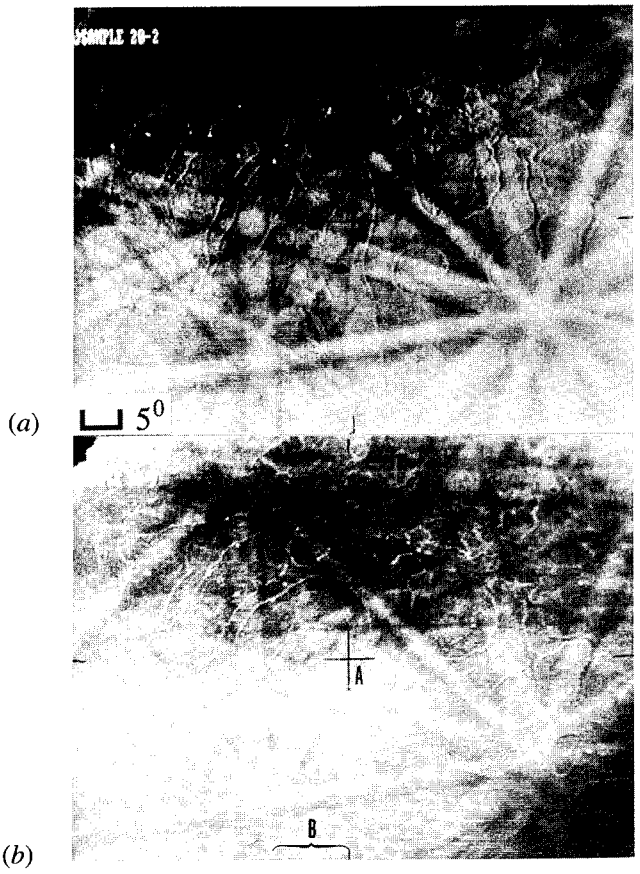


Fig. 7—Two example EBSPs (from a set of 13) taken on either side of the apex of the tent-shaped relief in Figure 6. The orientation difference between the patterns is below the ~1 deg detection limit. Measurements made from all of the patterns are presented in Table II.

Table II. Electron Backscattering Pattern Measurements

Position*	Angle *a* (deg)**	Angle *b* (deg)†
1	2.2, 2.0	10.0, 10.0
2	2.5	10.5
3	2.5	10.2
4	2.2, 2.5	10.0, 10.0
5	—	10.5
6	2.5	10.2
7	—	10.5
8	2.7	—
9	2.7	10.0
10	2.5	9.5
11	3.0	10.0

*Position is measured across the plate.
**Angle *a* is the angle from the center of the screen to the pole on the pattern indicated in Figure 7(b).
†Angle *b* is the rotational angle subtended at the center of the screen by the directions illustrated in Figure 7(b).

To date, however, only one instance of a tent-shaped relief has been reported during the formation of $AlAg_2$ plates from an α Al-Ag matrix.[32] Others observed are of the IPS type.[9]

In the case of fcc/bcc transformations, the crystallographic degeneracy operative in fcc/hcp transformations is not available. Explanation of a tent-shaped relief in terms of the PTMC would thus require that two lattice-invariant deformations operate in essentially opposing directions in conjunction with a single orientation relationship variant of the matrix and product phases. As yet, no such mechanism has been proposed, and it is doubtful if one exists. Hence, an additional ground is provided for ruling out a martensitic mechanism for the formation of a tent-shaped surface relief on a single precipitate plate.

Kinsman *et al.*[14] suggested that when the strain energy associated with an IPS becomes sufficiently large, the development of growth ledges moving in a direction making a large angle (appreciably greater than 90 deg) with respect to the direction of motion of the first-formed growth ledges would become energetically more favorable, leading to the evolution of tent-shaped reliefs. In support of this mechanism, they noted that optical interference fringes and "fiducial lines in the austenite such as scratches or terraces produced by thermal faceting" associated with tent-shaped reliefs "tend to be displaced within the plate but not in the matrix phase." Watson and McDougall[13] similarly observed that "doubly tilted regions were self accommodating and produced no

Table III. Tilt Angle Measurements

Position*	Left-Hand Side (deg)	Right-Hand Side (deg)
1	9.8	8.4
2	10.5	11.2
3	8.4	7.7

*Position is measured along the plate.

observable plastic strains in the matrix."* Table IV summarizes the data available on the ranges of tilts found among tent and IPS reliefs associated with proeutectoid ferrite plates in Fe-C alloys (Watson and McDougall[13] and the present investigation) and with proeutectoid α plates in a hypoeutectoid Ti-Cr alloy.[17] The average tilt angle for tent-shaped reliefs on ferrite plates is seen to be roughly one-half that for IPS reliefs on these plates. This ratio is somewhat less for proeutectoid α plates. These ratios suggest that simultaneous evolution of these tilts from both broad faces in the case of tents results in mutual accommodation of the accompanying strains. Watson and McDougall[13] showed that an IPS relief associated with a ferrite plate produces accommodation strains in the surrounding austenite (transformed to martensite during quenching to room temperature) primarily on one side of the plate. Taken in conjunction with the data of Table IV, this suggests that in a tent-shaped relief morphology, the IPS relief produced by each moving plate broad face would seek strain accommodation at the expense of the tilt produced by the opposing broad face. Much of the resulting reduction in tilt angle probably occurred by plastic deformation, as described for various processes by Christian.[5] Diminution with reaction time of the tilt angle of IPS reliefs generated by γ $AlAg_2$ plates was proposed to take place diffusionally because it developed during prolonged isothermal reaction,[9] unlike the very brief transformation times required to form proeutectoid ferrite plates in Fe-C alloys and low-alloy steels.

Since the tent-shaped reliefs often appeared to be fairly symmetric about their apex, the evolution of an IPS into a tent must begin at an early stage in plate thickening. If IPS relief formation does precede the evolution of a tent, the average "flow" direction for growth ledges on the uptilted side of the plate should be toward the specimen surface. When tent formation begins, the flow direction should be either reversed or replaced with one making a greater than 90 deg angle with respect to the original flow direction. While growth ledge activity seems unnecessary on the side of the plate which is not tilted during an IPS relief stage of plate thickening, ledgewise growth must begin on the untilted side when tent formation commences, and should also be directed toward the specimen interior. The circumstance that growth ledges on both sides of a ferrite plate should then grow into the specimen seems as puzzling (from the standpoint of shear strain energy reduction) as it is unavoidable. The requirement that ledge growth on the initially untilted side of a plate begin at the same time as ledges essentially reverse their growth direction on the tilted side of the plate may be fulfilled by extension of growth ledges around the (lateral) circumference of the plates when tent formation begins.

*These investigators did not determine experimentally whether or not the tent-shaped reliefs they observed associated with single ferrite crystals.[13]

Table IV. Summary of IPS and Tent-Tilt Data

Relief Morphology	α Fe-C	α Ti-Cr
Tent	8 to 11 deg (present work)	2 to 2.5 deg[17]
IPS	15 to 25 deg[13]	6 deg[17]

It should also be noted that Watson and McDougall did show micrographs of some well-formed IPS reliefs, whereas Kinsman *et al.* found this relief morphology to be largely absent. Examination of the micrographs published by Watson and McDougall makes clear that the plates that they studied were both narrower and had more nearly parallel sides than the isosceles triangle-like cross-sectioned sideplates examined by Kinsman *et al.* However, the sideplates, and perhaps also the intragranular plates, studied by Watson and McDougall may have formed during continuous cooling both before and after an interval of isothermal transformation, while the sideplates investigated by Kinsman *et al.* developed only during isothermal reaction. The higher strain energies presumably associated with the isosceles triangle-like sideplates may have made easier the formation of growth ledges at the broad faces of the plates studied by Kinsman *et al.*, especially where the sideplates joined the grain boundary ferrite allotriomorphs from which they developed.

The development of tent-shaped surface reliefs is equivalent to a change in the lattice-invariant deformation during the growth of a martensite plate. A reviewer has enquired why this is considered unlikely during a martensitic transformation but not during ledgewise diffusional growth. During growth by shear, the lattice-invariant deformation is accomplished by the glide of lattice dislocations emplaced in glissile orientation in the advancing interphase boundary. These dislocations are an intrinsic component of the matrix:martensite interface. Reorienting simultaneously the Burgers vector of all such dislocations in a given facet seems kinetically (and likely also mechanistically) infeasible. On the other hand, except in the case of fcc/hcp transformations, growth ledges do not appear to arise from an intrinsic component of the interphase boundary structure[35] during diffusional growth by the ledge mechanism. Hence, changing the Burgers vector of the growth ledges is much more readily accomplished, particularly when these ledges originate by sympathetic nucleation.

V. SUMMARY

The question of whether tent-shaped surface reliefs associated with proeutectoid ferrite plates are single crystals[14] or pairs of plates formed back-to-back by sympathetic nucleation or mutual accommodation of shears[6,27] was investigated experimentally in 0.07 and 0.50 percent C low-alloy steels. Electron channeling contrast and EBSPs formed in SEMs were employed to

study this question. All of the more than 50 ferrite sideplates studied proved to be single crystals. Any misorientation within them must have been <1 deg. No evidence of even a small-angle grain boundary along the length of the ferrite plates was found. Tent-shaped reliefs are inconsistent with a martensitic mechanism for plate growth. Earlier considerations[12,14,32] on the formation of invariant plane-strain and tent-shaped surface reliefs through a diffusion-controlled growth ledge mechanism were further extended in an effort to explain further the origins of these reliefs in diffusional phase transformations.

ACKNOWLEDGMENTS

MGH is indebted to Dr. M. Halliday and C. Cooper for permission to use and assistance with the field-emission SEM for these experiments. He would also like to express his thanks to Professor J.F. Knott, the Head of the School of Metallurgy and Materials at the University of Birmingham, for the provision of laboratory facilities and help in finding support that made attending this conference possible.

REFERENCES

1. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. TMS-AIME*, 1953, vol. 194, p. 150.
2. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 129, 128, and 224.
3. B.A. Bilby and J.W. Christian: *The Mechanism of Phase Transformations in Metals*, Institute of Metals, London, 1956, p. 121.
4. D. Hull: *Bull. Inst. Met.*, 1954, vol. 2, p. 134.
5. J.W. Christian: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 371.
6. H.M. Clark and C.M. Wayman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 59.
7. R. Smith and J.S. Bowles: *Acta Metall.*, 1960, vol. 8, p. 405.
8. C. Laird and H.I. Aaronson: *Acta Metall.*, 1969, vol. 17, p. 505.
9. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, p. 845.
10. M.H. Wu, Y. Hamada, and C.M. Wayman: *Metall. Mater. Trans. A*, in press.
11. B.C. Muddle, J.F. Nie, and G.R. Hugo: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1841-56.
12. H.I. Aaronson: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, p. 270.
13. J.D. Watson and P.G. McDougall: *Acta Metall.*, 1973, vol. 21, p. 961.
14. K.R. Kinsman, E. Eichen, and H.I. Aaronson: *Metall. Trans. A*, 1975, vol. 6A, p. 303.
15. G.R. Speich: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 353.
16. M.G. Hall, H.I. Aaronson, and G.W. Lorimer: *Scripta Metall.*, 1975, vol. 9, p. 533.
17. H.J. Lee and H.I. Aaronson: *Acta Metall.*, 1988, vol. 36, p. 787.
18. H.I. Aaronson and C. Wells: *Trans. AIME*, 1956, vol. 206, p. 1216.
19. G.R. Srinivasan and C.M. Wayman: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, p. 310.
20. C.J. Smithells: *Metals Reference Book*, 5th ed., Butterworth and London, 1976, p. 321.
21. M.G. Hall, J.K. Hulbert, and P.N. Zarucki: *Scanning*, 1991, vol. 13, p. 217.
22. D.J. Dingley, C. Harper, and S. Long: *Proc. EMAG 81*, Adam Hilger, Bristol, United Kingdom, 1981, Institute of Physics Conference Series, vol. 81, p. 63.
23. V. Randle: *Microtexture and Its Applications*, Institute of Materials, London, 1992, p. 11.
24. R.F. Mehl, C.S. Barrett, and D.W. Smith: *Trans. AIME*, 1933, vol. 105, p. 215.
25. H.I. Aaronson: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 387.
26. E.S.K. Menon and H.I. Aaronson: *Acta Metall.*, 1987, vol. 35, p. 549.
27. H.K.D.H. Bhadeshia: *Acta Metall.*, 1981, vol. 29, p. 1117.
28. H.I. Aaronson and W.T. Reynolds, Jr.: *Scripta Metall.*, 1988, vol. 22, p. 567.
29. M.G. Hall and H.I. Aaronson: *Scripta Metall. Mater.*, 1992, vol. 25, p. 2569.
30. U. Dahmen: *Scripta Metall.*, 1981, vol. 15, p. 73.
31. H.I. Aaronson, C. Laird, and K.R. Kinsman: *Phase Transformations*, ASM, Metals Park, OH, 1970, p. 313.
32. C. Laird and H.I. Aaronson: *Acta Metall.*, 1967, vol. 15, p. 73.
33. R.F. Hehemann, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1972, vol. 3, p. 1077.
34. U. Dahmen: *Scripta Metall.*, 1987, vol. 21, p. 1029.
35. H.I. Aaronson: *J. Microsc.*, 1974, vol. 102, p. 275.

Effects of Precursor Matrix Events on Subsequent Nucleation

KENNETH C. RUSSELL

A precursor transformation may actually increase the chemical driving force for subsequent nucleation of certain phases while decreasing the free energy of the system. Spinodal decomposition of Fe-C austenite has been proposed as a means of providing C-poor regions in which subsequent nucleation of ferrite or martensite is easier. However, the C-C interaction potential in austenite precludes such a reaction. Statistical fluctuations have also been proposed as a means of providing C-poor regions in austenite. However, the activity coefficient of C in austenite varies only slightly with C content, so that the scale of these fluctuations is far too small to provide attractive sites for martensite or ferrite nucleation. Statistical concentration fluctuations observed above the θ solvus in Al-Cu alloys bear a striking resemblance to Guinier Preston (GP) zone critical nuclei which will form only at temperatures that are hundreds of degrees lower. The role played by these fluctuations in subsequent critical nucleus formation is as yet unclear. A proposal that solute segregation to dislocations increases the driving force for nucleation is shown to be rendered invalid by the requirement that the chemical potentials in the system be uniform.

I. INTRODUCTION

THE free energy change, ΔG_v, on producing unit volume of product phase from a large quantity of the parent phase is an important factor in determining the rate of phase transformation. The more negative ΔG_v, the greater is the driving force which is available to overcome the surface and strain energy contributions which usually hinder solid-state phase transformations. Clearly, a phase transformation may not occur at all, except in atomic-scale fluctuations, if ΔG_v is positive.

A plot of the algebraic sum of the various contributions of the free energy of formation of a cluster of the new phase *vs* cluster size typically exhibits a maximum for a cluster size known as the critical nucleus. The (positive) free energy of forming this critical nucleus is known as ΔG^*, with the steady-state nucleation rate proportional to exp $(-\Delta G^*/kT)$.[1] Such structural defects as dislocation lines and loops and grain and interphase boundaries may catalyze nucleation processes by providing part of the free energy needed to overcome the activation barrier, thereby reducing ΔG^* and increasing the nucleation rate. Clusters resembling the new phase may be formed by statistical fluctuations even at temperatures where the parent phase is stable. It is possible, in principle, that these clusters may be larger than the critical nucleus size at lower temperatures. If so, a rapid quench would render these clusters supercritical and capable of sustained growth. The result would be a burst of transient nucleation. The steady-state nucleation rate would of course not be affected.

KENNETH C. RUSSELL, Professor, is with the Department of Materials Science and Engineering and the Department of Nuclear Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

In the simplest case, the calculation of ΔG_v and ΔG^* is straightforward.[1] This article addresses some more subtle issues of the effects of precursor events in the matrix phase on the driving force for nucleation. In particular, it considers the effects of:

(a) Spinodal decomposition;
(b) Statistical concentration fluctuations;
(c) Solute segregation to dislocation lines and interfaces; and
(d) Plastic deformation.

II. CHEMICAL DRIVING FORCE

Figure 1 is a schematic isothermal, isobaric free energy diagram plotting the molar Gibbs free energy *vs* composition for the α and β phases. The α phase of composition x^α is metastable to β phase formation. The maximum molar driving force is for formation of β of composition x^β and is given by the line *c-d*. The intersections of the extended tangent to a free energy curve at a particular composition with the two ordinates correspond to the chemical potentials of the components in the phase at that composition.[1] The distance *c-d* is seen to be the difference between the chemical potentials of A and B atoms in the two phases, multiplied by the appropriate number of atoms to give one mole of β of composition x^β. The chemical driving force for nucleation, ΔG_v, is just distance *c-d* divided by the molar volume of β.

It is thermodynamically possible to form β phase of a range of compositions around x^β, though with a smaller ΔG_v. The driving force for β precipitation from α phase of composition x^α drops to zero where the extended tangent to G_α at x^a intersects the G^β curve. These intersections are seen in Figure 1 to occur at nearly pure B and near the equiatomic composition.

At low temperatures, where atomic mobility is limited, phase transformations may only be able to occur at

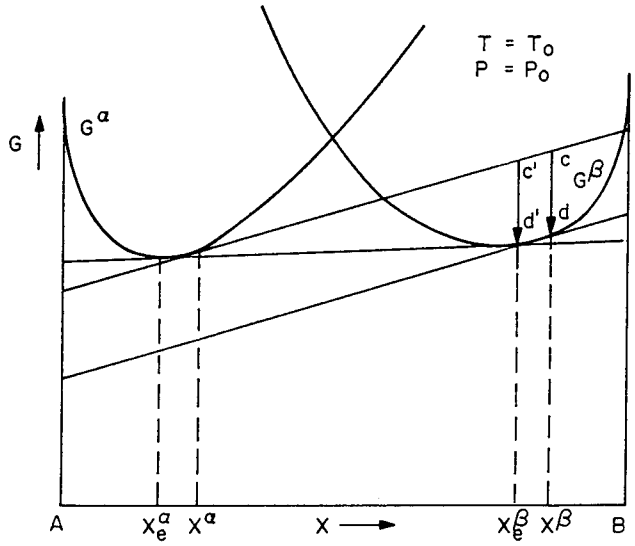


Fig. 1—Schematic isothermal, isobaric plot of free energy *vs* composition for α and β phases. The construction showing the graphical technique for obtaining ΔG_v is also shown.[1]

constant composition. Figure 1 shows that α of composition x^α cannot transform to β in this way. However, α phase richer in B may have $G^\alpha \geq G^\beta$, so that transformation at constant composition is possible.

The locus of points on a temperature-composition (T-x) diagram where $G^\alpha = G^\beta$ is known as the T_0 line. Such lines are often plotted on temperature-composition diagrams to indicate the equality of free energies and to indicate that above the line, transformation at constant composition is thermodynamically impossible.

Figure 2 plots ΔG_v vs temperature at constant composition for transformation of three Fe-C alloys to martensite. ΔG_v is seen to be most negative at the lowest temperatures and lowest carbon contents. Both T_0 and M_s (the temperature at which the first martensite forms) are seen to increase with increasing carbon content. Martensitic transformation is impossible above T_0, and in fact occurs only some 200 degrees below, due largely to strain energy and interfacial energy considerations.

A precursor transformation may have strong and sometimes unexpected effects on ΔG_v for subsequent nucleation processes. Figure 3[3] illustrates one such case, where the metastable phase is a liquid of composition, x^l. The stable state of the system is a two-phase mixture of liquid of composition x^{l1} and β phase of composition x^β. However, the extended tangent at x^l lies below G^β, so that $\Delta G_v > 0$ and the β phase cannot form directly. This situation changes after decomposition into the two liquid phases. Then the extended common tangent lies above G^β, so that $\Delta G_v < 0$ and the β phase may form. As drawn, the transformation into two liquids would occur by nucleation and growth, but a slightly more A-rich liquid would undergo spinodal decomposition.

The foregoing shows that although the precursor decomposition into two liquids lowered the free energy of the system, ΔG_v for β phase formation went from positive to negative.

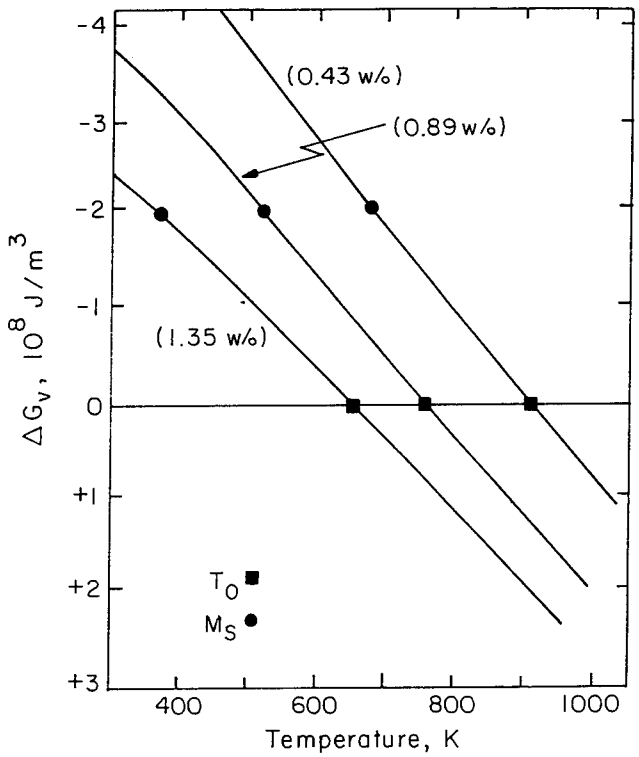


Fig. 2—Chemical driving force *vs* temperature diagram for transformation of three Fe-C austenitic alloys to martensite. T_0 is the temperature at which the austenite and martensite phases have the same free energy, and M_s is the temperature at which the first martensite is observed to form.[2]

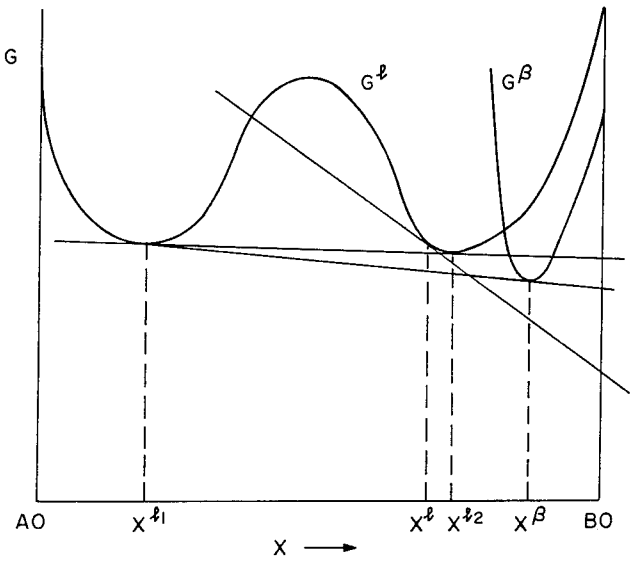


Fig. 3—Schematic free energy diagram which shows that metastable liquid of composition x^l cannot decompose directly to the equilibrium phases, liquid of composition x^{l1} and β phase of composition x^β. The driving force for β phase formation is positive until the preliminary liquid-liquid separation has taken place.[3]

III. SPINODAL DECOMPOSITION IN FE-BASED ALLOYS

Entin[4] suggested that phase transformation in Fe-C austenite might occur in small, coherent C-poor regions in which ΔG_v for martensite or ferrite nucleation is more negative than in the beginning alloy.

The enthalpy of mixing of carbon in either austenite or ferrite is strongly positive, which would tend to suggest the possibility of separation. However, coherent phase separation is possible only when the coexisting phases have the same crystal structure, or differ only slightly, as for example in one being ordered. The potential C-rich phases differ so greatly from either ferrite or austenite that coherent phase separation appears impossible.

The possibility also exists that phase separation, either by spinodal decomposition or nucleation and growth, may occur solely on the interstitial carbon sublattice, with the Fe atoms staying in place. Decomposition would then give C-rich and C-poor regions; martensite or ferrite nucleation would be favored in the latter because of a more negative ΔG_v.

The work of Nagakura *et al.*,[5] illustrates the tendency of C atoms to cluster in the ferrite matrix. Tempering was at 343 K, so lattice atom motion was impossible. Light-dark contrast in the electron micrographs clearly shows that the C atoms have segregated into very fine regions, apparently by spinodal decomposition. Figure 4 is from the same study and clearly shows the tendency for carbon to segregate. Taylor *et al.*[6] made an extensive study of aging of virgin Fe-Ni-C martensites. They observed fine-scale structural modulation along ⟨203⟩ directions. These modulations were not associated with the martensitic substructure, as would be expected for heterogeneous nucleation and were concluded to be the results of spinodal decomposition on a single octahedral interstitial sublattice, known as the "Zener sublattice."

Khachaturyan[7] had earlier presented a theory for the habit plane of coherent plates which are derived by a tetragonal lattice deformation. Taylor *et al.* found that for their case the measured habit plane was less than one degree from that calculated based on elastic anisotropy and tetragonal distortions.

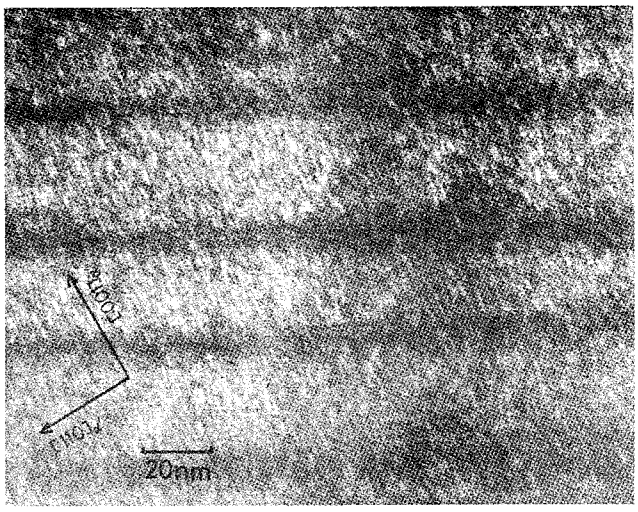


Fig. 4—High-resolution electron micrograph shows contrasting patches due to the concentration modulations in the Fe-1.39 wt pct C alloy tempered at 343 K for 1 h. (Micrograph courtesy Dr. M. Kusunoki (née Toyoshima).)

Sato *et al.*[8] analyzed octahedral defects in the body-centered cubic (bcc) lattice using an eigenstrain method in lattice theory. They found that C atoms in ferrite indeed tend to cluster, on second nearest neighbor sites. Machlin[9] had earlier derived a C-C potential in Fe which was long range and oscillatory. He predicted in ferrite a strong attraction between C-C nearest neighbors and a much weaker attraction between second nearest neighbors.

Aaronson *et al.*[10] made an extensive analysis of thermodynamics of Fe-C alloys. They found from their data analysis that in ferrite, carbon atoms attract, and that sites near a carbon atom are more likely to be occupied than other sites. Internal friction studies have shown that C atoms in ferrite tend to occur as pairs.[11]

Thus, spinodal decomposition of carbon on the Zener sublattice in ferrite has both been theoretically predicted and observed. The agreement between experiment and more recent theory is impressive.

The thermodynamic situation is very different in austenite, where spinodal decomposition on the C sublattice has been proposed. Aaronson *et al.*[10] concluded in 1966, after careful study of available data, that carbon atoms repelled each other. Chipman[12] showed in 1967 that C in austenite followed dilute solution thermodynamics fairly closely, even at the eutectoid. Any tendency toward spinodal decomposition of austenite would be indicated by the carbon activity becoming independent of carbon content. The results of both Aaronson *et al.*[10] and Chipman[12] show no sign of such a variation. To the best of the author's knowledge, no activity data exist which indicate spinodal decomposition in Fe-C austenites.

Machlin[9] predicted a strong repulsion between C-C nearest neighbors and a weak attraction between second nearest neighbors in austenite. The prediction of second nearest neighbor attraction should perhaps be viewed with some skepticism in view of the model's erroneous predictions in ferrite noted earlier.

Many studies have been made of martensite nucleation from Fe-Ni austenite. Dodd *et al.*,[13] Garner *et al.*,[14] and Garner and McCarthy[15] found that neutron or heavy ion irradiation produced a spinodal-like transformation in alloys in the Fe ~35 at. pct Ni range. Irradiation in the 750 to 950 K range gave decomposition into regions of approximately 25 pct and 50 at. pct Ni on about a 0.1- to 0.5-μm scale. Some of their results are shown in Figure 5.

The T_0 line drops rapidly with Ni content in Fe-Ni alloys, so there is a greater ΔG_v for martensite nucleation in the low Ni regions. Martensite was indeed observed to form at room temperature in the low Ni regions but not in the high Ni regions nor in the nonirradiated alloy. These results are consistent with the experiments of Kaufman and Cohen,[16] who found the M_s temperature for Fe-25 at. pct Ni alloys to lie well above room temperature, and M_s for alloys above 30 at. pct Ni to lie well below.

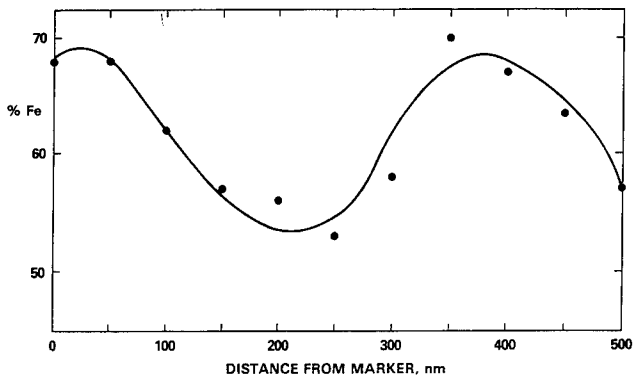


Fig. 5—Compositional oscillations of iron along a random-direction trace in Fe-35 wt pct Ni neutron irradiated to 117 displacements per atom at 898 K.[13]

The spinodal transformation observed by Garner *et al.*[14] was either induced by or greatly enhanced by irradiation. Since the transformation may have been driven by irradiation, it is not clear whether the segregation resulted in a lowering of the free energy or not. The purely thermal spinodal in the system leads to much lower amplitude concentration oscillations,[17] which would be of limited use in enhancing martensite formation.

IV. CONCENTRATION FLUCTUATIONS

A macroscopically uniform solid solution is always nonuniform on the microscopic level. Statistical fluctuations, not unlike those in gases, are always present. These ensure that although a certain region may on the average contain, say, $\bar{n}$ solute atoms, with time the actual number fluctuates around that number. From time to time, there have been suggestions that these denuded regions might provide attractive nucleation sites, especially at lower temperatures where diffusion is slow. This possibility was analyzed by Russell[18] in 1971.

Landau and Lifshitz[19] give the number of fluctuations to n from mean value $\bar{n}$ per unit volume as

$$\omega(n) = \frac{N_0}{(2\pi\bar{n})^{1/2}} \exp\left(-\frac{(n-\bar{n})^2}{\overline{\Delta n^2}}\right) \qquad [1]$$

where N_0 = atomic sites per unit volume and $\overline{\Delta n^2}$ is the mean square fluctuation, given by $\overline{\Delta n^2} = kT/(\partial\mu'/\partial n)_{T,P}$, where μ' is the chemical potential of the solute. In dilute solutions, $\overline{\Delta n^2} = \bar{n}$, and $\omega(n)$ is temperature independent.

If the system has a tendency to cluster, the fluctuations from the mean become more prevalent; the reverse obtains in systems which tend to order. Near a spinodal line, $\partial\mu'/\partial n$ approaches zero, so that $\overline{\Delta n^2}$ becomes very large, thus allowing large relative fluctuations. More sophisticated models which account for incipient surface creation would not greatly change this analysis.

The T_0 line for martensite formation decreases with solute content for Fe-C and Fe-Ni alloys, and in a number of other martensite forming systems. Martensite should thus be easier to form in a system denuded of solute. In a system obeying dilute solution thermodynamics, the number of regions per unit volume denuded or nearly denuded of solute is given by

$$\omega(n) = \frac{N_0}{(2\pi\bar{n})^{1/2}} \exp(-\bar{n}/2) \qquad [2]$$

Equation [2] indicates that in a laboratory-sized sample, one may expect to find approximately one region denuded of approximately 100 solute atoms. The chance of finding a region denuded of 200 or more atoms is effectively nil. Russell[18] utilized the thermodynamic data of Chipman[12] to show that even at the Fe-C eutectoid, austenite followed dilute solution thermodynamics closely enough that Eq. [2] was applicable.

Olson and Roitburd[20] discuss models for the martensite nucleus. At very high driving forces or in the presence of lattice softening, a nonclassical nucleus model with a diffuse interface is appropriate. Ferrous martensite forms at relatively low driving forces and without lattice softening in Fe-C austenite. According to Gibbs,[21] the small values of ΔG_v imply a large critical nucleus, so ferrous martensite probably has a classical critical nucleus composed of many thousands of atoms. Nonclassical nuclei generally tend to be even larger. Equation [2] shows that the number density of denuded regions useful as nucleation sites decreases very strongly with site size. The small regions which might be denuded of solute by fluctuations would clearly be of no assistance in nucleating large particles.

Ferrite nucleates in austenite at relatively low driving forces, therefore with large critical nuclei. The foregoing arguments thus apply to nucleation of ferrite in carbon-denuded regions of austenite.

Calculating the probability of finding a denuded region and then the free energy of forming a nucleus in the region is just a variation of calculating the free energy of forming a nucleus of a composition different from the matrix. The only advantage of appealing to denuded regions is that the regions need not be formed at the martensite nucleation temperature, where atomic mobility may be extremely low.

The usual heat treatment for precipitation consists of holding the sample at a temperature just above the solvus, followed by a quench to well below the solvus, where the precipitates form. The equilibrium precipitate may form directly in the quench (as in Cu-Co) or be preceded by a complex series of metastable phases as in Al-Cu alloys. Figure 6 shows the Al-rich end of the Al-Cu phase diagram, along with solvi for the various phases which may form.

Usually, the dilute solid solution above the solvus is visualized as being largely random, with some tendency to clustering. In this picture, most of the solute would be present as single atoms, accompanied by a smaller number of pairs and an even smaller number of trimers. Matsubara and Cohen[23] and Cohen[24] present X-ray diffraction measurements in sharp contrast with this "nearly random" picture. They found in the prototypical Al-1.7 at. pct (4 wt pct) Cu system at 793 K, just above the solvus for the equilibrium θ phase, that 70 pct of the Cu atoms are present in the form of trimer and larger clusters, as compared to 10 pct in a random solution, as

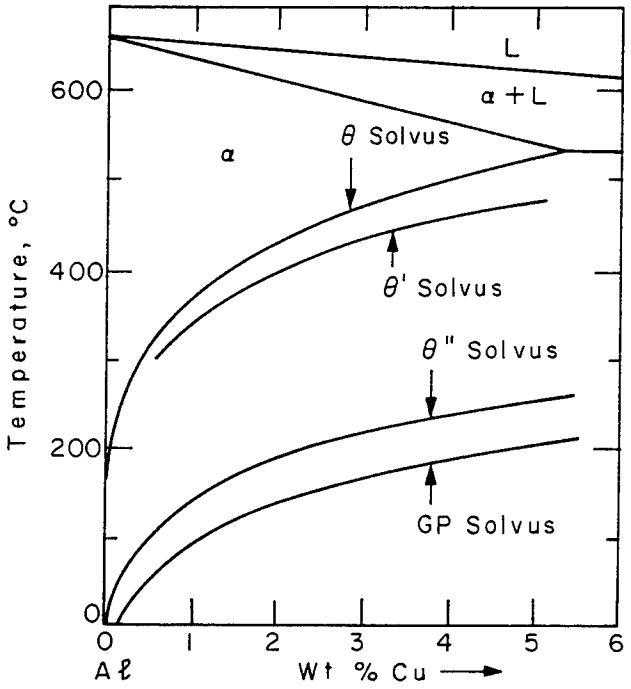


Fig. 6—Al-rich end of Al-Cu phase diagram showing equilibrium θ solvus and solvi for metastable θ' and θ'' phases and for GP zones.[22]

is assumed in the popular regular solution model.[25] In particular, 30 pct of the Cu was present in seven-atom aggregates located on two adjacent (100) planes, as compared to 0.5 pct in a random solution. The Guinier Preston (GP) zones observed after aging are also on (100) planes, have about the same number density per unit volume as the seven-atom fluctuations, and contain only some ten atoms.[26] Cohen suggested that these seven-atom aggregates present at 793 K provided much of the clustering needed for nucleation at lower temperatures and that thermally activated nucleation was unnecessary.

The Matsubara and Cohen[23] and Osamura *et al.*[26] measurements raise at least two serious questions:

(a) Why are the observed cluster densities so much greater than predicted by simple solution theory?
(b) Do these cluster densities indeed prove that the clusters stable at lower temperatures form in large part above the solvus?

Huge concentration fluctuations may occur just above the solvus near a critical point and give rise to the phase transformation on cooling the alloy just across the solvus. It should be noted that the GP zone solvus is somewhere below 473 K, over 300 K lower than the 793 K at which the clusters were observed.

Matsubara and Cohen[23] analyzed their diffraction results on Al-Cu alloys to derive pair potentials. They found an oscillatory potential which indicated a tendency for like neighbors in the first two shells and for unlike neighbors in the third and fourth shells. Though unexpected, such a potential is consistent with Harrison,[27] who shows that the Al-Al potential is long range and oscillatory.

The question remains as to whether these platelike aggregates present above the solvus will give a transient burst of super-critical nuclei upon quenching. Such clusters would, of course, not lead to steady-state nucleation. Furthermore, the packing in a cluster of n^* atoms would have to be appropriate to the critical nucleus at the lower temperature. For example, a fluctuation consisting of n^* atoms in a line would hardly constitute a critical nucleus at the lower temperature, even though the fluctuation contained the appropriate number of atoms.

Herman and Fine[28] studied the same alloy as Matsubara and Cohen[23] and found that GP zones containing many hundreds of Cu atoms reverted to the solid solution at 473 K. These observations raise the following question: why, if the zones are almost nucleated at 793 K, do they only grow at temperatures nearly 400 K lower, and why did much larger zones revert to solid solution over 300 K below the temperature at which such heavy clustering was observed? The stability of the clusters should increase with decreasing temperature, simply from LeChatelier's principle. It is reasonable to expect that if the fluctuations needed for nucleation were present in the stable alloy, nuclei would form with minimal undercooling and be stable up to just below the solvus.

V. EFFECTS OF DISLOCATIONS

Dislocation lines are known both from theory and experiment to be effective catalysts for nucleation of new phases. The effect is due to the strain energy of the particle and dislocation line together being less than the strain energy of the dislocation and particle apart.[1]

Both substitutional and interstitial impurities are known to segregate to or away from dislocations, again to reduce strain energy. It was suggested by Wu, Hamada, and Wayman[29] that such segregation involving dislocations in solute might increase ΔG_v for martensite plate nucleation.

Solute segregates to or away from dislocations to equalize the chemical potential in the system. Hirth and Lothe[30] begin with equality of chemical potential to derive their equation for the concentration of misfitting solute atoms around a dislocation

$$\frac{C}{N-C} = \frac{Co}{N-Co} \exp\left[-\frac{P\Delta V}{kT}\right]$$

where C = solute concentration;
Co = solute concentration in bulk;
N = sites per unit volume;
P = hydrostatic component of dislocation stress field; and
ΔV = difference between atomic volume of solute and matrix atom.

Concentration thus varies to account for a variable activity coefficient; activity and chemical potential do not vary. Since ΔG_v for nucleation is defined in terms of chemical potentials, it is unchanged by equilibrium segregation to dislocations. Aaron and Aaronson[31] came to a similar conclusion on the invariance of ΔG_v with equilibrium solute segregation.

The free energy of the matrix phase is increased by dislocations and other defects created by cold work,

which provides E_{cw}, the driving force for the nucleation of recrystallized grains. One might enquire as to the possibility of E_{cw} contributing to the driving force for other nucleation processes. Typical E_{cw} are small by nucleation standards. The energy is mostly stored in dislocations, which at a density of $10^{14}/m^2$ give something less than 10^6 J/m^3. In precipitation reactions, typically ΔG_v is the order of $RT/V_m \approx 10^9$ J/m^3.[1]

It has been suggested that E_{cw} might assist in martensite nucleation. This suggestion fails on two counts. First, E_{cw} is some 100 times smaller than the typical ΔG_v for martensite nucleation (Figure 2). Second, the nature of the transformation interface is such that the martensitic phase usually inherits the dislocations present in the parent phase. Changes in elastic constants and dislocation arrangement would be second-order effects.

VI. COMMENT ON SOLID SOLUTION MODELING

The very popular regular solution model assumes random mixing of atoms and a mixing enthalpy from nearest neighbor interactions. The next step in sophistication is assuming weaker interactions with second- and higher-order neighbors. These models describe solid miscibility gaps well and are often used to calculate ΔG_v for nucleation in metastable solid solutions.

This article has quoted several studies which show that a high degree of clustering exists even in stable solutions, observations sharply at odds with the predictions of simple solution models. In addition, the common assumption of the strain energy of one atom of carbon in ferrite "blocking" all nearby sites was found to give the opposite of what was observed and was calculated in a sophisticated model.

The long-range, oscillatory, interatomic potentials found to exist for C-C interactions in Fe, for Cu-Cu interactions in Al, and for Al-Al interactions in pure Al raise grievous doubts about the applicability of the regular solution model and its derivatives. It may be prudent to reduce the status of the regular solution model for metallic solutions from that of a good approximation, as it is often taken, to that of a rough approximation. Realistic calculations of solution thermodynamics may require use of more sophisticated schemes, such as the cluster-variation formalism, as discussed recently by Asta *et al.*[32]

VII. CONCLUSIONS

A precursor transformation, either by spinodal decomposition or nucleation and growth, may increase the driving force for nucleation even though the free energy of the system has decreased. Spinodal decomposition on the carbon sublattice has been both predicted and observed in Fe-C ferrite. There exists neither experimental nor theoretical evidence for such decomposition in austenite. Irradiation-induced spinodal decomposition may assist martensite formation in Fe-Ni alloys. Statistical concentration fluctuations are always present in solid solutions but are far too small to provide sites for martensite nucleation in austenite. The role of the solute clusters observed above the GP zone solvus in Al-Cu alloys in subsequent GP zone formation is not clear. The chemical driving force for nucleation is unchanged by equilibrium segregation to dislocations.

ACKNOWLEDGMENT

I am pleased to acknowledge the financial support of the conference sponsors, which made my participation possible. Partial support by the Office of Naval Research under Grant N00014-93-11213 is gratefully acknowledged. I am grateful to Professor H.I. Aaronson for suggesting the topic for this article.

REFERENCES

1. K.C. Russell: *Adv. Colloid Interface Sci.*, 1980, vol. 13, pp. 205-318.
2. L. Kaufman, S.V. Radcliffe, and M. Cohen: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, pp. 313-52.
3. J.W. Cahn: *J. Am. Ceram. Soc.*, 1969, vol. 52, p. 118.
4. R.I. Entin: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, pp. 295-311.
5. S. Nagakura, Y. Hirotsu, M. Kusunoki, T. Suzuki, and Y. Nakamura: *Metall. Trans. A*, 1983, vol. 14A, pp. 1025-1031.
6. K.A. Taylor, L. Chang, G.B. Olson, G.D.W. Smith, M. Cohen, and J.B. Vander Sande: *Metall. Trans. A*, 1989, vol. 20A, pp. 2717-37.
7. A.G. Khachaturyan: *Theory of Structural Transformations in Solids*, John Wiley & Sons, Inc., New York, NY, 1983, pp. 244-49.
8. A. Sato, Y. Watanabe, and T. Mura: *J. Phys. Chem. Solids*, 1988, vol. 49, pp. 529-40.
9. E.S. Machlin: *Trans. TMS-AIME*, 1968, vol. 242, pp. 1845-47.
10. H.I. Aaronson, H.A. Domian, and G.M. Pound: *Trans. Metall. Soc. AIME*, 1966, vol. 236, pp. 753-67.
11. D. Keefer and C.A. Wert: *J. Phys. Soc. Jpn.*, 1963, vol. 18, Suppl. III, p. 110.
12. J. Chipman: *Trans. TMS-AIME*, 1967, vol. 125, p. 236.
13. R.A. Dodd, F.A. Garner, J.-J. Kai, T. Lauritzen, and W.G. Johnson: in *Radiation-Induced Changes in Microstructure: 13th International Symposium (Part I)*, F.A. Garner, N.H. Packan, and A.S. Kumar, eds., ASTM STP 955, American Society for Testing and Materials, Philadelphia, PA, 1987, pp. 788-804.
14. F.A. Garner, H.R. Brager, and J.M. McCarthy: in *Radiation-Induced Changes in Microstructure: 13th International Symposium (Part I)*, F.A. Garner, N.H. Packan, and A.S. Kumar, eds., ASTM STP 955, American Society for Testing and Materials, Philadelphia, PA, 1987, pp. 775-87.
15. F.A. Garner and J.M. McCarthy: in *Physical Metallurgy of Controlled Expansion Invar-Type Alloys*, K.C. Russell and D.F. Smith, eds., TMS-AIME, Warrendale, PA, 1990, pp. 187-206.
16. L. Kaufman and M. Cohen: *Trans. Am. Inst. Min. Metall. Eng.*, 1956, vol. 206, p. 1393.
17. K.C. Russell and F.A. Garner: *Metall. Trans. A*, 1992, vol. 23A, pp. 1963-76.
18. K.C. Russell: *Metall. Trans.*, 1971, vol. 2, pp. 5-12.
19. L.D. Landau and E.M. Lifshitz: *Statistical Physics*, Pergamon Press Ltd., London, 1958.
20. G.B. Olson and A.L. Roitburd: in *Martensite*, G.B. Olson and W.S. Owen, eds., ASM INTERNATIONAL, Metals Park, OH, 1992, pp. 149-74.
21. J.W. Gibbs: *The Scientific Papers of J. Willard Gibbs*, Longmans, Green, and Co., 1906. Reprinted by Dover Publications, NY, 1961.
22. E. Hornbogen: *Aluminum*, 1967, vol. 43, pp. 9-15.
23. E. Matsubara and J.B. Cohen: *Acta Metall.*, 1983, vol. 31, pp. 2129-35.
24. J.B. Cohen: *Metall. Trans. A*, 1992, vol. 23A, pp. 2685-97.

25. R.A. Swalin: *Thermodynamics of Solids*, John Wiley & Sons, Inc., New York, NY, 1962.
26. K. Osamura, N. Otsuka, and Y. Murakami: *Philos. Mag. B*, 1982, vol. 45, pp. 583-99.
27. W.A. Harrison: *Electronic Structure and the Properties of Solids*, Freeman, San Francisco, CA, 1980, p. 388.
28. H. Herman and M.E. Fine: *Trans. Metall. Soc. AIME*, 1962, vol. 224, pp. 503-05.
29. M.H. Wu, Y. Hamada, and C.M. Wayman: *Metall. Mater. Trans. A*, in press.
30. J.P. Hirth and J. Lothe: *Theory of Dislocations*, McGraw-Hill, New York, NY, 1968.
31. H.B. Aaron and H.I. Aaronson: *Metall. Trans.*, 1971, vol. 2, pp. 23-37.
32. M. Asta, C. Wolverton, D. de Fontaine, and H. Dreyssé: *Phys. Rev. B*, 1991, vol. 44, pp. 4907-24.

On the Prebainitic Phenomenon in Some Alloys

M.K. KANG, Y.Q. YANG, Q.M. WEI, Q.M. YANG, and X.K. MENG

The prebainitic phenomenon in several copper alloys and ferrous alloys was investigated by means of internal friction measurement and composition analysis. The internal friction peaks of zinc or carbon atoms diffusion in β brasses or in austenite of ferrous alloys, respectively, were related to the zinc or carbon-depleted region formation within the bainitic incubation period observed by composition analysis. The possibility of bainite shear nucleation in solute atoms' depleted regions, formed by diffusing of solute atoms to a certain level, was pointed out.

I. INTRODUCTION

THE prebainitic phenomenon is one of the important subjects of bainitic transformation which has been greatly debated. In the 1930s, Kurjumov[1] pointed out that carbon-depleted austenite can form during the incubation period of steels holding at medium temperature and then the austenite transforms into ferrite. In the 1940s, a similar view was suggested by Klier and Lyman.[2] The existence of the carbon-depleted region in austenite was confirmed indirectly by the X-ray method by Entin[3] in the 1960s and with the method of ultrasonics[4] and high-temperature X-ray diffraction[5] in the 1970s. However, Aaronson *et al.*[6] thought it impossible to decompose the austenite into carbon-depleted and rich regions, *i.e.*, spinodal decomposition, according to their thermodynamic calculation. Hsu *et al.*[7] showed indirectly by means of internal friction that it is a soft mode nucleation process of bainite in the incubation period. Garwood[8] and Takezawa and Sato[9] proposed that the bainite in β brasses can form in the zinc-depleted region by shear. Feng *et al.*[10] pointed out that many defects exist in austenite and that, by Monte-Carlo simulation, the carbon-depleted region appears in austenite during the incubation period and then bainite forms by shear in that region. Although the prebainitic phenomenon is an important aspect of understanding the mechanism of bainite formation, the description of solute atoms' behavior in the bainitic transformation, the direct measurement of the composition change during bainitic incubation period, and direct observation of the nucleation site are scarce at present. Thus, it is necessary to study these phenomena in detail.

M.K. KANG, Professor, and X.K. MENG, Ph.D., are with the Department of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an, 710072, People's Republic of China. Y.Q. YANG, Q.M. WEI, and Q.M. YANG, formerly Ph.D. Candidates with Northwestern Polytechnical University, are Postdoctoral Researchers, Nanjing University, Nanjing, 21008, Beijing University of Science and Technology, Beijing, 10085, and Zhejiang University, Hangzhou, 310027, respectively, People's Republic of China.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

II. EXPERIMENT

The tested β brasses were prepared by melting high-purity Cu and Zn or Cu, Zn, and Al in a graphite crucible, respectively, and casting into ingots with diameters of 80 mm, followed by homogenizing at 1073 K for 48 hours and forging into rods of 20-mm and 10-mm diameter, respectively. The composition is Cu-42.85Zn (weight percent) and Cu-26.84Zn-4.22Al (weight percent). The specimens for internal friction measurement with a dimension of $1 \times 1 \times 60$ mm and for metallography were cut from the rods. These specimens were betanized in a molten salt bath at 1063 K for 2 minutes and then quenched into water to obtain ordered β' or a nitrite-nitrate salt bath holding at 523 K for 20 minutes (Cu-Zn) or for 4 minutes (Cu-Zn-Al) within the incubation period and, hence, also obtaining β'. No bainite was observed by using an optical microscope or a transmission electron microscope (TEM).

The composition of tested steels is shown in Table I. In addition to commercial 9CrSi, the steels were molten in a vacuum arc furnace and cast into ingots. After homogenizing the steel at 1423 K for 72 hours and then forging it into rods, specimen 2 # and 4 # for internal friction measurement were made from the rods and heat-treated in vacuum-packed quartz tubes at 1423 K for 10 hours and then quenched into water. The microstructure of so-treated specimens is austenite at room temperature. Some of the specimens austenitized in a salt bath at 1423 K for 20 minutes were then quenched in water to obtain an austenitic structure or held at 573 K for 25 minutes, at which the bainite incubation period is as long as 4 hours. No decomposition product in austempering was found using an optical microscope or a TEM. The specimens of 9CrSi for composition analysis were homogenized at 1423 K for 84 hours, followed by austenizing in a salt bath at 1223 K for 20 minutes, and were then quenched in water; the microstructure is a mixture of martensite and austenite. Other specimens were austempered at 553 K for 400 seconds, at which the bainitic incubation period is 1000 seconds,[11] and then quenched in water; this microstructure is also a mixture of martensite and austenite.

Internal friction measurements were carried out in an inverted torsion pendulum controlled by computer. The frequency range is 0.57 to 2.0 Hz. The carbon content

Table I. Chemical Composition of Tested Steels (Weight Percent)

Alloy	C	Mn	Si	Cr
2#	1.96	3.69	0.37	—
4#	1.75	3.69	0.37	—
5#	1.24	3.06	—	—
9CrSi	0.89	0.043	1.47	1.15

in steels was measured by a PHILIPS* 595 scanning

*PHILIPS and EDAX are trademarks of Philips Electronics Instruments Corp., Mahwah, NJ.

Auger Micro-Probe, and the composition change in β brasses was measured by a Hitachi H-800 TEM under transmission-scanning mode (STEM) and an EDAX* PV9100 energy dispersion spectrometer.

III. RESULTS AND ANALYSIS

A. *The Experimental Results of Internal Friction*

Figures 1 and 2 are the experimental results of internal friction for Cu-Zn and Cu-Zn-Al, respectively. Internal friction peaks appear near 473 K. They are of relaxation type, because the peak temperature increases as the frequency increases and it is independent of heating rate. The same results were obtained for specimens 2 # and 4 # of steels, as shown in Figures 3 and 4. The peak temperature (T) and associated frequency (f) are shown in Table II.

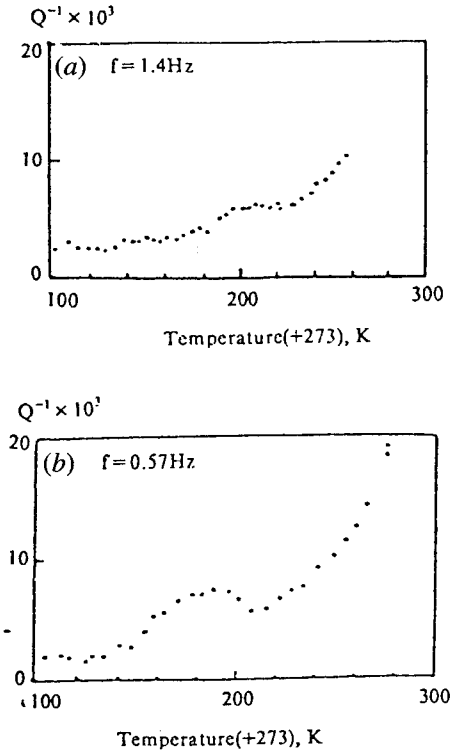


Fig. 1—The internal friction of Cu-Zn alloy.

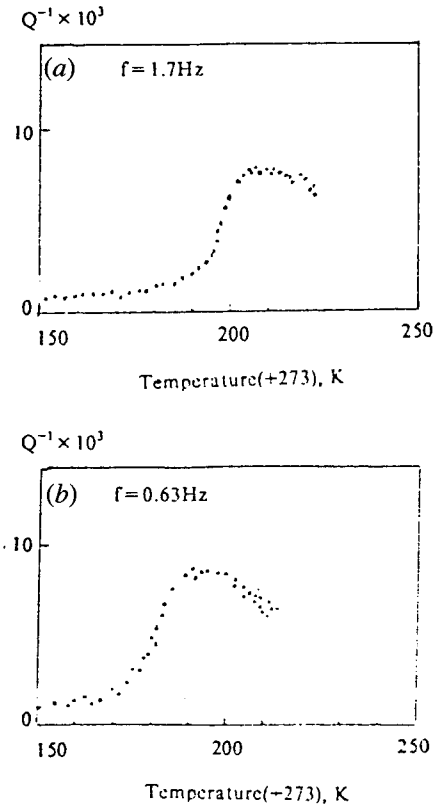


Fig. 2—The internal friction of Cu-Zn-Al alloy.

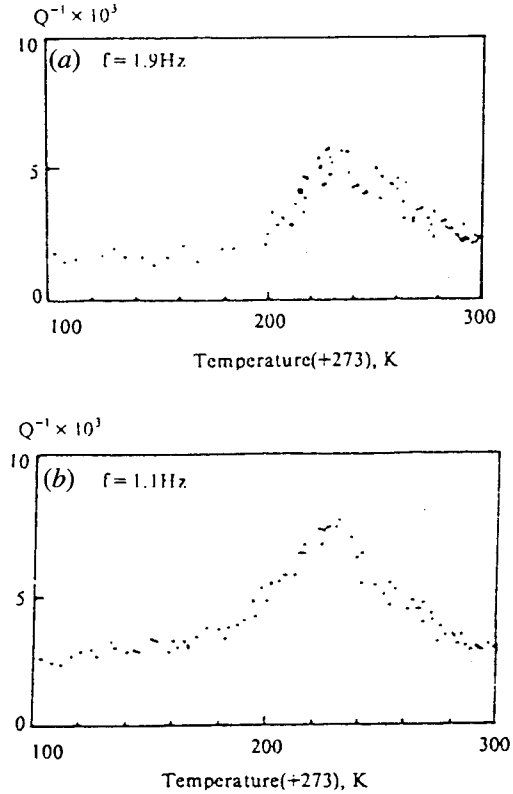


Fig. 3—The internal friction of steel 2#.

B. *The Calculation of Activation Energy*

The activation energy, Q, is calculated with the equation[12]

$$Q = \frac{RT_1T_2}{(T_2-T_1)} \ln\left(\frac{f_2}{f_1}\right) \quad [1]$$

where R is the universal gas constant, f_1 and f_2 are the frequencies, and T_1 and T_2 are peak temperatures. The activation energy of Cu-Zn alloy is 115.27 KJ/mol and that of Cu-Zn-Al alloy is 115.87 KJ/mol, which are similar to the diffusion activation energy of zinc atoms in ordered β brass.[13] It is evident that the internal friction peaks are induced by the diffusion of zinc atoms in ordered β brasses. That the activation energy of specimens 2 # and 4 # of steels is 117.21 KJ/mol, similar to the diffusion activation energy of carbon atoms in austenite,[14] shows that the internal friction peak is induced by the diffusion of carbon in austenite.[15]

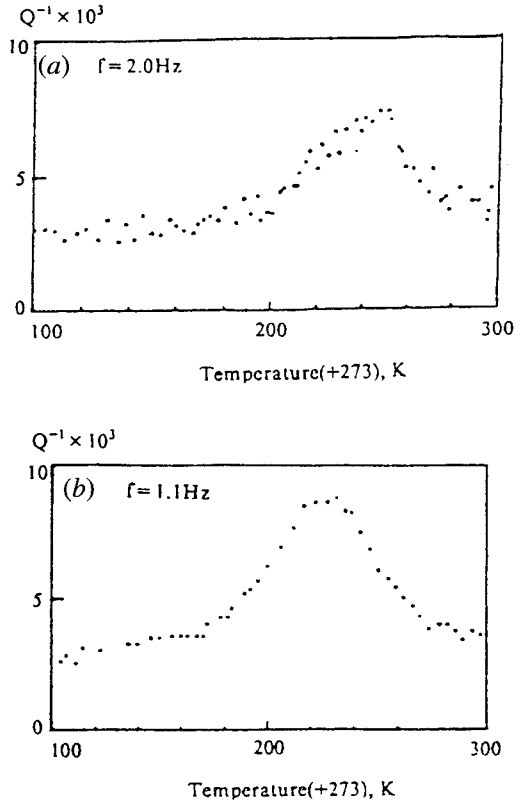


Fig. 4—The internal friction of steel 4#.

Table II. Peak Temperature (T) of Internal Friction and Frequency (f) of Some Tested Alloys

Alloy	T (K)	f (Hz)
Cu-Zn	479	1.4
	465	0.57
Cu-Zn-Al	478	1.70
	463	0.63
2#	508	1.9
	499	1.1
4#	509	2.0
	497	1.1

The width of the internal friction peaks mentioned earlier is larger than that of the Debye relaxation peak, and the shape of the peaks is not symmetric. This kind of internal friction peak is similar to that caused by segregation or precipitation of solute atoms.[16]

C. *Composition Measurement Results*

In order to show directly the composition change within the incubation period of bainite, the composition measurements and comparisons of austempered specimens with those of quenched specimens were conducted.

Figures 5 and 6 show the solute distribution of quenched and austempered specimens in Cu-Zn and Cu-Zn-Al alloys, respectively. Obviously, solute atoms distribute homogeneously in quenched specimens (Figures 5(a) and 6(a)), but there exist solute atoms in the depleted and rich regions in austempered specimens within the incubation period, as shown in Figures 5(b) and 6(b). The carbon atomic-depleted and rich regions are also formed in austenite in steels within the bainitic incubation period. The distribution curves of carbon atoms in quenched and austempered specimens are shown in Figures 7 and 8. The results indicate that there is (1) a very small fluctuation of the carbon content in quenched specimens and (2) a large fluctuation of the carbon content in austempered specimens because of the formation of the carbon-depleted region and rich region.

The results mentioned earlier show that the phenomena of solute atom depletion and richness appear because

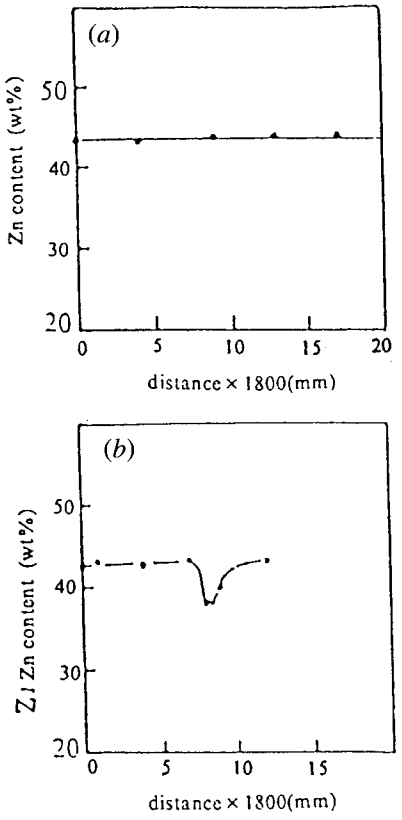


Fig. 5—The distribution of Zn content in Cu-Zn alloy: (*a*) quenched and (*b*) austempered at 523 K for 20 min.

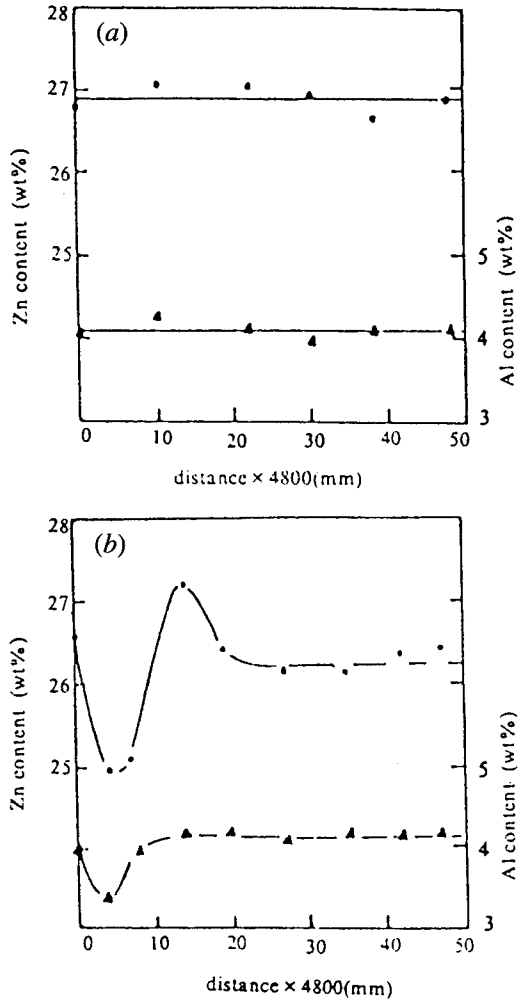


Fig. 6—The distribution of Zn and Al contents in Cu-Zn-Al alloy: (*a*) quenched and (*b*) austempered at 523 K for 4 min.

of the segregation of solute atoms caused by the interaction of solute atoms with defects and stress present in the specimens which are heated up to the incubation temperature. When the incubation period is finished, the martensite start temperature, M_s, of the solute atoms depleted region is increased, bainitic transformation will take place in the region where the driving force is larger than the resistance of bainite formation.

IV. EXAMPLES

Thermodynamic consideration of the bainitic formation in Fe-0.89 wt pct C(X_0 = 0.04) alloy at the initial stage, for example, is as shown in Figure 9.[17] It can be seen that if the carbon content of the depleted region is decreased within 0.22 to 0.03 wt pct C (X_0 = 0.01 to 0.0014) at the B_s point (773 K), the chemical driving force would range from −1300 to −1590 J/mol, which is enough to drive the phase transformation with martensite type.

Another example is shown in Figure 10. In a Cu-Zn-Al alloy, a nucleus of α_1 plate (bainite) occurs in the depleted Zn and Al regions, where solute atoms diffuse to a certain level, and with the observation *in situ* by the TEM specimen heating holder.

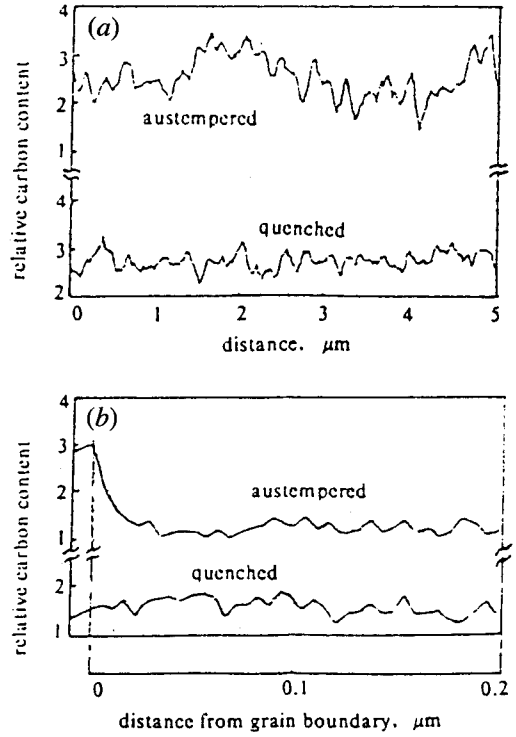


Fig. 7—The distribution of carbon content in steel 5#: (*a*) quenched and (*b*) near grain boundary.

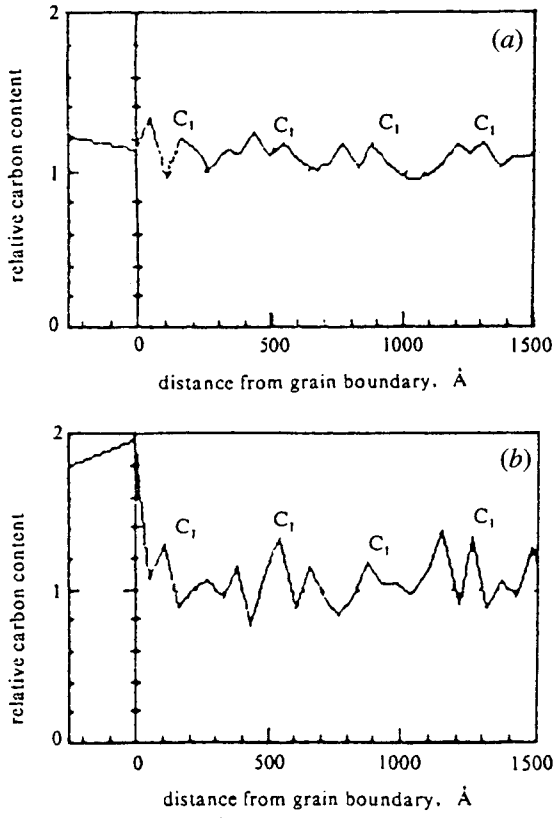


Fig. 8—The distribution of carbon near grain boundary: (*a*) quenched and (*b*) austempered at 553 K for 400 s.

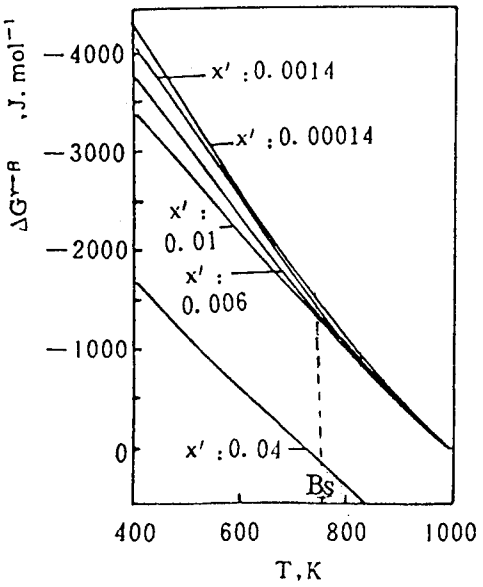


Fig. 9—Free energy changes of the bainitic transformation in an alloy Fe-O.89 wt pct C.

V. DISCUSSION

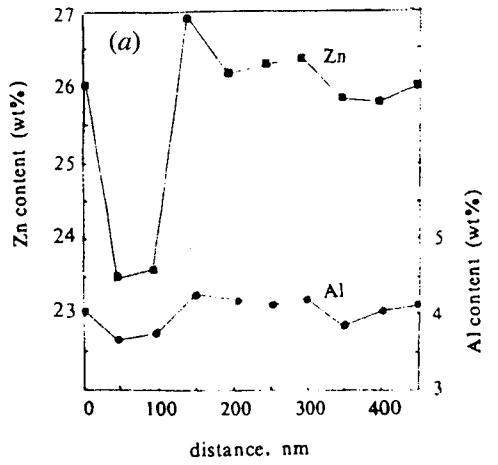
Because of the internal stress formed during quenching from high temperature, a large number of dislocations exist in the specimens. In order to reduce the free energy of the system, carbon atoms will cluster in the tensile stress region of the dislocations. For the same reason, the zinc and aluminum atoms with larger effect atomic radii also cluster in that region.[9] Therefore, the solute atoms' depleted region is formed in the compressed region of dislocations.

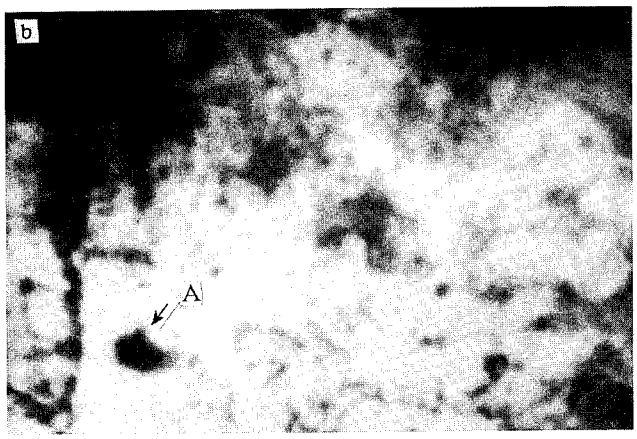
The tempering cluster of carbon in body-centered cubic or in body-centered tetragonal martensite can occur near room temperature. The diffusion constant of carbon, D_0, in α-Fe is $0.2 \times 10^{-5} m^2/s$ and that in γ-Fe is $2.0 \times 10^{-5} m^2/s$. The diffusion activation energy of carbon, Q_c, in α-Fe is 84KJ/mol and that in γ-Fe is 140 KJ/mol. From the formula,

$$D = D_0 \exp(-Q_c/RT) \qquad [2]$$

we can obtain that the diffusion coefficient of carbon in α-Fe at room temperature (293 K), D_C^α, is 2.64×10^{-21} m^2/s and that in γ-Fe at 493 K, D_C^γ, is 3.66×10^{-20} m^2/s. It is obvious that D_C^γ at 493 K is larger than D_C^α at 293 K. If the diffusion clustering of carbon in martensite can occur at room temperature, then that in austenite can also occur at 493 K from the viewpoint of kinetics due to the previous comparison. Shen *et al.*[16] even observed *in situ* by high-temperature TEM the electron diffraction effect caused by segregation or precipitation of solute atoms in a Cu-Zn-Al alloy.

After the solute atoms' depleted region formation to a certain content level, bainite can nucleate martensitically in this region because its M_s temperature increases. We think that there are two factors which change the M_s temperature: (1) the strength change of matrix caused by the temperature rising from room temperature to holding temperature and (2) the composition change in the





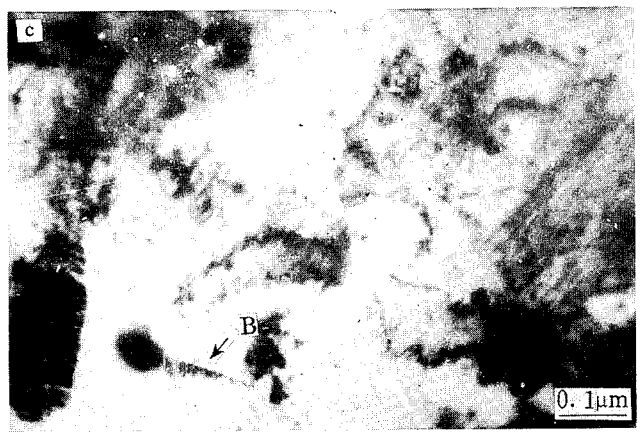


Fig. 10—Nucleation on the solute atoms' depleted region with the observation *in situ* at the temperature 493 K. (*a*) Formation of the solute atoms depleted and rich regions with austempering for 60 min within the incubation period; (*b*) stain (arrow A) at the solute atoms' depleted region before specimen heating in the heating holder of a TEM; and (*c*) specimen austempering for 21 min in the heating holder and the bainitic nucleation (arrow B) on the solute atoms' depleted region.

depleted region formed by solute atoms' diffusion. Considering the M_s change, ΔM_s^s, caused by strength change, we have

$$\Delta M_s^s = \frac{\partial M_s}{\partial \sigma} (\Delta\sigma_1 + \Delta\sigma_2) \qquad [3]$$

where $\partial M_s/\partial\sigma$ is the change rate of M_s temperature with matrix strength, $\Delta\sigma_1$ the mean strength change of matrix caused by temperature increase, and $\Delta\sigma_2$ that of the depleted region caused by segregation. Increasing holding temperature makes the matrix strength decrease and the M_s temperature increase. Although the segregation of solute atoms may increase the matrix strength in its entirety, the strength in the solute atoms' depleted region still decreases, which makes M_s temperature in that region increase. Therefore, the M_s temperature in the depleted region increases with the decrease in the strength in that region. Considering the M_s change, ΔM_s^d, caused by the composition change in the solute atoms' depleted region, we get

$$\Delta M_s^d = \frac{\partial M_s}{\partial X}\Delta X \qquad [4]$$

where $\partial M_s/\partial X$ is the change rate of M_s temperature with composition and ΔX the difference of solute content between the depleted region and the matrix. It is well known that M_s in steel increases with the decrease in carbon content, and the same tendency can be found in β brasses everywhere.[18,19,20] So, the M_s temperature in the depleted region also increases with the decrease of the solute content in that region. Obviously, both factors raise the M_s temperature in the solute atoms' depleted region. The summation of ΔM_s^s and ΔM_s^d gives the total increased quantity, ΔM_s, of M_s temperature in the depleted region:

$$\Delta M_s = \Delta M_s^s + \Delta M_s^d \qquad [5]$$

It has been indicated[21] that the strain energy of displacive transformation is related to the ratio of thickness to length of the nucleus of the product phase. The thinner and longer the nucleus, the smaller the transformation resistance. Comparing the ratio and the shear strain energy which dominates the total strain energy of bainite with those of martensite in three steels, one can find that the ratio and, therefore, the shear strain energy of bainite are smaller than those of martensite.

From the preceding discussion, it can be seen that although the M_s temperature increased by the solute atoms' depleted region is below the holding temperature when the incubation period is finished, the chemical driving force produced enough to overcome resistance of bainite by shear formation.

VI. CONCLUSION

The solute atoms' depleted region, formed by solute atoms diffusing to a certain level, appears at first in the matrix when the alloy is held at a temperature within the bainitic incubation period and, hence, the chemical driving force is enough to drive bainite formation in that region by shear after the incubation period is finished.

REFERENCES

1. G. Kurjumov: *Trans. AIME*, 1933, vol. 105, pp. 253-55.
2. E.P. Klier and T. Lyman: *Trans. AIME*, 1944, vol. 158, pp. 394-442.
3. R.I. Entin: in *Decomposition of Austenite by Diffusional Processes*, V.F. Zakay and H. I. Aaronson, eds, Interscience Publishers, London, 1962, pp. 259-311.
4. C. Lim and M. Wutting: *Acta Metall.*, 1974, vol. 22, pp. 1215-22.
5. Z. Bojarski and T. Bolds: *Acta Metall.*, 1974, vol. 22, pp. 1223-34.
6. H.I. Aaronson, H.D. Domian, and G.M. Pound: *Trans. AIME*, 1966, vol. 236, pp. 753-67.
7. G.H. Zhang, S.C. Cheng, and T.Y. Hsu: *Acta Metall. Sinica*, 1986, vol. 22, pp. A372-A376.
8. R.D. Garwood: *Metall. Trans.*, 1973, vol. 4, pp. 1375-78.
9. L. Takezawa and S. Sato: *Metall. Trans. A*, 1990, vol. 21, pp. 1541-45.
10. H. Feng, M.K. Kang, and C.X. Li: *Mater. Sci. Progr., People's Republic of China*, 1988, vol. 2, pp. 32-35.
11. M.K. Kang, D.M. Chen, S.P. Yang, and G.L. Hu: *Metall. Trans. A*, 1992, vol. 23, pp. 785-95.
12. G. Lefshiz: *Physical Properties of Metals and Alloys*, Mashigiz, Moscow, 1956, ch. 9.
13. A.B. Kuper, D. Lazarus, J.R. Manning, and C.T. Tomiruka: *Phys. Rev.*, 1956, vol. 104, pp. 1536-41.
14. Z.L. Wu and Q.M. Wang: *Acta Phys. Sinica*, 1958, vol. 14, pp. 354-59.
15. T.S. Ke and B.W. Yang: *Acta Phys. Sinica*, 1957, vol. 13, pp. 409-16.
16. H.M. Shen, Y.M. Rui, Z.F. Zhang, F. Yan, and Y.N. Wang: *3rd Chin. Conf. on Martensitic Transformations*, Qinghuangdao, China, 1990, pp. 137-39.
17. Xie Zhao Yang and Kang Mo Kuang: *Mater. Sci. Progr., People's Republic of China*, 1990, vol. 2, pp. 161-67.
18. P.E.J. Flewitt and J.M. Towner: *J. Inst. Met.*, 1967, vol. 95, pp. 273-79.
19. M. Alers: *Scripta Metall.*, 1974, vol. 8, pp. 213-18.
20. N. Mwamba and L. Delaey: *ICOMAT-82, J. Phys.*, 1982, pp. c4-639.
21. Q.M. Wei: Ph.D. Thesis, Northwestern Polytechnical University, Xian, China, 1991,, pp. 30-53.

Thermodynamics and Kinetics of the Formation of Widmanstätten Ferrite Plates in Ferrous Alloys

MASATO ENOMOTO

Experimental data on the formation of Widmanstätten/bainitic ferrite in ferrous alloys (*i.e.*, the Widmanstätten start temperature, partition of alloying elements, incomplete transformation, lengthening kinetics, *etc.*) are examined on the basis of thermodynamic calculations and kinetic analyses. A morphological change of ferrite from grain-boundary allotriomorph to Widmanstätten plate occurs well above the T_0 temperature, except in high Mn and Ni alloys, but does so in the regime of carbon diffusion control in all alloys. Under the assumption that the plate tip consists of a pair of ledges of the height equal to the tip radius, the reported lengthening kinetics of ferrite plates can be accounted for very well by the diffusion-controlled motion of these ledges in a wide range of carbon supersaturation. It is also shown that the transformation stasis (incomplete transformation) observed below the kinetically defined B_s in some iron alloys cannot be unequivocally attributed to either the completion of the precipitation of no-partitioned ferrite or the loss of the driving force for subsequent shear transformation.

I. INTRODUCTION

THE atomistic mechanism for the formation of Widmanstätten plates has attracted a considerable amount of attention. Because a well-defined surface relief often accompanies the plate formation, many authors have considered that it takes place by the cooperative motion of atoms. However, a number of observations[1–5] are available which are apparently inconsistent with that view: the surface-relief effects appear well above the T_0 temperature (upper limiting temperature for the composition invariant transformation); the growth of precipitate plates which exhibit surface reliefs is accompanied by the composition variation; plates which exhibit complex surface reliefs are actually monocrystals, *etc.*

On the other hand, numerous studies have been conducted on the thermodynamic and kinetic aspects of the plate formation, as reviewed recently.[6,7] If Widmanstätten plates are truly formed by displacive mechanism, the solute concentration has to be inherited from the matrix, at least at the instance of formation. However, comparisons with experimental data have shown that the W_s temperatures are likely to be higher than T_0 in many ferrous and nonferrous alloys.[8–11]

According to hot-stage *in situ* observations, the lengthening of Widmanstätten or bainite plates occurs continuously.[12,13,14] Thus, several theories have been proposed to analyze the lengthening of plates under the assumption that the growth is controlled by diffusion of solute in the matrix.[15–19] Whereas these theories can account well for lengthening kinetics of ferrite plates in Fe-C binary alloys,[20,21] it is widely accepted that the experimentally observed lengthening rates are much lower than theoretical predictions in alloyed steels.[22,23,24] A contrary view is offered that the lengthening rates of bainite sheaves and subunits are far greater than calculations based on the carbon-diffusion controlled growth model if the strain energy in the plates is incorporated.[25]

A peculiar transformation behavior is observed in the bainite transformation in some alloyed steels: the ferrite transformation appears to stop prematurely and resumes at prolonged holding accompanying the precipitation of carbides at ferrite:austenite interfaces. The mechanism for the appearance of such stasis has been extensively discussed.[2,6,7,26–31] However, there is still a lot of controversy on this subject.

In this report, these problems are considered further on the basis of thermodynamic calculations and analyses of plate-lengthening kinetics in iron alloys. The ledge mechanism has been invoked to analyze the thickening kinetics of precipitate plates.[3,32–34] However, in the transmission electron microscope (TEM) micrograph of intragranular α plates in a Ti-Fe alloy, an array of fine misfit dislocations and coarse but closely spaced growth ledges are observed at the leading edge of the plate.[35] Though the hot-stage *in situ* observation of growing plate tips is yet to be made, it is necessary to consider the lengthening of ferrite plates in terms of the ledge mechanism, though theories for the *disordered* plate tip have gained considerable success.[20,21]

II. RELATION AMONG THE MORPHOLOGY OF FERRITE, T_0 TEMPERATURE AND PARTITIONING OF ALLOYING ELEMENTS

A. *Summary of W_s, T_0, and the Temperature of Onset of Alloying Element Partition in Fe-C-X Ternary Alloys*

In Reference 10, the Ae_3, W_s, T_0 temperatures and the temperature of the onset of alloying element partition (designated as AEP) were plotted against the concentration of an alloying element, X, in eight Fe-0.5 at. pct

MASATO ENOMOTO, Professor, is with the Department of Materials Science, Faculty of Engineering, Ibaraki University, Hitachi, 316, Japan.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

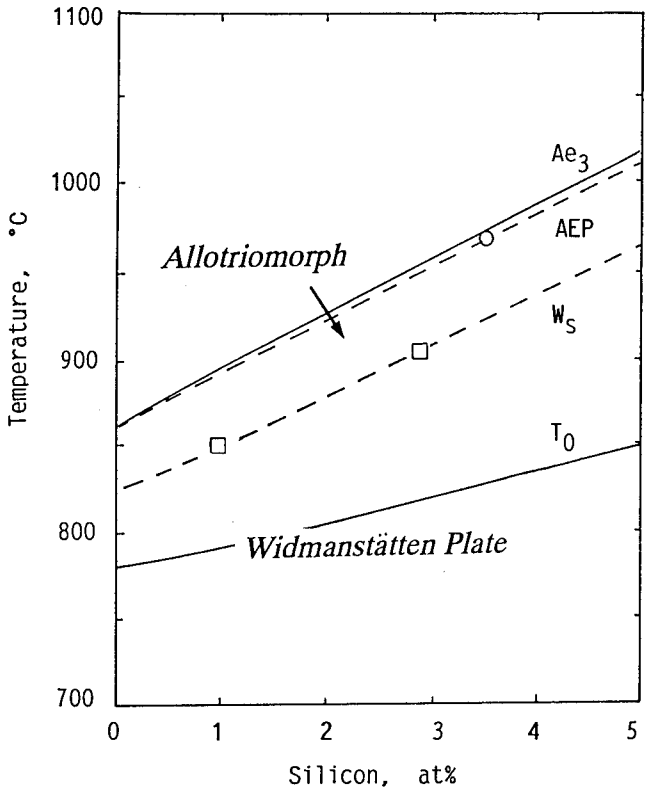


Fig. 1—Variation of Ae_3, temperature of alloying element partition (AEP), Widmanstätten start (W_s) temperature, and T_0 temperature with the silicon concentration in Fe-0.5 at. pct C-Si alloys. The bulk partition of silicon between ferrite and austenite occurs only above the AEP curve. Above W_s, only grain boundary allotriomorphs are found. They are gradually replaced by Widmanstätten plates as the temperature crosses the W_s curve and goes down further.

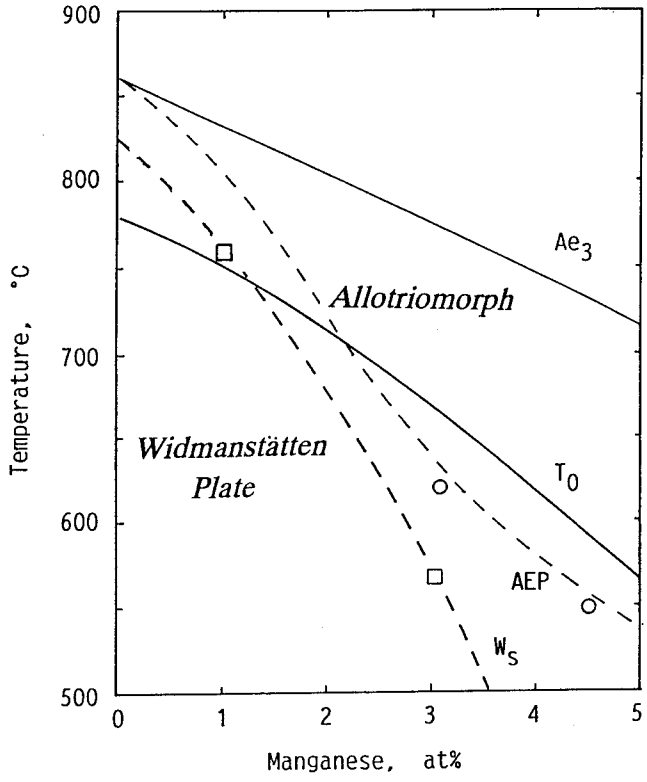


Fig. 2—Same plot as Figure 1 for Fe-0.5 at. pct C-Mn alloys.

T_0 are calculated from the Hillert-Staffanson regular solution model.[36] At lower X concentrations, the PLE to NPLE transition temperature calculated from the theory of local equilibrium[37] accounts very well for the observed temperature of the onset of partition.[38,39] However, the theory deviates from experimental results at higher Mn (and Ni) concentrations.[40] Accordingly, the AEP curves were calculated at low X concentrations and were connected somewhat arbitrarily to experimental data for the Mn concentration larger than ~3 at. pct in Figure 2. The W_s is reproduced from Aaronson *et al.*[41] and is purely experimental.

Figure 1 shows that the W_s is about 50 °C higher than T_0 in the Fe-C system, and the difference tends to increase with the Si concentration. Hence, the displacive mechanism by which the concentration of both carbon and silicon is inherited from the matrix cannot operate during the formation of Widmanstätten ferrite at these temperatures.

The partition of Si occurs only in the narrow region between the Ae_3 and AEP curves. The dominant morphology of ferrite is still the grain-boundary allotriomorph around W_s. Thus, this figure shows that allotriomorphs of no-partitioned ferrite (or paraferrite) are formed extensively in the strip between the AEP and W_s curves and also somewhat below W_s. Thus, allotriomorphic and Widmanstätten ferrite are the same in terms of the partitioning of carbon and silicon. The variation of these four temperatures with the X concentration and their relative positions in the diagram are similar for other elements, such as Cr, Cu, Co, Mo, Al, *etc.*[10]

On the other hand, the AEP and W_s curves take somewhat different paths in Fe-C-Mn and Ni alloys. As seen in Figure 2, these two curves and T_0 cross each other with increasing Mn concentration. This means that in a 3 pct Mn alloy, Mn atoms of very low diffusivity are partitioned during the formation of ferrite allotriomorph, even below T_0. Hence, T_0 can hardly be a measure of the onset of shear transformation in this alloy. Also, the diagram shows that Widmanstätten ferrite is formed only *below* T_0 when, for example, 2 at. pct Mn is contained in the alloy. It is emphasized that in view of the W_s *vs* T_0 relation in other alloys, such a situation is the exception rather than the rule. The mechanism for the suppression of Widmanstätten morphology in the presence of manganese or nickel is not known.

B. *Possible Mechanism for the Morphology Change from Allotriomorphic to Widmanstätten Ferrite*

In all alloys, the morphology change occurs in the regime of growth under carbon diffusion control, namely either NPLE or paraequilibrium (PARA) mode.* The former mode involves the diffusion of substitutional solute in the vicinity of ferrite:austenite interface and is compatible with the view that the lattice reconstruction (the breaking of old atomic bonds in the parent lattice and the formation of new ones in the lattice of the product phase) takes place in the transformation. On the other

*It is noted that the paraequilibrium $\gamma/(\alpha + \gamma)$ boundary lies between the Ae_3 and NPLE temperatures. Thus, W_s is also well below the upper limiting temperature of paraequilibrium transformation.

hand, the latter mode does not require the reconstruction of substitutional lattice. However, the loss and regaining of bonds occurs, at least with respect to carbon. It is often considered that the reconstructive mechanism which involves the substitutional lattice operates during the formation of grain-boundary allotriomorphs. On this view, it is possible that the switching from NPLE to PARA mode of growth corresponds to the morphology change. However, the diffusion distance of alloying elements associated with the formation of allotriomorphs of no-partitioned ferrite is often estimated to be much smaller than the lattice spacing of austenite.[42]

Hence, it is necessary to explain the major morphology change in the regime of carbon diffusion control. One possibility is the Mullins-Sekerka instability theory.[43] In fact, the spacing of secondary Widmanstätten ferrite sideplates in Fe-C alloys was shown to be consistent with the considerations based on the instability theory.[44,45] A simple explanation of morphology change can be offered in terms of the ledge mechanism.[46] Figure 3 shows the schematic illustration of an early stage of ferrite formation. A small area of austenite is transformed from face-centered cubic (fcc) to body-centered cubic (bcc) structure and forms a nucleus of ferrite in the grain-boundary region. At a certain undercooling, the formation of nuclei may be triggered by defect structures of the parent lattice (*i.e.*, grain boundaries and possibly special sites on them). It is very natural to consider that ledges exist at the nucleus:matrix interface, at least by the time the nucleus starts to grow. The carbon partition is likely to occur between the nucleus and the matrix, as long as the area becomes sufficiently large to be perceived as the bulk bcc structure. Hence, as each ledge starts to move, the carbon diffusion field begins to build up around the riser. At small supersaturations, the velocity of ledges is expected to be small. Then, the overlap of diffusion field of ledges occurs extensively, and hence, ledges stay near the grain boundary. In contrast, at larger supersaturations, several first-formed ledges can move fast and form a plate-like protuberance. Late-nucleated ledges see a somewhat smaller supersaturation and are likely to form allotriomorphs. According to the computer simulations previously conducted,[46,47] the critical supersaturation at which the motion of ledges switches from the low to high supersaturation regime depends on the ledge nucleation rate, the distribution of nucleation sites on the grain boundary, *etc.* Hence, no major change of transformation mechanism, such as reconstructive to displacive, is envisaged.

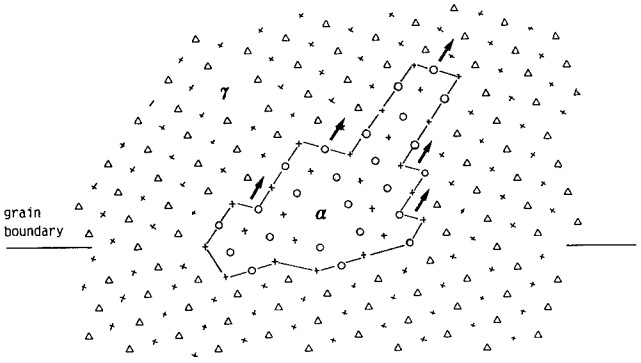


Fig. 3—Schematic illustration of early stages of ferrite formation in the grain-boundary region.

Widmanstätten plates grow invariably in the matrix grain which has a definite orientation relationship with the precipitate plate.[48,49] It is tentatively proposed that in such a grain, ledges of a specific crystallographic orientation are preferentially formed and migrate in the same orientation to form a protuberance. On the other hand, in the grain having an irrational orientation relationship, it is possible that ledges can be formed in a number of orientations and, on account of the severe overlap of diffusion field, behave as in the effectively low supersaturation regime.

III. TRANSFORMATION BEHAVIOR BELOW KINETICALLY DEFINED-B_s

A TTT curve for the initiation of ferrite transformation exhibits a distinct bay in iron alloys containing carbide-forming alloying elements. The upper temperature limit of the lower C curve, or more conveniently the bay temperature, is defined as the kinetic B_s. A most notable feature of transformation in these alloys is the occurrence of stasis below the kinetic B_s, a phenomenon called incomplete transformation.[2,50] In fact, the fraction transformed when stasis occurred was shown to be much less than the fraction expected from the metastable equilibrium between no-partitioned ferrite and carbon-enriched austenite.[2,10] Accordingly, it is considered by some authors[26,27] that the stasis appears when carbon is enriched in austenite to the T_0 composition at the reaction temperature and thus, the driving force for shear transformation is diminished to zero.

It is possible to test this supposition by calculating the ferrite fraction at transformation stasis, applying the lever rule to the T_0 composition curve. Obviously, the metastable equilibrium fraction of no-partitioned ferrite can be calculated from the lever rule with respect to the extrapolated paraequilibrium $\gamma/(\alpha + \gamma)$ boundary.[10] Figures 4 and 5 compare the results with experiment in Fe-0.13 wt pct C-2.99 wt pct Cr[30] and Fe-0.064 wt pct C-1.80 wt pct Mo[31] alloys (curves of 0 J/mol). It is readily seen that not only the metastable equilibrium fraction (designated as f_p), but also the T_0 fraction (f_t) disagree with experiment. Previously, the shear strain energy attending the formation of Widmanstätten ferrite was calculated to be larger than ~1 kJ/mol from the Eshelby's theory, and this leads to the decrease in the Ae_3 and T_0 temperatures more than 300 °C.[51] Here, instead, the amount of strain energy necessary to get agreement with the measured ferrite fractions at stasis are evaluated. The results are shown in Figures 4 and 5. Approximately 600 and 900 J/mol are necessary to fit f_p and f_t curves, respectively, to the experimental data in these alloys. The expression for the free energy of iron proposed by Ågren[52] was used in the calculation. It gives values of free energy little different from the Orr and Chipman's data.[53] However, the Kaufman *et al.* data[54] are smaller by 200 ~ 300 J/mol than these in

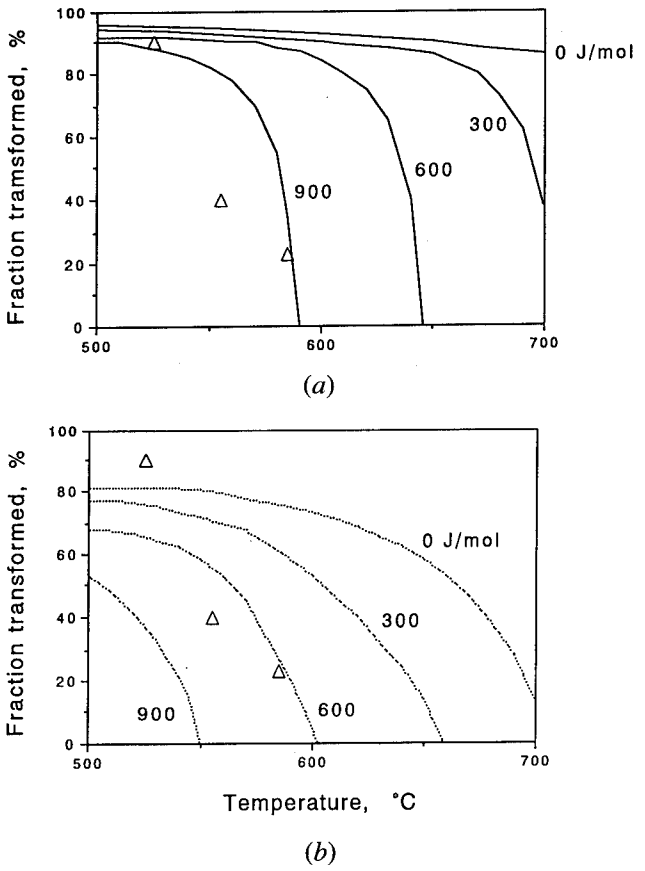


Fig. 4—Comparison of the measured fraction transformed at transformation stasis and calculated fraction of ferrite in an Fe-0.13 wt pct C-2.99 wt pct Cr alloy. (*a*) f_p and (*b*) f, are, respectively, the metastable equilibrium and T_0 fractions of ferrite.

this temperature range.* In either case, it is difficult to ascertain which value represents more closely the strain energy in bainitic ferrite, and thus, whether the stasis is due to the completion of the precipitation of no-partitioned ferrite or to the loss of the driving force for shear transformation.

*If Kaufman's data[54] were used in the calculation of T_0 in Figures 1 and 2, T_0 becomes even lower than those shown in the figures.

It is noted that the transformation stasis is not commonly observed in all alloys.[55] Even in the same alloy system, whether or not stasis occurs depends upon the alloy composition.[30] Hence, though the strain energy may play some role in this peculiar transformation behavior, it seems more reasonable to attribute this phenomenon to the characteristics of alloying elements (*e.g.*, chemical affinity with carbon in austenite) and the kinetics of carbide formation at ferrite:austenite boundaries.

IV. LEDGEWISE LENGTHENING OF WIDMANSTÄTTEN FERRITE PLATES

A. *Previous Theories of Plate-Lengthening Rate*

The studies of plate-lengthening kinetics initiated by Zener[15] are based on the assumption that the plate tip and the broad face are surrounded by disordered boundaries; that is, the atom attachment is allowed at all lattice sites of the boundary. Table I summarizes the features of theories developed thereafter and utilized in the discussion of the lengthening of Widmanstätten plates.[16–19] With the exception of Zener-Hillert (Z-H),[17] the plate shape is assumed to be a parabolic cylinder.* Also, the solute concentration is assumed to be constant along the plate interface, though it is modified in Trivedi[19] to take into account the effects of capillarity and interfacial reaction kinetics. Whereas Ivantsov[16] only presented the relation between the solute supersaturation and the lengthening rate, Z-H and Trivedi assumed the maximum growth rate principle to determine the lengthening rate and the plate tip radius simultaneously at a given supersaturation. Figure 6 shows the variation of $v\rho/D$ with solute supersaturation calculated from these theories, where v is the lengthening rate, ρ, the plate tip radius, and D, the diffusivity of solute. The ratio ρ/ρ_c (ρ_c is the critical radius at which the lengthening rate is zero) is not the same in Z-H and Trivedi and depends on the supersaturation in the latter theory. The interface kinetic coefficient, μ_0, is assumed to be infinite for Trivedi's curve shown here.

*Horvay-Cahn[18] treated the problem in a more general form. When one of the two principal radii of curvature of elliptical paraboloid is infinitely large, the solution becomes equivalent to that of Ivantsov.[16]

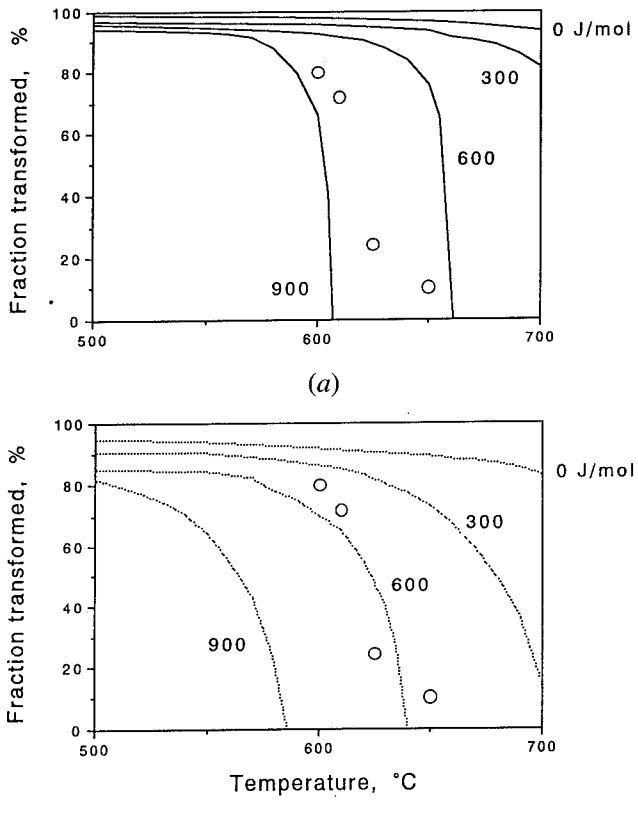


Fig. 5—Same comparison as Figure 4 for an Fe-0.064 wt pct C-1.80 wt pct Mo alloy.

Table I. Theories of Plate-Lengthening Rate

Author(s)	Year	Assumed Plate Shape	Boundary Condition	Maximum Growth Rate Principle	Relation Between Ω and $p = V_p/2D$
Ivantsov	1947	parabolic cylinder	constant concentration	—	$\Omega = \sqrt{(\pi p)} \exp(p)$ erfc $(\sqrt{p})$
Zener-Hillert	1957	circular cylinder	constant concentration along the tip	$\rho/\rho_c = 2$	$p = \Omega/8(1 - \Omega)$
Horvay-Cahn	1961	elliptical paraboloid $(\rho_1/\rho_2 = 0)$	constant concentration	—	same as Ivantsov
Trivedi	1970	parabolic cylinder	concentration varies with boundary curvature and interfacial reaction kinetics	ρ/ρ_c decreases with Ω, approaches 2 as $\Omega \to 1$ for $q = \infty$	$\Omega = \sqrt{(\pi p)} \exp(p)$ erfc $(\sqrt{p}) \times [1 + (V/V_c)\,\Omega S_1(p) + (\rho_c/\rho)\,\Omega S_2(p)]$ and $\partial V/\partial\rho = 0$
Atkinson	1981	ledge	constant concentration gradient along the riser, no flux on the terrace	—	numerically obtained

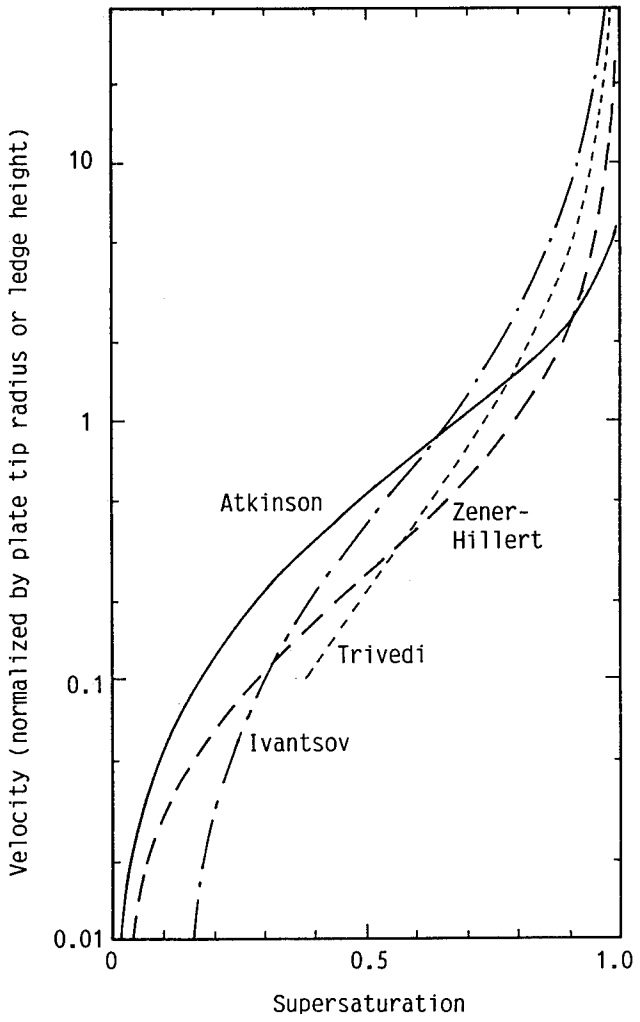


Fig. 6—Variation of the dimensionless velocity of a plate tip (lengthening rate, $v\rho/D$) with solute supersaturation calculated from theories included in Table I.

B. *Theory of Growth of Single Ledge*

The interfacial structure of the leading edge of Widmanstätten ferrite plates in iron alloys has not been studied. However, it is reasonable to assume that the edge comprises growth ledges. The height of ledges observed in the Ti-Fe alloy[35] is 20 to 30 nm. If the ledges are 1 or 2 atm high, the plate tip is essentially disordered, and the diffusion field and lengthening kinetics may be described adequately by previous theories. However, in view of the fact that the risers are significantly larger than that, it is necessary to consider the plate lengthening in terms of the ledge mechanism. As a simplest case, the plate tip is assumed to consist of a pair of single ledges, the height being equal to the observed tip radius (Figure 7). Under this assumption, a finite difference computer simulation of plate lengthening has been conducted and good agreement was obtained with experiment data in Fe-C alloys.[47]

At higher supersaturations, it is convenient to use the solution obtained from the rigorous treatment of the growth of a single ledge by Atkinson.[56] The features of this theory are included in Table I. In contrast to other theories, the concentration gradient is assumed to be constant along the riser. In the original article,[56] the $p(=v\rho/2D)$ *vs* Ω(solute supersaturation) relation is shown only up to $\Omega \sim 0.88$. Since the experimental data to be compared were obtained at even higher supersaturations, the numerical analysis to obtain the p *vs* Ω relation was extended to $\Omega \geq 0.9$. In the analysis, $\alpha(p)$, the inverse of the concentration gradient at the riser, g, is expressed as

$$\alpha(p) = -1/g = -2 \sum_{n=0}^{\infty} f_n(p)$$

where f_n represents the coefficients of the Fourier expansion of the (normalized) solute concentration in the

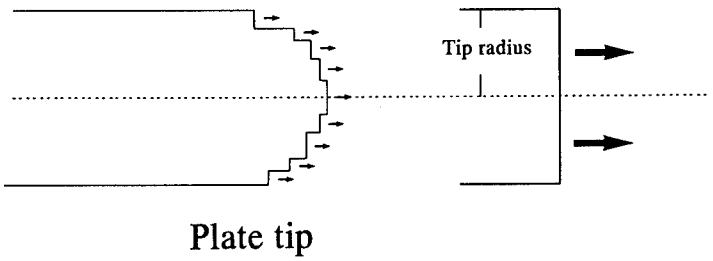


Fig. 7—Schematic illustration of the plate tip comprising an array of growth ledges.

matrix along the riser, *i.e.*, $U(0, y) = 2g\Sigma f_n \cos(n\pi y)$, ($0 \leq y \leq 1$ is the coordinate along the riser). The f_n terms are obtained solving simultaneous equations, which are derived from the functional equation derived from the boundary condition at the riser. The set of equations is truncated to $n \leq 10$. The velocity *vs.* supersaturation curve is calculated from the relationship

$$\Omega = 2p\alpha(p)$$

The result is presented in Figure 6. It is seen that whereas the ledgewise lengthening is faster at low supersaturations, it becomes remarkably smaller than the theories of the growth of a disordered plate tip at high supersaturations.

Figure 8 shows a schematic illustration of the diffusion flux and concentration contours around the plate tip in the disordered and ledgewise lengthening of a plate. As a representative of the former growth, Ivantsov's model is considered. In the disordered growth, solute attachment (or desorption) is permitted all along the plate interface. As a consequence, at low supersaturations, the diffusion flux emanated from the side of the plate can influence the growth of the tip significantly. Since no atom attachment is allowed on the terrace of a ledge, the riser can move faster than the tip of a parabolic cylinder.

At high supersaturations, the tip moves fast, and thus, the influence of diffusion flux from the side part diminishes. As a result, the tip of a parabolic cylinder takes full advantage of the so-called point effect of diffusion. In contrast, the diffusion flux may not become so high in front of a flat ledge riser. This accounts for a slower lengthening rate predicted from the ledgewise growth. In fact, the difference between the two models becomes more than an order of magnitude at very high supersaturations.

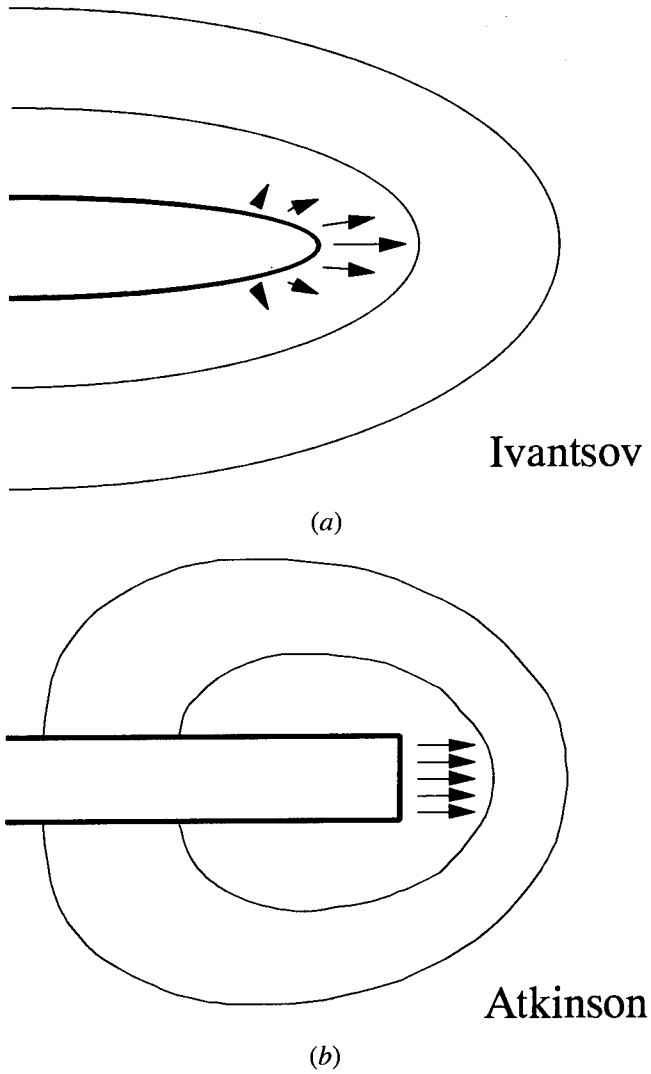


Fig. 8—Schematic illustration of the diffusion flux of solute and concentration contours around the plate tip for (*a*) disordered growth and (*b*) ledgewise lengthening.

C. *Comparison with Experiment*

A considerable amount of experimental data are available on the lengthening of Widmanstätten ferrite plates in iron alloys.[13,21–24] Trivedi's theory was shown to give the best account of the lengthening rates of Widmanstätten and bainite plates in Fe-C binary alloys.[20] On the other hand, in high-purity Fe-C-Ni alloys, the measured lengthening rates of bainite plates were reported to be much lower than the calculations using the Horvay-Cahn analysis.[23] A kind of drag effect on the interface motion by nickel atoms was considered to be responsible for the discrepancy.[20] Later on, the data were reanalyzed using Trivedi incorporating the interfacial reaction kinetics to obtain an agreement within a factor of 5.[57] Here, these data are re-examined in terms of the ledgewise lengthening.

Figures 9 and 10 show the comparison of all theories with the Simonen et al.[21] data on the lengthening of Widmanstätten ferrite at 700 °C and 450 °C, respectively. The same comparison as Figure 9 was made with respect to the Z-H and Trivedi theories at a slightly different temperature.[20] Except these two theories in which

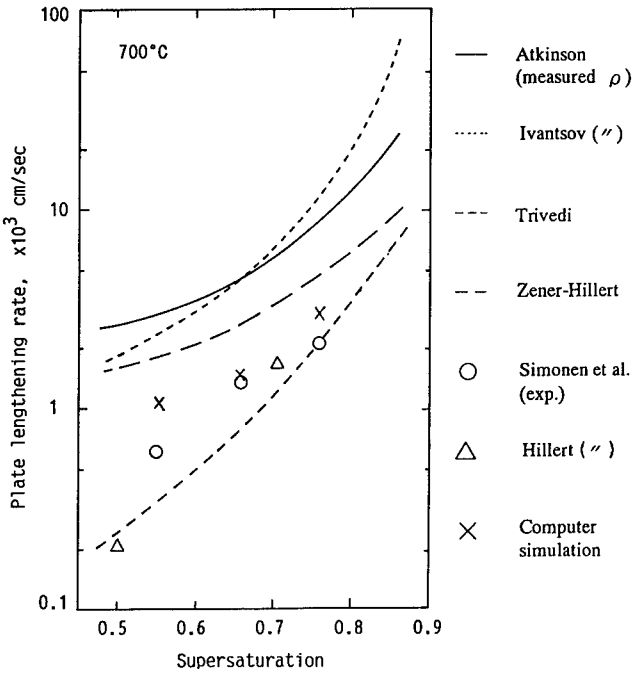


Fig. 9—Comparison of Simonen et al.'s data on the lengthening rate of Widmanstätten ferrite in Fe-C alloys at 700 °C[21] with calculation. The results of computer simulation (cross)[47] are included.

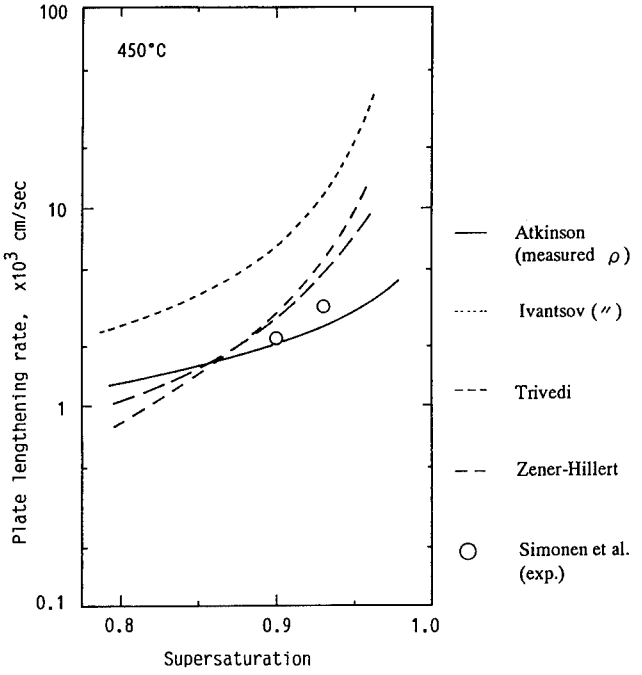


Fig. 10—Same comparison as Figure 9 at 450 °C.

the tip radius is automatically obtained, the measured value of the tip radius was used in the calculation from the Ivantsov and Atkinson theories. The carbon diffusivity was averaged with respect to the carbon concentration using the Ågren's expression for the carbon diffusivity in austenite.[58] Due to the diffusion geometry around the plate tip, the diffusivity near the interface might play a more significant role than if taken into account by averaging. The diffusivity evaluated at the boundary concentration was approximately two to five times larger than the averaged value as the temperature varied from 700 °C to 400 °C.

In the computer simulations previously conducted,[47] the ledge nucleation rate was adjusted to produce the reported ledge spacing on the broad face of plates.[3] The difference between the computer simulation (cross) and the analytical theory of single ledge (solid curve) in Figure 9 is due to the complex diffusion-field interaction between ledges forming the plate tip and ledges on the broad face. It is seen that both Trivedi and computer simulation agree very well with experiment at this temperature. At 450 °C (Figure 10), the theory of single ledge, Z-H, and Trivedi all agree well with the experimental data. It is noted that a nonzero interface kinetic coefficient is assumed in the calculation by Trivedi. The computer simulation was not conducted at this temperature, because no data on the ledge spacing were available. A large difference from the solid curve (single ledge) is not expected at this supersaturation.

Figure 11 compares the experimental lengthening rates of bainite plates in Fe-C-Ni alloys[23] with calculation from the theory of ledge growth. Two values of the tip radius were used: one used by original authors[23] and the other calculated from Trivedi's theory.[56] As seen in the figure, the theory of ledgewise lengthening

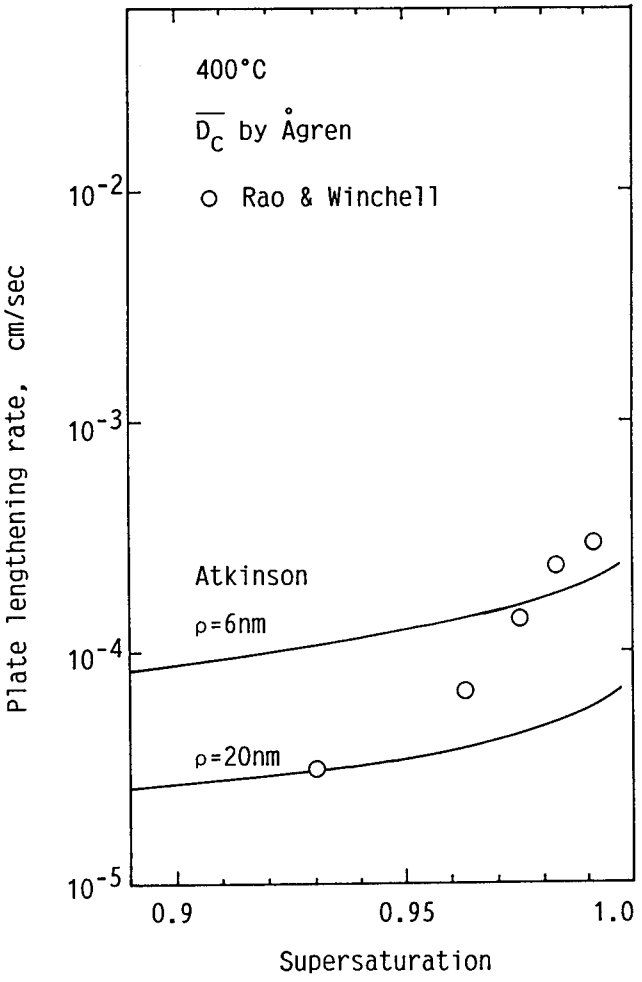


Fig. 11—Comparison of Rao and Winchell's data on the lengthening of bainite plates in Fe-C-Ni alloys[23] at 400 °C with calculation.

gives the rates of the same order of magnitude with experiment. It is emphasized that this agreement was obtained solely from the assumption of local equilibrium of carbon at the bottom of a ledge riser. That is, the interfacial kinetic coefficient, capillarity, or the energy of plate interfaces was not invoked.

Then, a question arises as to what governs the plate tip radius or the height of the leading ledge. According to measurement,[21] the plate tip radius tends to decrease with decreasing temperature or with increasing supersaturation at a given temperature. Hence, the size of ferrite nucleus is likely to be a primary factor to be considered. The grain-boundary structure, particularly the size of repetitive units of atom arrangement, may play some role. Also, the tip radius may vary during the ledge migration due to coalescence. The probability of coalescence is small at high supersaturations, in accord with the observation. Obviously, more work is necessary on the atomic structure of plate tip and early stages of ferrite formation.

V. CONCLUSIONS

The mechanism of the Widmanstätten plate formation was discussed on the basis of thermodynamic calculation of the formation temperature, theories of plate lengthening kinetics, and comparisons thereof with experimental data. The following points are especially noted:

1. The reported W_s temperature is about 50 °C higher than the calculated T_0 temperature in many ternary Fe-C-X alloys. Hence, the displacive mechanism in which the atomic bonds of solute atoms (both carbon and alloying element) with surroundings in the new phase are inherited from the parent phase is not directly responsible for the plate formation.
2. The W_s appears to become lower than the T_0 temperature only in alloys containing a substantial amount of Mn or Ni. The mechanism of a marked restraining effect of Mn or Ni on the formation of Widmanstätten ferrite plate is not understood.
3. The partition of alloying elements between ferrite and austenite ceases at temperatures above W_s. Thus, both allotriomorphic and Widmanstätten ferrite are the same in terms of the partitioning of carbon and alloying element.
4. The mechanism for the occurrence of transformation stasis (incomplete transformation) is not identified from thermodynamic calculations which takes into account the possible effects of strain energy on the equilibrium between bainitic ferrite and austenite.
5. Under the assumption that the plate lengthening can be simulated by the migration of a pair of single ledges, the lengthening rate is predicted to be larger at low supersaturations and to be remarkably smaller at high supersaturations than from theories of plate lengthening used thus far.
6. The ledgewise lengthening was shown to give a good account for reported lengthening rates of Widmanstätten and bainite plates in Fe-C and Fe-C-Ni alloys in a wide range of reaction temperature and supersaturation.

REFERENCES

1. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, pp. 845-56.
2. R.F. Heheman, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1972, vol. 2, pp. 1077-94.
3. K.R. Kinsman, E. Eichen, and H.I. Aaronson: *Metall. Trans. A*, 1975, vol. 6A, pp. 303-17.
4. H.J. Lee and H.I. Aaronson: *Acta Metall.*, 1988, vol. 36, pp. 787-93.
5. M.G. Hall and H.I. Aaronson: Paper presented at *Pacific Rim Conf. on the Roles of Shear and Diffusion in the Formation of Plate-shaped Transformation Products*, Hawaii, ASM International, 1992.
6. H.K.D.H. Bhadeshia and J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 767-97.
7. H.I. Aaronson, W.T. Reynolds, Jr., G.J. Shiflet, and G. Spanos: *Metall. Trans. A*, 1990, vol. 21A, pp. 1343-80.
8. M. Hillert: "The Growth of Ferrite, Bainite and Martensite," Internal Report, Swedish Institute for Metal Research, 1960.
9. H.I. Aaronson: *The Mechanism of Phase Transformations in Crystalline Solids*, Institute of Metals, London, 1969, pp. 270-81.
10. M. Enomoto and H. Tsubakino: *Mater. Trans. JIM*, 1991, vol. 32, pp. 642-57.
11. Zi-Kui Liu: Report No. TRITA-MAC-0486, Royal Institute of Technology, Stockholm, 1992.
12. K. Tsuya: *J. Mech. Lab. Jpn.*, 1955, vol. 1, p. 1.
13. H. Yada and T. Ōoka: *J. Metall. Soc. Jpn.*, 1967, vol. 31, pp. 766-71.
14. M. Nemoto: *High Voltage Electron Microscopy*, P.R. Swann, C.J. Humphreys, and M.J. Goringe, eds., Academic Press, London, 1974, p. 230.
15. C. Zener: *Trans. AIME*, 1946, vol. 167, pp. 550-95.
16. P.G. Ivantsov: *Dokl. Akad. Nauk SSSR*, 1947, vol. 58, p. 567.
17. M. Hillert: *Jernkontorets Ann.*, 1957, vol. 141, pp. 757-89.
18. G. Horvay and J.W. Cahn: *Acta Metall.*, 1961, vol. 9, pp. 695-705.
19. R. Trivedi: *Metall. Trans.*, 1970, vol. 1, pp. 921-27.
20. M. Hillert: *Metall. Trans. A*, 1975, vol. 6A, pp. 5-19.
21. E.P. Simonen, H.I. Aaronson, and R. Trivedi: *Metall. Trans.*, 1973, vol. 4, pp. 1239-45.
22. R.H. Goodenow, S.J. Matas, and R.F. Hehemenn: *Trans. Met. Soc. AIME*, 1963, vol. 227, pp. 651-58.
23. M.M. Rao and P.G. Winchell: *Trans. Met. Soc. AIME*, 1967, vol. 239, pp. 956-60.
24. J.M. Oblak and R.F. Heheman: *Transformation and Hardenability in Steels*, Climax Molybdenum Company, Ann Arbor, MI, 1967, pp. 15-30.
25. A. Ali and H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1989, vol. 5, pp. 398-402.
26. H.K.D.H. Bhadeshia and D.V. Edmonds: *Acta Metall.*, 1980, vol. 28, pp. 1265-73.
27. H.K.D.H. Bhadeshia and A.R. Waugh: *Acta Metall.*, 1982, vol. 30, pp. 775-83.
28. P.G. Boswell, K.R. Kinsman, G.J. Shiflet, and H.I. Aaronson: *Mechanical Properties and Phase Transformations in Engineering Materials*, S.D. Antolovich, R.O. Ritchie, and W.W. Gerberich, eds., TMS-AIME, Warrendale, PA, 1986, pp. 445-66.
29. G.J. Shiflet and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1413-32.
30. H. Goldstein and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1465-78.
31. W.T. Reynolds, Jr., F.Z. Li, C.K. Shui, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1433-63.
32. C. Laird and H.I. Aaronson: *Acta Metall.*, 1969, vol. 17, pp. 505-19.
33. C. Atkinson, K.R. Kinsman, and H.I. Aaronson: *Scripta Metall.*, 1973, vol. 7, pp. 1105-14.
34. M. Enomoto: *Acta Metall.*, 1987, vol. 35, pp. 947-56.
35. H.J. Lee and H.I. Aaronson: *J. Mater. Sci.*, 1988, vol. 23, pp. 150-60.
36. B. Uhrenius: *Hardenability Concepts with Applications to Steel*, D.V. Doane and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, pp. 28-81.

37. D.E. Coates: *Metall. Trans.*, 1973, vol. 4, pp. 1077-86.
38. H.I. Aaronson and H.A. Domian: *Trans. Met. Soc. AIME*, 1966, vol. 236, pp. 781-95.
39. M. Enomoto: *Trans. ISIJ*, 1988, vol. 28, pp. 826-35.
40. M. Enomoto: *Phase Transformations '87*, G.W. Lorimer, ed., Institute of Metals, London, 1987, pp. 462-65.
41. H.I. Aaronson, P.G. Boswell, and K.R. Kinsman: *Mechanical Properties and Phase Transformations in Engineering Materials*, S.D. Antolovich, R.O. Ritchie, and W.W. Gerberich, eds., TMS-AIME, Warrendale, PA, 1986, p. 467-73.
42. J.R. Bradley and H.I. Aaronson: *Metall. Trans. A*, 1981, vol. 12A, pp. 1729-50.
43. W.W. Mullins and R.F. Sekerka: *J. Appl. Phys.*, 1963, vol. 24, pp. 323-29.
44. R.D. Townsend and J.S. Kirkaldy: *Trans. ASM*, 1968, vol. 61, pp. 605-19.
45. H.I. Aaronson, D. Eylon, C.M. Cooke, M. Enomoto, and F.H. Froes: *Scripta Metall.*, 1989, vol. 23, pp. 435-40.
46. M. Enomoto: *Metall. Trans. A*, 1991, vol. 22A, pp. 1235-45.
47. M. Enomoto, G. Spanos, and R.A. Masumura: *Interface Dynamics and Growth*, Materials Research Society Symp. Proc., Materials Research Society, Pittsburgh, PA, 1991, vol. 237, pp. 119-30.
48. J.D. Watson and P.G. McDougall: *Acta Metall.*, 1973, vol. 21, pp. 961-73.
49. A.D. King and T. Bell: *Metall. Trans. A*, 1975, vol. 6A, pp. 1419-29.
50. R.F. Hehemann and A.R. Troiano: *Met. Prog.*, 1956, vol. 70(2), p. 97.
51. H.I. Aaronson, M.G. Hall, D.M. Barnett, and K.R. Kinsman: *Scripta Metall.*, 1975, vol. 9, pp. 705-12.
52. J. Ågren: *Metall. Trans. A*, 1979, vol. 10A, pp. 1847-52.
53. R.L. Orr and J. Chipman: *Trans. Met. Soc. AIME*, 1967, vol. 239, pp. 630-33.
54. L. Kaufman, E.V. Clougherty, and R.J. Weiss: *Acta Metall.*, 1963, vol. 11, pp. 323-35.
55. W.T. Reynolds, Jr., S.K. Liu, F.Z. Li, S. Hartfield, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1479-91.
56. C. Atkinson: *Proc. R. Soc. Lond.*, 1981, vol. A378, pp. 351-68.
57. W.T. Reynolds, Jr. and H.I. Aaronson: *Scripta Metall.*, 1985, vol. 19, pp. 1171-76.
58. J. Ågren: *Scripta Metall.*, 1986, vol. 20, pp. 1507-10.

Diffusion in Growth of Bainite

MATS HILLERT

Edgewise growth rates for Widmanstätten ferrite and bainite in low alloy steels can be represented with an empirical equation showing proportionality to the square of the supersaturation of the austenite. The proportionality constant has a value in reasonable agreement with the assumption of rate control by carbon diffusion. The growth rates are too low to give a noticeable supersaturation of carbon in the growing ferrite.

The experimental B_s for low alloy steels does not seem to be related to the T_0 line, nor does B_s evaluated from the incomplete transformation to bainite for an alloy steel. By assuming rate control by carbon diffusion, the empirical equation can be used to calculate the growth rate under paraequilibrium or no partition, local equilibrium (NPLE) conditions. Experimental growth rates for a similar steel falls in-between.

The fact that paraequilibrium does not seem to apply is taken as an indication that the α/γ interface for Widmanstätten ferrite and bainite is not of a purely martensitic type.

I. INTRODUCTION

THE classical method to study transformations in alloys was to measure the change of some physical property during continuous heating or cooling in a furnace. By combination with microscopical studies of the resulting structures, it was possible to identify various transformations. Hultgren in 1920[1] extended this technique by interrupting the gradual cooling by quenching from various temperatures and also by transferring the specimen to another furnace of a constant temperature and keeping it there for some time before quenching. He studied a W steel and observed a transformation occurring isothermally below the temperature range of pearlite formation and resulting in "secondary ferrite" of a Widmanstätten appearance. For instance, he wrote, "At 525 °C one minute is sufficient for ferrite to appear while at 550 °C no ferrite appears in fifteen minutes." This structure would today be called carbide-free bainite and is well known from steels containing Si. Hultgren explained the formation of the secondary ferrite by referring to the extrapolated A_3 line in the phase diagram (line AB in Figure 1). In particular, he stated that "the composition of the separated ferrite follows the extended equilibrium line AC."

In another context, Hultgren[2] mentioned "needles of troostite" that sometimes occur in hardened carbon steel and more frequently in alloy steels. This is probably what today is called lower bainite, although Hultgren believed that martensite needles had first formed and had then been tempered.

Robertson in 1929[3] improved the isothermal technique used by Hultgren by immersing the specimen in a molten metal, which is necessary in order to prevent transformation in a plain carbon steel during cooling, although it was not necessary in Hultgren's W steel. However, it seems that Robertson did not intentionally interrupt the isothermal transformation by quenching after various times, as Hultgren did. Anyway, Robertson made a very detailed study of the transformation products formed between 480 °C and 380 °C in a eutectoid carbon steel, *i.e.*, bainite. He proved that it cannot be regarded as tempered martensite but forms by decomposition of austenite "initiated by the $\gamma \rightarrow \alpha$ change." He even claimed to have seen minute plates of ferrite in the fine structure down to 390 °C.

Robertson explained the fact that bainite can form at a lower temperature than pearlite by postulating that the carbon content of ferrite increases at lower temperatures, thus decreasing the need of carbon diffusion. He even proposed that ferrite could inherit all the carbon from the austenite at a low temperature and no cementite would form under such conditions. It seems that he thus proposed a gradual change from bainite to martensite as the reaction temperature is lowered.

Davenport and Bain in 1930[4] also used molten metal but combined this with Hultgren's technique of interrupting the isothermal transformation by quenching after various times. Thus, they were able to construct the first time-temperature-transformation (TTT) diagram with a typical transformation curve. On the lower side of the C-shaped curve, they observed the same structure as Robertson. They described it as a dark-etching structure with a more or less acicular shape. They suggested that the name troostite should be reserved for this structure which corresponds to upper bainite in today's nomenclature. At lower temperatures, they observed more needle-like shapes and introduced the name martensite-troostite for the structure today called lower bainite. Davenport and Bain did not add much to the discussion of the nature of bainite given by Robertson. In agreement with Hultgren they stated that they "suspect that the final constituent designated martensite-troostite first forms in the manner of martensite but is subsequently more or less tempered."

Numerous experimental studies of TTT diagrams, initiated by the work of Davenport and Bain, served to focus considerable interest on the microstructure today called bainite and finally led to the adoption of that name.

MATS HILLERT, Professor Emeritus, is with the Division of Physical Metallurgy, Royal Institute of Technology, S-10044 Stockholm, Sweden.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

Robertson started his study in order to test the popular idea that the various transformation products of austenite form in the succession austenite → martensite → troostite → sorbite → pearlite and the final structure in a specimen depends on how far this series of reactions has proceeded. His conclusion was that austenite instead transforms directly to any of these structures, depending on the temperature.

It is fascinating that in spite of his results and the confirmation by Davenport and Bain, it is still a popular idea that the first stage in bainite formation is a rapid shear transformation (more or less identical to the martensitic transformation) and that the supersaturation of carbon, which is thus created, is only gradually relieved by a secondary reaction through which either carbide precipitates or carbon diffuses into the adjoining austenite. There are certainly many reasons why this idea has survived and has even been extended to include upper bainite. One reason may be that already Robertson felt the need to explain how bainite can form at a lower temperature than pearlite and thus proposed that bainitic ferrite forms with an increasing amount of carbon at lower temperatures which would allow it to grow with a higher rate. This idea was expressed by several authors during the 1930s but it is not always clear if the authors simply refer to the increased solubility according to the $\alpha/\alpha + \gamma$ phase boundary (the AC line in Figure 1), which was often believed to go to high carbon contents, or if they actually proposed a supersaturation with respect to austenite. It is self-evident that there will be a supersaturation with respect to cementite. As an example, one may mention Wever and Lange[5] who in 1932 gave exactly the same description as Robertson but illustrated their discussion with a phase diagram (Figure 2), showing that ferrite in equilibrium with austenite would increase its carbon content at lower temperatures and reach a value of 1 pct at 130 °C.

In his attempt to describe all the various transformations of austenite with different well-defined mechanisms, Zener[6] tried the idea that bainite forms initially without any diffusion of carbon and he showed that the upper temperature for bainite formation, B_s, would then coincide with the T_0 line. By further assuming that carbon will rapidly diffuse into the adjoining austenite and thus make it less prone to transform to ferrite, Zener was able to give a qualitative explanation of the "incomplete" transformation to bainite, sometimes observed close to B_s in alloy steels. His idea was that bainite would no longer form when the remaining austenite had reached the T_0 line through the increasing carbon content.

In contrast, Hultgren[7] retained his idea of bainitic ferrite growing with a carbon content according to the $\alpha/\alpha + \gamma$ phase boundary and illustrated the edgewise growth of bainite with the sketch reproduced in Figure 3. It illustrates that the bainitic transformation is initiated by the growth of Widmanstätten plates of ferrite. By electron microscopy, Modin and Modin[8] confirmed that Hultgren's picture applies also to bainite found at lower temperatures.

Ko and Cottrell[9] found that bainite forms with a surface relief effect, as does martensite. They concluded that both structures form by the same shear of the lattice

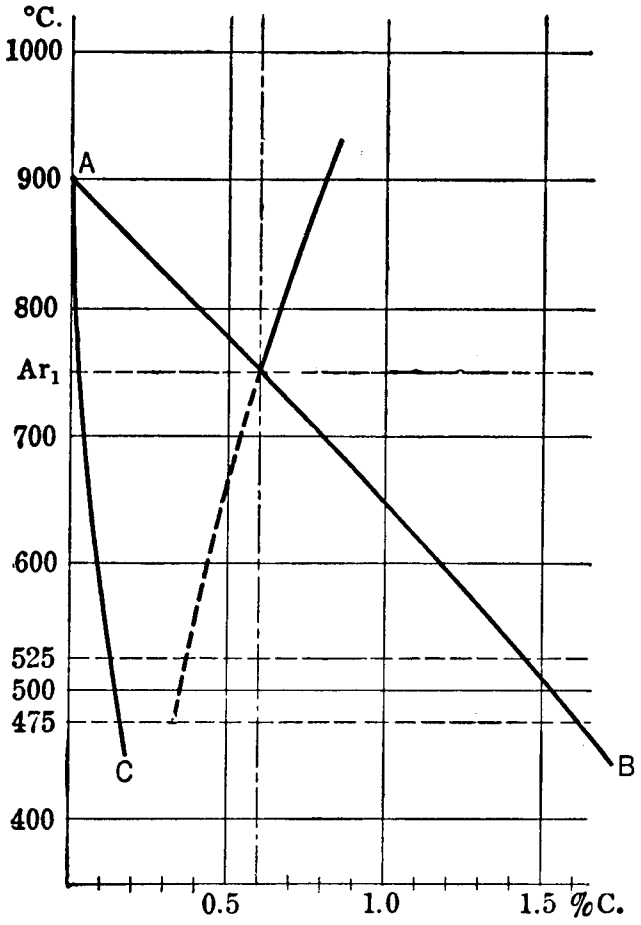


Fig. 1—The Fe-C phase diagram with extrapolated phase boundaries representing conditions of local equilibrium.[1]

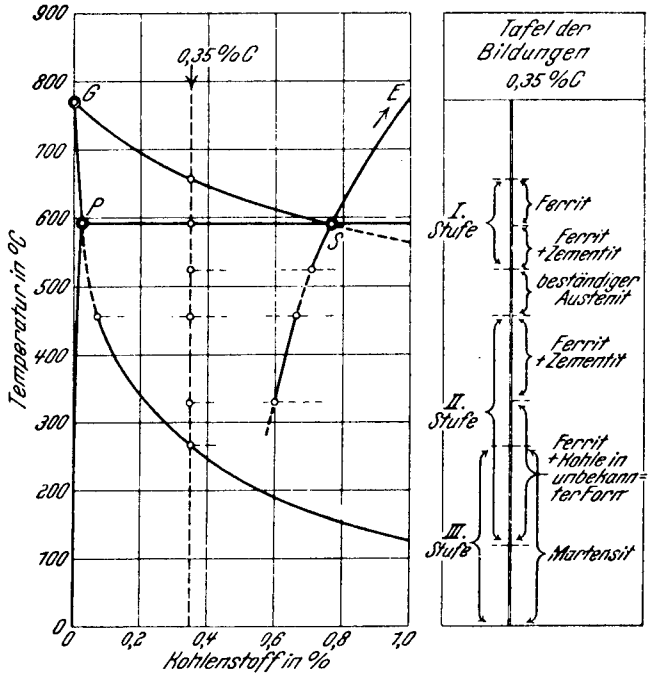


Fig. 2—An Fe-C phase diagram with the $\alpha/\alpha + \gamma$ boundary extrapolated to high carbon contents at low temperatures.[5]

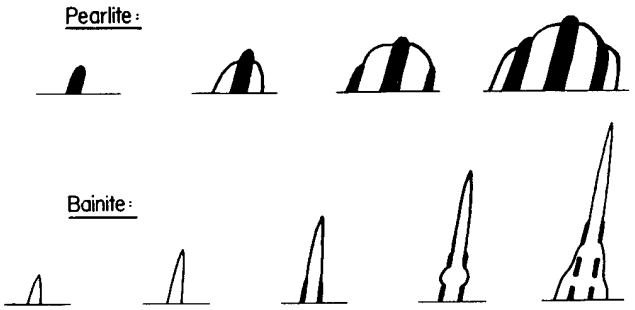


Fig. 3—The shapes of pearlite and bainite during growth. The edgewise growth of bainite is governed by the growth of Widmanstätten α.[7]

but proposed that the growth rate of bainite is governed by carbon diffusion. They also proposed that there is no fundamental difference in the mode of formation of Widmanstätten ferrite, upper bainite, and lower bainite. The latter viewpoint was shared by Hillert,[10] who modified an equation by Zener for diffusion controlled edgewise growth of a plate and applied it to Widmanstätten ferrite and bainite. In a subsequent article, Hillert[11] presented experimental growth rate data for low alloy steels with 0.21 to 0.81 pct C and compared them with calculated values. He found that the experimental growth rates were somewhat lower and proposed that there is some resistance in the interface that must be overcome. Soon after, an article by Kaufman *et al.*[12] appeared in which they applied the same growth equation to the same data and to similar data for 0.96 and 1.43 pct C by Speich and Cohen.[13] They used an improved description of the Fe-C phase diagram and still found that the experimental data were somewhat lower but concluded that there was agreement within the uncertainty of the calculations. Later measurements by Townsend and Kirkaldy[14] were in essential agreement with the previous ones. The growth rate equation was improved by Trivedi,[15] and a test of new data[16] confirmed that the experimental growth rates may be somewhat lower than predicted by carbon diffusion. It would thus seem that there should then have been general agreement that Widmanstätten ferrite and bainitic ferrite grow edgewise with a rate low enough to allow carbon to diffuse away and let ferrite grow without a supersaturation of carbon.

In a classical debate on the Bainite reaction,[17] Aaronson and Kinsman argued for a growth mechanism based upon ledges and controlled by the rate of carbon diffusion whereas Heheman argued for a martensitic type of shear mechanism. He found it "difficult to argue against these diffusion controlled models" but nevertheless seemed to favor a proposal by Oblak and Heheman[18] that the edgewise growth occurs in small steps where individual subunits grow rapidly without partition of carbon to austenite, *i.e.*, without diffusion of carbon. Aaronson and Kinsman objected that the edgewise growth of bainite has not been observed to occur in such a discontinuous fashion and presented results using thermionic electron microscopy which showed continuous growth. This was later confirmed by Nemoto[19] using dynamic experiments in a high voltage electron microscope. His results are reproduced in

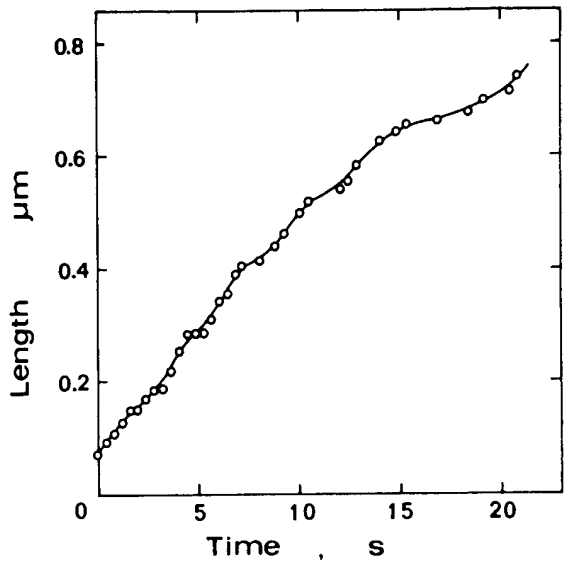


Fig. 4—Growth of bainite tip at about 380 °C in a steel with 0.51 °C and 9.1 pct Ni.[19]

Figure 4 and show with a very high resolution that the growth rate is indeed very continuous. It should be added that the average growth rate given by Figure 4 agrees well with the more macroscopic growth rates observed by Goodenow *et al.*[20] for a similar steel.

It would seem that the old controversy regarding the carbon content of bainitic ferrite should have been resolved by the preceding results. However, from the *Proc. Int. Conf. on Bainite at the World Materials Congress*, 1988,[21] and the *Special Issue on the Bainitic Transformation*, Aug. 1991, of *Mater. Trans.* JIM, it seems that there are still some researchers who prefer believing that bainitic ferrite grows with a supersaturation of carbon and even without any diffusion of carbon. As an example, Bhadeshia and Christian[22] came to the conclusion that the only "key characteristic" for which bainite is different from martensite is that diffusion of carbon plays some role in the nucleation of bainite but not martensite. (An additional difference that martensite does not show the incomplete reaction phenomenon known for bainite is probably a misprint.) As a consequence, it may be justified again to consider the role of carbon diffusion although it has been thoroughly discussed, for instance, by Aaronson *et al.*[23] and by Enomoto and Tsubakino.[24] The present discussion will mainly be based upon the edgewise growth rate.

II. EDGEWISE GROWTH RATE OF BAINITE

A calculation of the diffusion controlled growth rate is usually based on the assumption of local equilibrium at the moving interface and requires information on the extrapolation of phase boundaries. Hillert[11] used a simple linear extrapolation of the A_3 line. It would thus be interesting to repeat his calculation using a phase diagram obtained by a thermodynamic assessment of the system.[25] The modified Zener equation is

$$\frac{v}{D} = \frac{(x^{\gamma/\alpha} - x^0)^2}{x^0} \cdot \frac{RT}{8\sigma V_m} \qquad [1]$$

where v is the growth rate, x^0 is the mole fraction of carbon in the alloy, and $x^{\gamma/\alpha}$ is the value in the γ phase in front of the growing α. In the derivation, the interfacial energy σ was assumed to be isotropic, which is certainly not the case for a Widmanstätten plate. A similar situation arose in an analysis of plate-like precipitation in solidification of cast iron,[26] and it was then argued that one should insert the low value for the flat sides.

Figure 5 shows $(vx^0/DT)^{0.5}$ as a function of $x_e^{\gamma/\alpha} - x^0$, where $x_e^{\gamma/\alpha}$ is the equilibrium value according to the phase diagram. All the data can be reasonably well accounted for by a straight line with a slope of 1580 (m, K)$^{-0.5}$, corresponding to a value of $\sigma = 0.06$ J/m^2, which seems a reasonable value for a flat, immobile interface. Again, this result may be taken as support for the idea that the edgewise growth of Widmanstätten ferrite and bainite is controlled by carbon diffusion.

However, it should be emphasized that the above analysis has several uncertainties. First, the extrapolation of the A_3 line is still uncertain. Second, D varies with the carbon content, and it is uncertain how the effective D value should be chosen. In the present case, the highest value was taken. Third, the dependence of D on temperature and carbon content is obtained by an extrapolation from higher temperatures, which has a considerable uncertainty. The present work was based on an extrapolation proposed by Ågren.[27] Fourth, the validity of the growth equation has been questioned. The more ambitious treatment by Trivedi[15] has shown that one should introduce a correction factor which grows larger as the supersaturation decreases. However, even though his treatment may be correct for the case of isotropic interfacial energy, it cannot be applied directly to the very anisotropic case of a Widmanstätten plate. Thus, the older and much simpler equation was used in the present work. Fifth, there is no guarantee that all the data in Figure 5 should be represented by a single line. For the low-temperature data, the extrapolation to zero growth rate is very long and the use of a higher value of σ would give a result more similar to Hillert's original result. Anyway, the result indicates that the experimental growth rates are in reasonable agreement with diffusion control, supporting the opinion that bainitic ferrite grows without a supersaturation of carbon.

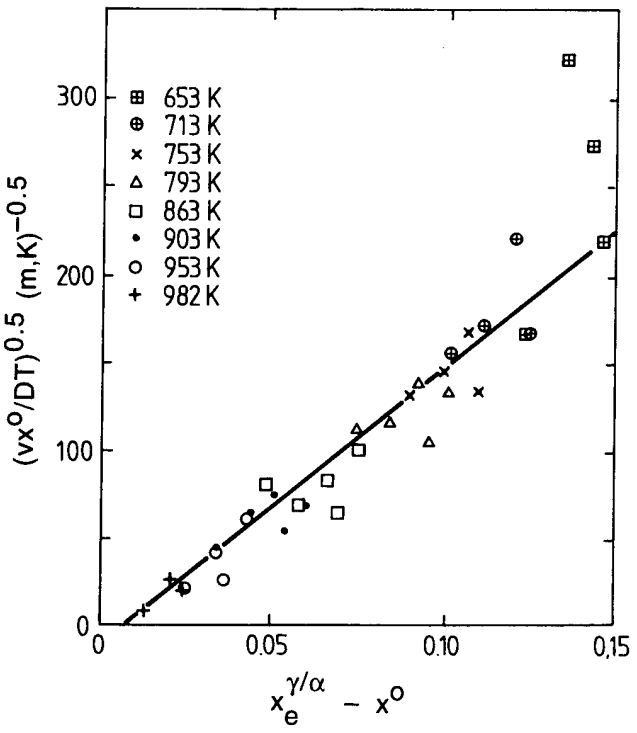


Fig. 5—Experimental edgewise growth rates of Widmanstätten α and bainite as function of the supersaturation of γ.

For a further analysis of the edgewise growth of Widmanstätten ferrite and bainite, we shall accept the straight line in Figure 5 as a reasonable empirical representation of experimental data. By neglecting the small intercept at $v = 0$, we shall write

$$v = 2{,}500{,}000 \cdot \frac{DT}{x^0} (x^{\gamma/\alpha} - x^0)^2 \text{ (m, K)}^{-1} \qquad [2]$$

III. COMPARISON OF GROWTH RATES FOR BAINITE AND PEARLITE

As already mentioned, Robertson's original proposal that the carbon content of bainitic ferrite increases at lower temperatures was made in order to explain why bainite can form at lower temperatures than pearlite. There still seem to be researchers who feel the need for such an explanation and thus favor the idea that the ferritic component of bainite develops by a different mechanism than diffusion.[28] Since the data presented in Figure 5 do not seem to indicate a difference between Widmanstätten ferrite and bainite, it should be of interest to examine if those data would predict that bainite should predominate in the low-temperature region. Thus, Eq. [2] was used to calculate transformation curves for two steels with 0.50 and 0.78 pct C. In Figure 6, they are compared with a curve for pearlite obtained by fitting a growth equation,

$$v = K \cdot D \cdot (x^{\gamma/\alpha} - x^{\gamma/\text{cementite}})^2 \qquad [3]$$

to data reported by Williams and Glover[29] for a high-purity Fe-C alloy. The value of D was taken at $(x^{\gamma/\alpha} + x^{\gamma/\text{cementite}})/2$, and K was adjusted to give the best fit.

Figure 6 demonstrates the well-known fact that bainite predominates at low carbon contents and at low temperatures. It is indicated that bainite can form below the pearlite nose even in very high carbon steels. It should be remembered that Figure 6 is based on experimental data. If the representation of the data with the growth rate equations based on diffusion control is accepted, then the reason why bainite can form below the pearlite nose is primarily the asymmetry of the phase diagram and, secondarily, the high diffusivity of carbon in γ obtained along the extrapolated γ/α phase boundary. The asymmetry of the phase diagram is demonstrated by Figure 7, where the dashed line is the middle line in the triangle where supercooled austenite is supersaturated with respect to cementite as well as ferrite. Along that line, pearlite would have the best chance to predominate and, intuitively, one may expect pearlite to predominate in a region around that line, *e.g.*, in the region between

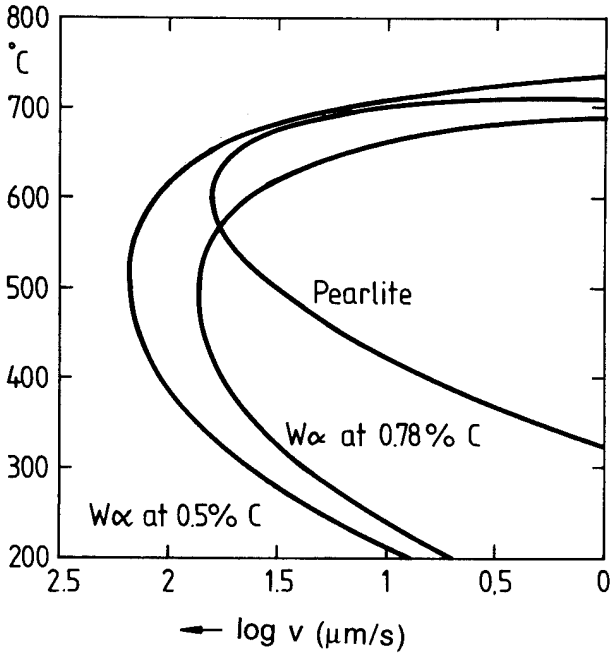


Fig. 6—Edgewise growth rates of Widmanstätten α evaluated for two steels from Figure 5, and experimental growth rates of pearlite.

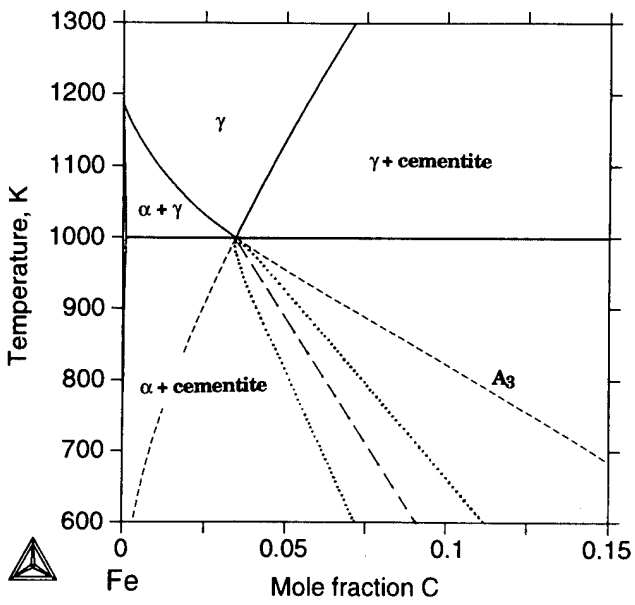


Fig. 7—Calculated Fe-C phase diagram showing the most preferred region for pearlite.

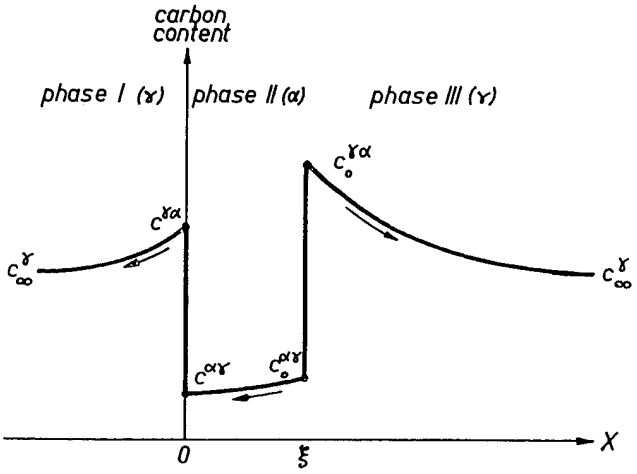


Fig. 8—Carbon concentration profile through a grain of α and the two neighboring grains of γ. The interface at $x = 0$ is immobile, and there, the carbon contents are lower than at the mobile interface,[11]

the two dotted lines. For a eutectoid or hypereutectoid alloy, one can thus understand why pearlite predominates in a high-temperature region and bainite in a lower temperature region. There is no need to invent a mechanism different from the diffusional one.

IV. THE QUESTION OF SUPERSATURATION

We should now turn to the fact that there have been many suggestions that bainitic ferrite forms with a supersaturation of carbon. In order to get a supersaturation, it is necessary but not sufficient to have a deviation from local equilibrium at the α/γ interface. As an example, it is well known that a grain-boundary allotriomorph of ferrite can have a flat side which is immobile while the opposite interface moves. The immobility of the flat side is due to the arrangement of the iron atoms at the interface. It is not due to the carbon atoms which are highly mobile and will most probably partition between ferrite and austenite at the interface according to a local equilibrium condition for carbon. As demonstrated by Figure 8 (taken from Reference 11), this kind of deviation from local equilibrium results in a lower carbon content of ferrite not a higher one.

In order to get a supersaturation of carbon, it is necessary to have a growth mechanism which yields such a high velocity of the interface that the mobility of carbon is no longer sufficient to maintain local carbon equilibrium at the interface. How a model for this phenomenon can be formulated was demonstrated by the present author,[11] the main parameter being the mobility of carbon in the interface relative to the growth rate. As the relative growth rate increases with a decreased carbon content of the steel or with a decreased temperature, there is a negligible supersaturation of the growing ferrite until a critical point is approached. A supersaturation is then developed and very quickly develops an instability where the growth rate suddenly increases and the transformation becomes diffusionless, *i.e.*, a martensitic transformation from the thermodynamic point of view. It is evident that this dramatic change in growth behavior does not take place within the experimental range covered by Figure 5, because there is no such break in the kinetic behavior.

It may be added that the preceding model of the possible transition from diffusion controlled to diffusionless growth has recently been developed in more detail by Ågren,[30] and his model yields the same conclusion.

A rough estimate of the critical point where the supersaturation develops dramatically can be made by using the simple relation

$$v = D/d \qquad [4]$$

where the mobility of carbon in the interface is represented by the diffusion constant in the interface, D, and where d is some measure of the width of the interface. By simply taking d as 10 Å and D as the diffusion coefficient in ferrite, we get the result shown in Figure 9. However crude this estimate may be, a comparison with the experimental growth rates in Figure 6 (the curve for 0.5 pct C being reproduced as a dashed line in Figure 9) demonstrates that in principle there may be a point of intersection which would roughly correspond to the critical point discussed earlier. Unfortunately, the choice of values for D and d is very uncertain and the exact position of the critical point is difficult to predict. However, Figure 9 clearly demonstrates that the growth rates at high temperatures are much too low to yield a noticeable supersaturation. The fact that the data in Figure 5 do not show any break in behavior is a strong indication that the critical point falls below the experimental conditions covered by those data.

Considering the fact that bainite in high alloy steels often grows with a rate one or more orders of magnitude lower than in low alloy steels whereas D and d should be roughly the same, it is even safer to conclude that bainitic ferrite in alloy steels does not grow with a noticeable supersaturation.

V. CONDITIONS FOR COMPLETELY DIFFUSIONLESS GROWTH

Most of the proponents of bainitic ferrite growing with a supersaturation of carbon seem to favor the idea of completely diffusionless growth. Following Zener, they thus like to identify the upper temperature of bainite formation, B_s, with the T_0 line or with a line displaced further down in order to account for strain effects. Such curves are presented in Figure 10, where the numbers represent the extra driving force available, expressed as J/mol. The start of martensite in Fe-C alloys according to Ohmori and Maki[31] is shown as a dashed line in Figure 10, and it is evident that the martensitic transformation occurs under conditions where the available driving force is of the order of 1500 or 2000 J/mol. (In this calculation, the effect of ordering of the carbon atoms in martensite was not considered.) It is generally believed that this driving force is required for the nucleation. Christian[32] has estimated the necessary driving force during growth to as little as 600 J/mol. Evidently, the rest of the driving force is used to give a very high growth rate. Growth rates similar to the speed of sound have been reported.[33] Under such conditions, there is not enough time for equilibration of temperature and a large part of the driving force may be absorbed due to the process being adiabatic. Figure 11 shows the critical line for a diffusionless adiabatic transformation, T_{adiab}, which falls below T_0. By adding the requirement of 600 J/mol, one gets a line, $T_{ad,600}$, which is surprisingly close to the experimental M_s, as demonstrated in Figure 12. This makes one wonder if the position of M_s is really determined by the nucleation process or has something to do with the growth conditions.

Figure 13 shows the data for B_s by Ohmori and Maki.[31] These data fall below T_0 but seem to be on their way to cross T_0. It is unfortunate that they did not show the extension of their B_s line to higher carbon contents. Possibly, the interference of pearlite made those measurements more difficult. Anyway, the slope of the B_s line is quite different from the slope of the T_0 line, and these data can hardly be taken as an indication that the T_0 line plays any role in the formation of bainite.

Steven and Haynes[34] made a very extensive study of the formation of martensite and bainite in steels. By regression analysis, they fitted an equation and claimed that it should be valid for carbon contents 0.1 to 0.55 pct and manganese contents 0.2 to 1.7 pct. The contents

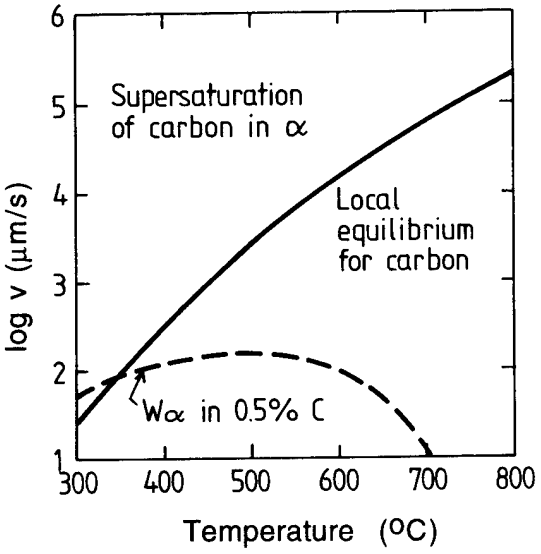


Fig. 9—Estimated critical growth rate required for the formation of supersaturated α. The dashed line represents experimental growth rates and is taken from Figure 6.

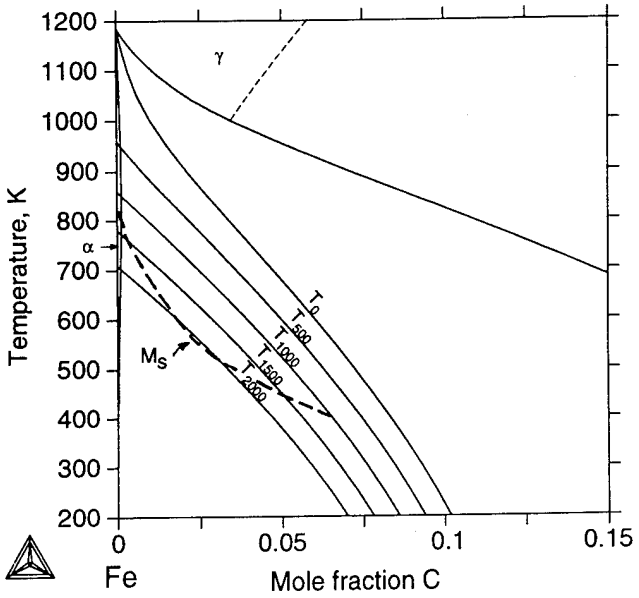


Fig. 10—The Fe-C phase diagram with lines calculated for diffusionless transformation $\gamma \rightarrow \alpha$, assuming various values for the driving force. M_s is the experimental temperature for the start of the martensitic transformation.

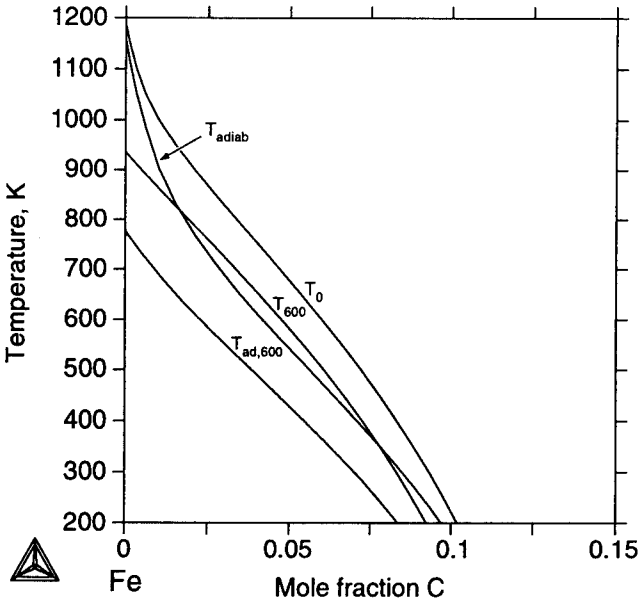


Fig. 11—Calculated curves where a diffusionless transformation $\gamma \rightarrow \alpha$ can take place adiabatically for two different driving forces.

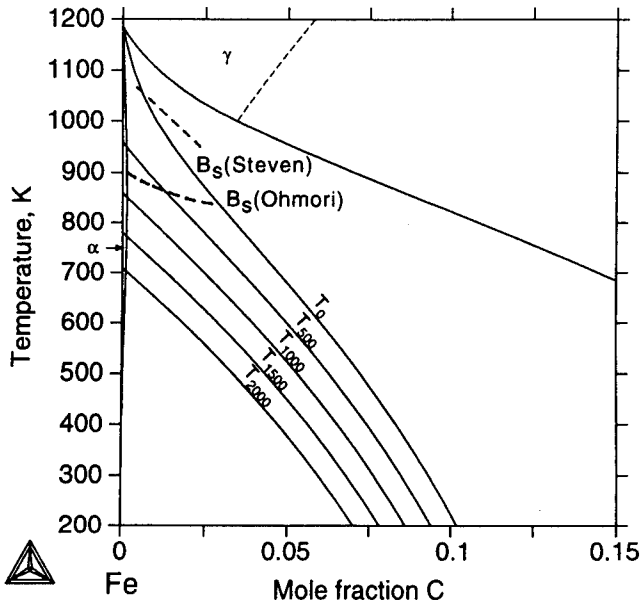


Fig. 13—Experimental values for the start of bainite in Fe-C alloys compared with lines for diffusionless transformation.

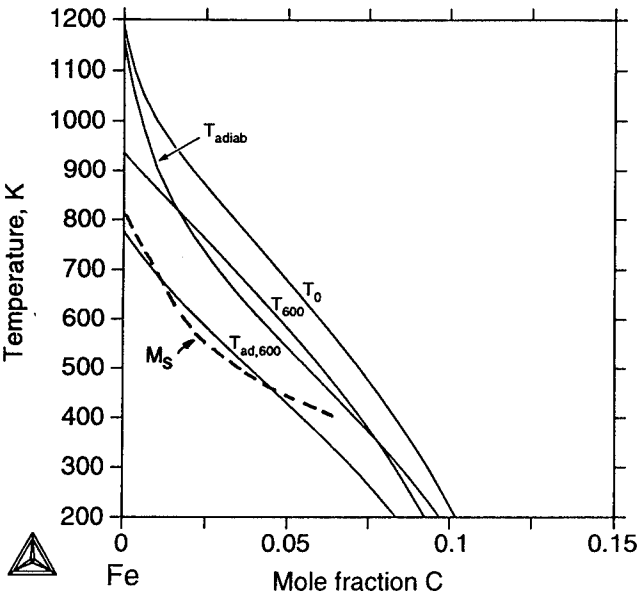


Fig. 12—Comparison of the experimental M_s line and lines for adiabatic growth.

of other alloying elements are allowed to be zero. For the lower manganese limit, their equation is B_s (°C) = 812 to 270 · [pct C]. This line is also included in Figure 13, and most of it falls well above the T_0 line.

VI. CONCLUSIONS FOR THE FE-C SYSTEM

For the Fe-C system and low alloy steels, we may draw the following conclusions.

(1) The edgewise growth rate of Widmanstätten ferrite, including bainite, can be reasonably well accounted for as controlled by carbon diffusion.

(2) The edgewise growth rate of Widmanstätten ferrite, including bainite, is generally too low to give a noticeable supersaturation of carbon.

(3) There is no break in the edgewise growth rate which indicates a change of mechanism that would distinguish bainite from Widmanstätten ferrite.

(4) The B_s line does not seem to be related to the T_0 line.

VII. BAINITE IN ALLOY STEELS

It is sometimes[22] stated that experimental support for the idea that bainite forms initially as supersaturated ferrite comes from the so-called incomplete transformation to bainite close to B_s. There, the formation of bainite in some alloy steels stops before the remaining austenite has increased its carbon content to the A_3 line. As an example, Bhadeshia and Edmonds[35] estimated the carbon content of austenite when the bainite formation stops by measuring the volume fraction of bainite and assuming that after its formation as supersaturated ferrite there is time for the carbon to diffuse into the austenite phase. They accepted Zener's suggestion that bainite can no longer form because of the increased carbon content, and the calculated carbon content of the remaining austenite would thus represent B_s of a steel with that carbon content and the initial alloy content (3.00 pct Mn and 2.12 pct Si). Following their proposal, the dashed line shown in Figure 14 was calculated from the same data[36] and is compared with the T_0 line. It intersects the T_0 line and has a completely different slope. This information can hardly be taken as support for the idea that bainite forms below T_0.

After referring to the incomplete transformation phenomenon as support for the hypothesis that the growth of bainitic ferrite occurs without any carbon diffusion, Bhadeshia and Christian[22] state that "there is as yet no reasonable alternative explanation of this phenomenon." Thus, it may be justified here to explain the role of alloy

elements in the bainitic transformation although this subject has been covered extensively by Aaronson, *e.g.*, in Reference 37.

Hultgren first discovered that alloy elements in steel do not partition between ferrite and carbide in the bainitic transformation. Believing in local equilibrium for carbon, he proposed that bainite forms under paraequilibrium, *i.e.*, under local equilibrium for carbon but without any diffusion of iron or substitutional alloy elements relative to each other. Hillert[38] pointed out that the experimental result showing no partition could just as well be explained by assuming complete local equilibrium at the interface if one considers the formation of a thin spike (of width $d = D/v$) of the alloy elements in front of the moving interface. He called this condition "false paraequilibrium," but today the notation NPLE (no partition, local equilibrium) introduced by Kirkaldy's group[39] is more common. Hillert also described how the conditions would change from false paraequilibrium to true paraequilibrium as D/v takes values below atomic dimensions. In a subsequent article,[40] he proposed that the transition from false to true paraequilibrium would not occur so quickly because the mobility of iron and alloy elements inside the interface is higher than in the lattice. As D^{lattice}/v decreases below atomic dimensions, the spike will move inside the interface and have an effect closely related to the so-called solute drag. A first attempt to model this process has been presented by Liu and Ågren.[41]

In an assessment of the effect of alloy elements on the growth of pearlite,[42] it was found that the effect in many cases can be described by assuming complete local equilibrium with a spike although D^{lattice}/v is much less than 1 Å. This was taken as an indication that the spike, when moved inside the interface, can have about the same effect as it has in front of the interface. On the other hand, the fact that bainite can form in steels with 9 pct Ni, although the maximum solubility of Ni in ferrite is only 6.5 pct, can only be explained by assuming that the effect is smaller for bainite. It seems reasonable to assume that the effect of a spike inside the interface depends upon the structure of the interface. If it is incoherent and moves by individual jumps of atoms across the interface, then there should always be time enough for some diffusion of iron and substitutional alloy elements relative to each other in the interface and the spike can develop there. If the interface is more coherent, the diffusivity inside the interface should be lower and the spike may not develop as well.

Kinsman and Aaronson[43] even proposed that the effect may be stronger inside the interface by attraction between a carbide-forming element and carbon. Thus, they explained the bay in the transformation curve for ferrite in Fe-C-Mo alloys.

In view of the preceding considerations, it would be interesting to examine where the actual value of $x^{\gamma/\alpha}$ falls relative to the lines for true and false paraequilibrium. Thus, the lines for (true) paraequilibrium and NPLE (false paraequilibrium) were included in Figure 14, and the position of the dashed line seems to indicate that bainite in this case grows under conditions similar to false paraequilibrium at high temperatures but moves in the direction of true paraequilibrium at lower temperatures.

Using the empirical Eq. [2], it is possible to predict the growth rate under true and false paraequilibrium by simply taking $x^{\gamma/\alpha}$ from these lines. The results are presented in Figure 15 together with a curve calculated for a pure Fe-C alloy with the same carbon content (0.43 pct C). It is immediately evident that the effect of the alloy elements would be rather small if bainite could grow under true paraequilibrium. On the other hand, the

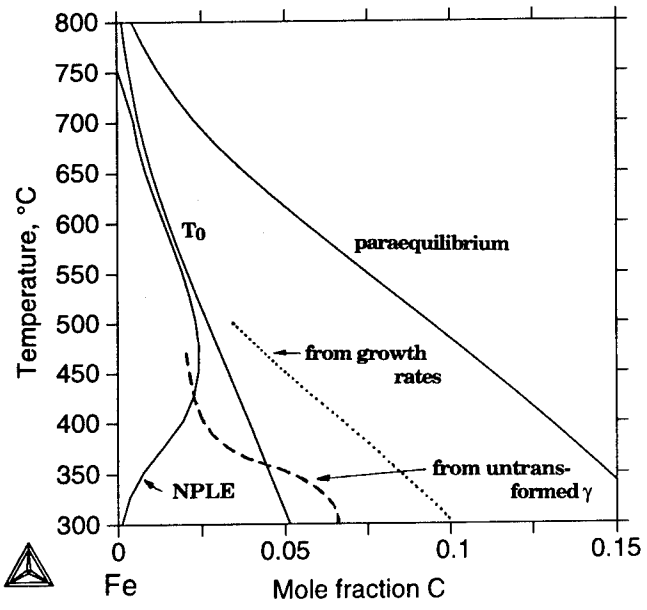


Fig. 14—Calculated lines for NPLE, paraequilibrium, and diffusionless transformation (T_0) in a steel with 3.00 pct Mn and 2.12 pct Si. The dashed line represents an estimate of the carbon content of the γ remaining after bainite has stopped growing. The dotted line is calculated in order to explain experimental growth rates in a similar steel.

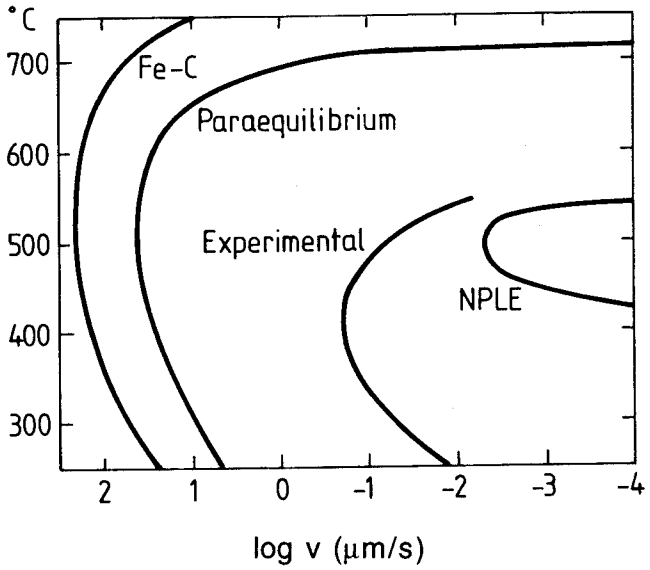


Fig. 15—Edgewise growth rates of Widmanstätten α or bainite in a steel with 3.00 pct Mn and 2.12 pct Si. Paraequilibrium holds if there is no diffusion of the alloy elements and NPLE if there is full local equilibrium at the interface but an alloy spike in the γ phase which allows α to grow with the same alloy contents as the initial γ had. The experimental curve is for a steel with 3.23 pct Mn.

calculated curve for false paraequilibrium (NPLE) shows an effect of many orders of magnitude. Unfortunately, no edgewise growth data have been reported for this steel. Instead, comparison is made with an experimental curve obtained from measurements on published micrographs[6] from a steel with a similar Mn content, 3.23 pct. That experimental curve falls between the curves for paraequilibrium and NPLE, being closer to the NPLE line at higher temperatures but moving closer to the paraequilibrium line at lower temperatures. The tendency is more evident if we use Eq. [2] and evaluate what carbon content is required in austenite at the interface in order to yield the growth rates observed experimentally. Those results are plotted with a dotted line in Figure 14. This curve also indicates that the growth conditions for bainite are close to false paraequilibrium at high temperatures and move closer to true paraequilibrium at lower temperatures but without ever getting very close.

Ideally, we would expect the dashed and dotted lines in Figure 14 to coincide, because they represent the same quantity, $x^{\gamma/\alpha}$, except for the difference that the dashed line has been obtained at zero growth rate (or rather zero rate of nucleation) and the dotted line from the experimental growth rates given in Figure 15. The difference between the two lines can be explained by several uncertainties in the experimental information. In addition, it may be worth while emphasizing that the stop of bainite formation may also be affected by the lack of nuclei for new plates or ledges if the broad faces of the plates, already formed, are immobile. The dashed line may thus fall to the left of actual $x^{\gamma/\alpha}$.

Even though there is no information on the properties of the α/γ interface with regard to the interaction with alloy elements and no quantitative predictions can be made, it is evident that the previous considerations, which are based upon the assumption of local carbon equilibrium, can provide a "reasonable alternative explanation" of the incomplete transformation phenomenon.

It is difficult to see how the mobility of iron and alloy elements in a glissile, martensitic interface can be very much higher than in the lattice. For such an interface, one should thus move away from false paraequilibrium and come very close to true paraequilibrium as D^{lattice}/v decreases below atomic dimensions. The experimental fact, presented in Figure 15, that actual growth rates are much lower than predicted by paraequilibrium indicates that there is some mobility of iron and the alloy element in the interface and may thus be taken as strong support for the opinion that bainite does not grow with a purely martensitic interface. On the other hand, the position of the dotted line in Figure 14 may be taken as an indication that the interface grows more and more coherent at lower temperatures in agreement with a suggestion by Purdy and Hillert[44] which has recently been modeled by Liu and Ågren.[45]

VIII. CONCLUSIONS FOR ALLOY STEELS

For bainite in alloy steels we may draw the following conclusions.

(1) The B_s line does not seem to be related to the T_0 line.

(2) Bainite grows under conditions falling between NPLE and paraequilibrium conditions.

(3) The growth conditions are closer to NPLE conditions at higher temperatures and move closer to paraequilibrium conditions at lower temperatures.

(4) Bainite does not grow with a purely martensitic type of interface but may become more and more coherent at lower temperatures.

ACKNOWLEDGMENTS

The author is grateful to Drs. John Ågren and Zikui Liu for stimulating discussions and to Dr. Bengt Hallstedt for collaboration in the numerical calculations.

REFERENCES

1. A. Hultgren: *A Metallographic Study on Tungsten Steels*, John Wiley, New York, NY, 1920.
2. A. Hultgren: *J. Iron Steel Inst.*, 1926, vol. 114, p. 421.
3. J.M. Robertson: *J. Iron Steel Inst.*, 1929, vol. 119, p. 391.
4. E.S. Davenport and E.C. Bain: *Trans. AIME*, 1930, vol. 90, p. 117.
5. F. Wever and H. Lange: *Mitt. Kaiser-Wilhelm Inst. Eisenf.*, 1932, vol. 14, p. 71.
6. C. Zener: *Trans. AIME*, 1946, vol. 167, p. 550.
7. A. Hultgren: *Trans. ASM*, 1947, vol. 39, p. 915.
8. H. Modin and S. Modin: *Jernkont. Ann.*, 1955, vol. 139, p. 481.
9. T. Ko and S.A. Cottrell: *J. Iron Steel Inst.*, 1952, vol. 172, p. 307.
10. M. Hillert: *Jernkont. Ann.*, 1957, vol. 141, p. 757.
11. M. Hillert: "The Growth of Ferrite, Bainite and Martensite," Internal Report, Swedish Institute for Metal Research, Stockholm, Sweden, 1960.
12. L. Kaufman, S.V. Radcliffe, and M. Cohen: *The Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, New York, NY, 1962, pp. 313-52.
13. G.R. Speich and M. Cohen: *Trans. TMS-AIME*, 1960, vol. 218, p. 1050.
14. R.D. Townsend and J.S. Kirkaldy: *Trans. ASM*, 1968, vol. 61, p. 605.
15. R. Trivedi: *Metall. Trans.*, 1970, vol. 1, p. 921.
16. E.P. Simonen, H.I. Aaronson, and R. Trivedi: *Metall. Trans.*, 1973, vol. 4, p. 1239.
17. R.F. Heheman and K.R. Kinsman: *Metall. Trans.*, 1972, vol. 3, p. 1077.
18. J.M. Oblak and R.F. Heheman: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 15-30.
19. M. Nemoto: *High Voltage Electron Microscopy*, P.R. Swann, C. J. Humphrey, and M.J. Goringe, eds., Academic Press, London, 1974, pp. 230-34.
20. R.H. Goodenow, S. Matas, and R.F. Heheman: *Trans. AIME*, 1963, vol. 227, p. 651.
21. *Proc. Int. Conf. on Bainite at the World Materials Congress*, Chicago, 1988, appeared in *Metall. Trans. A*, 1990, vol. 21A, pp. 767-893 and 1343-1545.
22. H.K.D.H. Bhadeshia and J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, p. 767.
23. H.I. Aaronson, W.T. Reynolds, Jr., G.J. Shiflet, and G. Spanos: *Metall. Trans., A*, 1990, vol. 21A, p. 1343.
24. M. Enomoto and H. Tsubakino: *Mater. Trans. JIM*, 1991, vol. 32, p. 642.
25. P. Gustafson: *Scand. J. Metall.*, 1985, vol. 14, p. 259.
26. M. Hillert: *Recent Research in Cast Iron*, Gordon and Breach, New York, NY, 1968, pp. 101-27.
27. J. Ågren: *Acta Metall.*, 1989, vol. 37, p. 181.
28. K. Tsuzaki, K. Fujiwara, and T. Maki: *Mater. Trans. JIM*, 1991, vol. 32, p. 667.
29. J. Williams and S.G. Glover: *The Mechanism of*

Phase Transformations in Crystalline Solids, Institute of Metals, London, 1969, pp. 293-94.
30. J. Ågren: *Acta Metall.*, 1989, vol. 37, p. 181.
31. Y. Ohmori and T. Maki: *Mater. Trans. JIM*, 1991, vol. 32, p. 631.
32. J.W. Christian: *ICOMAT-79*, Proc. Int. Conf. on Martensite Transformations, Cambridge, MA, 1979, MIT Press, Cambridge, MA, 1979, pp. 220-34.
33. R.F. Bunshah and R.F. Mehl: *Trans. AIME*, 1953, vol. 193, p. 1251.
34. W. Steven and A.G. Haynes: *J. Iron Steel Inst.*, 1956, vol. 183, p. 349.
35. H.K.D.H. Bhadeshia and D.V. Edmonds: *Acta Metall.*, 1980, vol. 28, p. 1265.
36. H.K.D.H. Bhadeshia and D.V. Edmonds: *Metall. Trans. A*, 1979, vol. 10A, p. 895.
37. W.T. Reynolds, Jr., F.Z. Li, C.K. Shui, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, p. 1433.
38. M. Hillert: "Paraequilibrium," Internal Report, Swedish Institute of Metal Research, Stockholm, Sweden, 1953.
39. D.E. Coates: *Metall. Trans.*, 1973, vol. 4, p. 2313.
40. M. Hillert: *Monograph and Report Series No. 33*, Institute of Metals, London, 1969, p. 231.
41. Z.-K. Liu and J. Ågren: *Acta Metall.*, 1989, vol. 37, p. 3157.
42. M. Hillert: *Solid → Solid Phase Transformations*, H.I. Aaronson, D.E. Laughlin, R.F. Sekerka, and C.M. Wayman, eds., TMS-AIME, Warrendale, PA, 1982, p. 789.
43. K.R. Kinsman and H.I. Aaronson: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 39-53.
44. G. Purdy and M. Hillert: *Acta Metall.*, 1984, vol. 32, p. 823.
45. Z.-K. Liu and J. Ågren: "Theoretical Calculation of the Transformation of Austenite into Ferrite," paper presented at *ICOMAT-92*, Monterey, CA, July 20–24, 1992.

The Fine Structure and Formation Mechanism of Lower Bainite

G. SPANOS

Isothermal transformation of austenite to lower bainite was studied by optical microscopy and transmission electron microscopy (TEM), in two high-purity Fe-C-4 wt pct Mn-2 wt pct Si alloys containing 0.4 and 0.6 wt pct carbon, in order to elucidate the fine structure and formation mechanism of lower bainite. The present results support a mechanism for lower bainite formation presented previously in which lower bainite sheaves result from the formation of an aggregate of fine ferrite crystals with thin austenite "gaps" between them; carbide precipitation occurs within these austenite gaps. This mechanism accounts for the carbides oriented at an angle to the sheaf axis repeatedly observed in lower bainite but is inconsistent with models based on the precipitation of a high-volume fraction of carbides within highly supersaturated ferrite formed by high-velocity shear. Observations by a number of other researchers are reviewed and shown to include morphological features consistent with the present mechanism. Finally, the orientation relationships typically observed among ferrite, austenite, and carbides in lower bainite are reviewed and shown in most (but not all) cases to be consistent with the view that carbides precipitate in austenite at ferrite-austenite boundaries, also in agreement with the present model.

I. INTRODUCTION

SINCE publication of the seminal article by Davenport and Bain[1] in 1930, the ferrite-plus-carbide microstructure known as bainite has been widely accepted as an important product of austenite decomposition in steels. In 1939, Mehl[2] first used the terms "upper bainite" and "lower bainite" to distinguish between morphologies of bainite formed at higher and at lower temperatures, respectively. It is now generally accepted that upper bainite consists of ferrite laths (or plates) containing carbides nucleated between the laths and lying roughly parallel to the lath axis,[3,4] while lower bainite is composed of sheaves of ferrite plates in which the majority of the carbides are oriented in a single direction which lies at an appreciable angle to the sheaf axis.[5–10]

In a recent article by Spanos *et al.* based primarily on transmission electron microscopy (TEM) observations in Fe-C-2 pct Mn* alloys, a model was proposed for the formation of lower bainite in which carbides precipitate within austenite gaps, or films, separating fine individual ferrite crystals comprising a lower bainite sheaf.[11] On this model, the sequence of events which lead to the formation of the lower bainite sheaf include the following: (1) the initial formation of a largely carbide-free ferrite spine, (2) the formation of the fine ferrite crystals atop the spine, (3) precipitation of carbides in the austenite gaps between the ferrite crystals, presumably at ferrite-austenite boundaries, and (4) continued growth of the ferrite crystals which rapidly consumes the austenite in the gaps, quickly causing the sheaf to resemble a single monolithic plate (Figure 1).[11] The rapid ferrite growth in the final step of this process can occur as a consequence of the return of supersaturation in the austenite resulting from the carbide precipitation in the gaps.

Alternatively, perhaps a more widely accepted view to date has been that the microstructure shown in Figure 1(a) is formed *ab initio* as a single, monolithic ferrite plate (*e.g.*, Figure 2),[12]* and that the carbides precipitate from within the (presumably) highly supersaturated ferritic component of lower bainite which formed by high-velocity shear.[5,13,14,15]

The present investigation was undertaken in order to study in more detail the mechanism proposed by Spanos *et al.*[11] by employing Fe-C-2 pct Si-4 pct Mn alloys. This combination of alloying elements was chosen, in part, to retain the austenite matrix at room temperature without making the transformation kinetics unreasonably sluggish in the lower bainite temperature regime. The purpose of the Si is to inhibit carbide precipitation in order: (1) to prevent obscuration of the details of the fine structure of the ferritic component of lower bainite and (2) to avoid carbon depletion of the austenite matrix adjacent to carbides associated with lower bainite. The resulting increase in carbon concentration of the austenite adjacent to ferrite crystals helps, in addition to the 4 pct Mn, to suppress the M_s temperature[16] below room temperature in these regions. In this way, the proposed austenite gaps[11] can be observed directly, without transformation to carbides and ferrite at the isothermal transformation temperature, or transformation to martensite during the quench to room temperature (as occurs in the Fe-C-2 pct Mn alloys studied by Spanos *et al.*[11]).

*Unless otherwise specified, all compositions given in this article are in weight percent.

*For this reason, the term lower bainite "sheaf" will be used throughout this article to refer to the microstructural feature that in the past has been termed a lower bainite "plate".[3,5,10]

G. SPANOS, Metallurgist, is with the Physical Metallurgy Branch, Naval Research Laboratory, Washington, DC 20375-5343.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

U.S. GOVERNMENT WORK
NOT PROTECTED BY U.S. COPYRIGHT

Additionally, a brief review is made of observations by a number of other researchers of morphological features which can be explained by the mechanism of lower bainite formation described previously[11] and confirmed here. Finally, an analysis and review of the crystallography of carbide precipitation in lower bainite is presented in order to provide insight into the question as to whether the carbides precipitate in austenite or in ferrite.

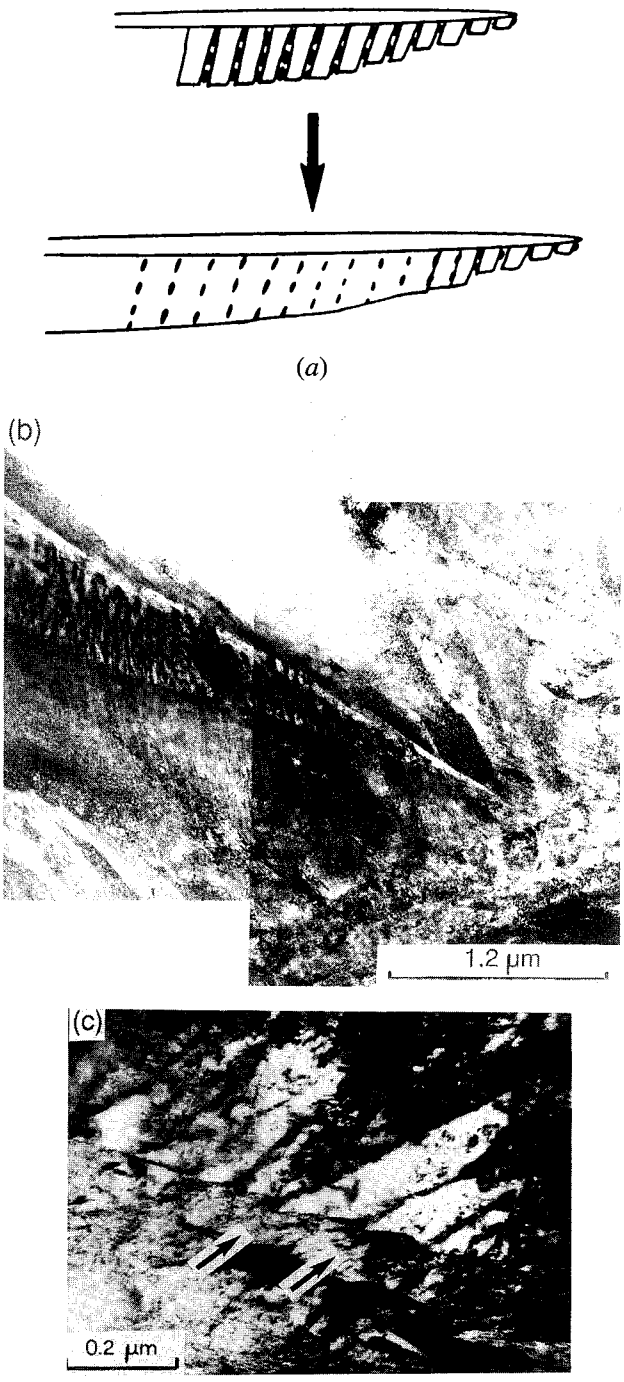


Fig. 1—(*a*) Schematic illustration of the mechanism of lower bainite formation proposed by Spanos *et al.*[11] (*b*) and (*c*) Bright-field TEM micrographs of lower bainite formed in an Fe-0.95 pct C-1.93 pct Mn alloy isothermally reacted at 250 °C for 15,000 s, reproduced from Ref. 11. Arrows in (*c*) highlight serrations near the tip of a lower bainite sheaf.

II. EXPERIMENTAL PROCEDURE

Two high purity Fe-C-Mn-Si alloys were used in this investigation; the compositions and the M_s temperatures, the latter determined by dilatometry, are listed in Table I. The alloys were prepared by vacuum induction melting of 45 kg heats, and the ingots were then hot rolled into bars with a cross section of about 0.04 × 0.04 m. After scale removal by grinding, 0.025 × 0.025 × 0.08 m bars of each alloy were encapsulated in quartz tubes containing 1/3 atm (at room temperature) of purified argon and homogenized for 3 days at 1250 °C. Chemical analysis was performed before and after homogenization to ascertain whether or not decarburization had occurred during this procedure; no significant loss of carbon was detected. Slices of 0.025 × 0.025 × 0.003 m were cut from the homogenized bars, annealed at 600 °C for 5400 seconds in a deoxidized molten lead bath to avoid cracking during subsequent rolling, and cold rolled to a thickness of 0.00025 m.

Individual specimens of 0.008 × 0.008 × 0.00025 m cut from the cold-rolled strips were then austenitized for 1500 seconds at 1200 °C in a deoxidized barium chloride salt bath,[17] isothermally reacted in a stirred, deoxidized Wood's metal bath at temperatures in the range of 200 °C to 350 °C, and finally quenched into room-temperature brine. Approximately 12 reaction times were employed at each reaction temperature in order to

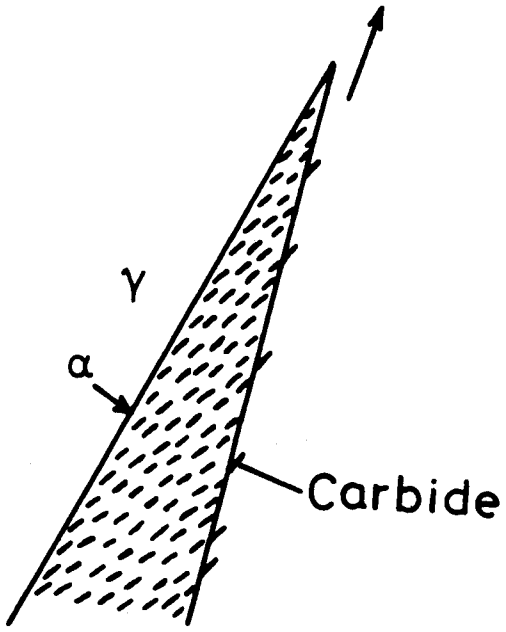


Fig. 2—Schematic illustration of a monolithic lower bainite plate. Taken from Ref. 12.

Table I. Alloy Compositions (in Weight Percent) and M_s Temperatures

C	Mn	Si	P	S	M_s (°C)
0.41	4.01	1.99	0.004	0.006	205
0.60	4.02	2.02	0.002	0.006	135

follow the transformation from beginning to end and identify suitable reaction times for TEM studies. Reaction times were chosen to yield relatively high-volume fractions of bainite. In this way, for bainitic growth by a diffusional mechanism, as supported by Aaronson and co-workers,[11,13,18–20] enough carbon would be rejected into the austenite matrix during the growth of the bainitic ferrite to suppress locally the M_s temperature below room temperature in the vicinity of the lower bainite and thus permit retention of large amounts of austenite.

Specimens for optical microscopy were etched with LePera's etchant.[21] Transmission electron microscopy specimens were prepared by chemically thinning heat-treated specimens to a thickness of 0.05 mm using a solution of 10 pct HF, 15 pct H_2O, and 80 pct H_2O_2. Disks 3 mm in diameter were then punched and thinned to perforation by jet polishing in a twin-jet FISCHIONE* unit with a solution consisting of 100 g of anhydrous sodium chromate dissolved in 500 ml of glacial acetic acid. The jet polishing was performed at room temperature with an operating voltage of 60 V, yielding a current between 20 and 25 mA.

*FISCHIONE is a trademark of Fischione Company, Pittsburgh, PA.

Transmission electron microscopy observations were performed on both a JEOL 200CX instrument operating at 200 kV and a PHILIPS* CM30 instrument operating at 300 kV.

*PHILIPS is a trademark of Philips Electronic Instruments Corp., Mahwah, NJ.

III. RESULTS

Although several different isothermal reaction temperatures were studied, the results for only a single temperature for each alloy are presented here. These representative temperatures were chosen to be low enough to ensure transformation to lower bainite, yet high enough above the M_s temperature to ensure no formation of athermal martensite during isothermal transformation, and also high enough to avoid formation of isothermal martensite at temperatures just above M_s.[22–25]

A. *Fe-0.4 pct C-4.01 pct Mn-1.99 pct Si reacted at 275 °C*

Optical micrographs of lower bainite reacted for two different transformation times are presented in Figure 3. The plate morphology typical of lower bainite is obvious at this level of magnification. Higher magnification TEM micrographs of these lower bainite sheaves are presented in Figure 4. In Figure 4(a), the dark features lying within the bright sheaves (S1) and at an angle to the axis of these sheaves are similar in appearance to typical lower bainitic carbides. However, they are actually regions of retained austenite, as evidenced by the dark-field micrograph of Figure 4(b) which was taken from a $(\bar{2}00)$ austenite reflection. Due to the high-silicon concentration in this alloy, carbide precipitation has been inhibited during this heat treatment, and the austenite gaps of the mechanism proposed by Spanos *et al.*[11] have now been directly imaged. Notice that many gaps within the three lower bainite sheaves (S1) are simultaneously brightly imaged from a single austenite reflection in Figure 4(b) because they have all formed within the same austenite grain. Only the ferrite in the lower bainite sheaf which appears very dark (S2) in the bright-field micrograph in Figure 4(a) is imaged brightly in the dark field in Figure 4(b). This is due to a ferrite reflection that was also included in the selected area diffraction aperture because it was immediately adjacent to (and partially overlapping) the austenite reflection used in Figure 4(b) (see the arrows in Figure 4(c)). It is thus obvious that this ferrite sheaf (S2) is of a different crystallographic orientation than the other three (designated S1 in Figures 4(a) and (b)). Although a long ferrite region predominantly free of the angular austenite gaps is observed along the top portion of the sheaf furthest to the right, and is pointed out by an arrow in Figure 4(a), clearly discernable ferrite "spines" with discrete ferrite-ferrite boundaries between the spine and the remainder of the sheaf, as reported by Spanos *et al.* in Fe-C-Mn alloys[11] (Figure 1(a)), were not frequently observed in the Fe-C-Mn-Si alloys studied here. It is unclear whether the presence of Si could affect the formation of the ferrite spines, or if the current lack of observation is largely apparent as a consequence of sectioning the thin foil in a plane above or below the spine.

The diffraction pattern corresponding to these micrographs is presented in Figure 4(c), with the austenite reflections indexed in the accompanying diagram. Another austenite pattern from gaps in bainite sheaves in a different region of the same specimen, corresponding to a $(\bar{2}1\bar{1})$ austenite zone axis in which the diffraction spots are more easily discernable, is presented in Figure 5.

The dark-field micrograph in Figure 6 shows many austenite gaps contained within multiple sheaves of lower bainite in a single austenite grain. Many similar observations were made in which austenite gaps within numerous sheaves were repeatedly imaged within a single austenite grain by using any of the individual austenite reflections in the corresponding diffraction pattern.

The sheaf near the right-center portion of Figure 6 demonstrates the presence of two different types of morphology of austenite gaps. Long austenite films lying roughly parallel to the sheaf axis are labeled A1 in this figure. A higher density of smaller austenite gaps which lie at an angle to the sheaf axis and are often continuous with the A1 gaps are designated A2 in Figure 6. These two types of gap morphology were often observed and will be referred to throughout the remainder of this article.

Bright-field micrographs of lower bainite formed in Fe-0.34 pct C-2 pct Mn (reproduced from Reference 11) and in the Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy studied here are presented in Figures 7(a) and (b), respectively, for comparison. In Figure 7(a), the gaps have already transformed to cementite and ferrite;[11] if any small amount of austenite did remain in the gaps after

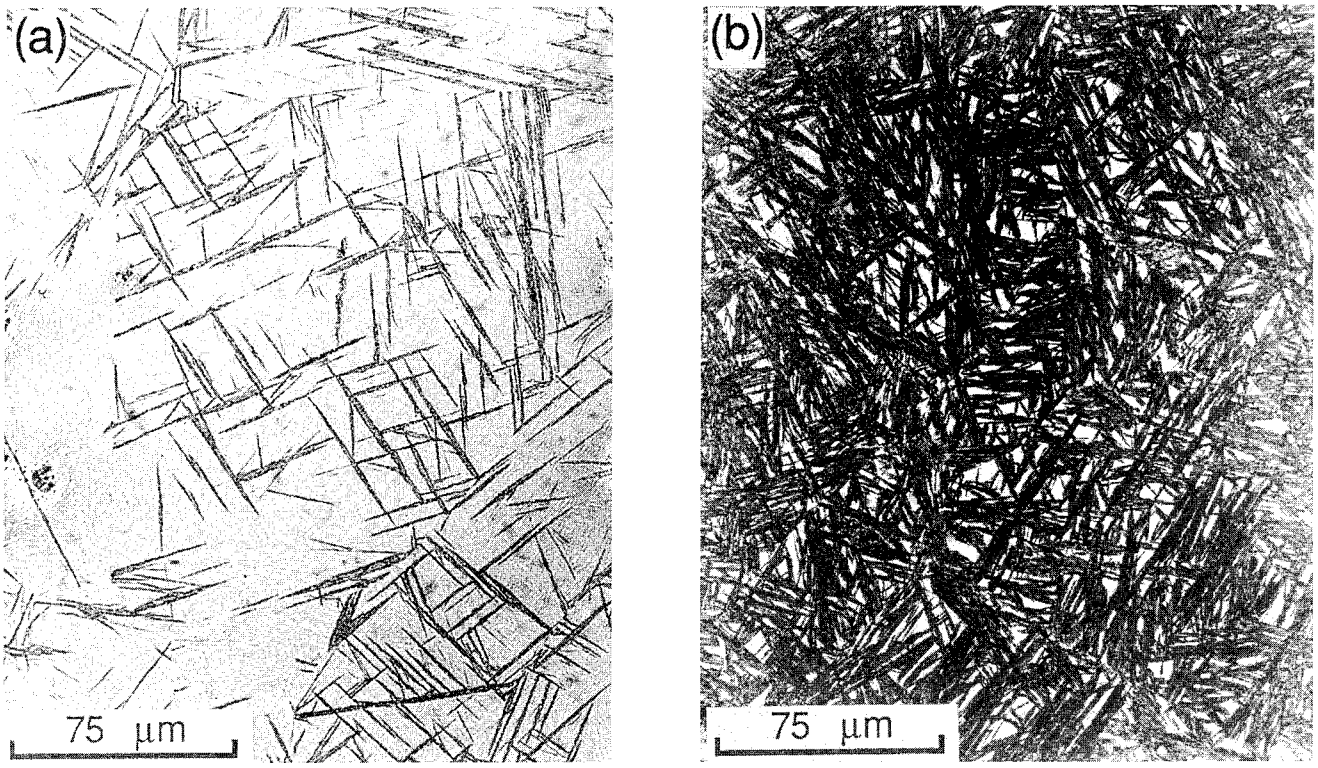


Fig. 3—Optical micrographs of bainite formed in an Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy reacted at 275 °C for (*a*) 20,000 s and (*b*) 40,000 s.

isothermal transformation, it would likely have transformed to martensite during the quench to room temperature in this alloy. On the other hand, in Figure 7(b), carbide precipitation has been inhibited due to the presence of Si in this alloy, and plenty of austenite remains in the gaps. The corresponding dark-field micrograph, taken from a $(\bar{1}1\bar{1})$ austenite reflection and presented in Figure 7(c), highlights the gaps shown in the bright field in Figure 7(b).

B. *Fe-0.60 pct C-4.02 pct Mn-2.02 pct Si Reacted at 200 °C*

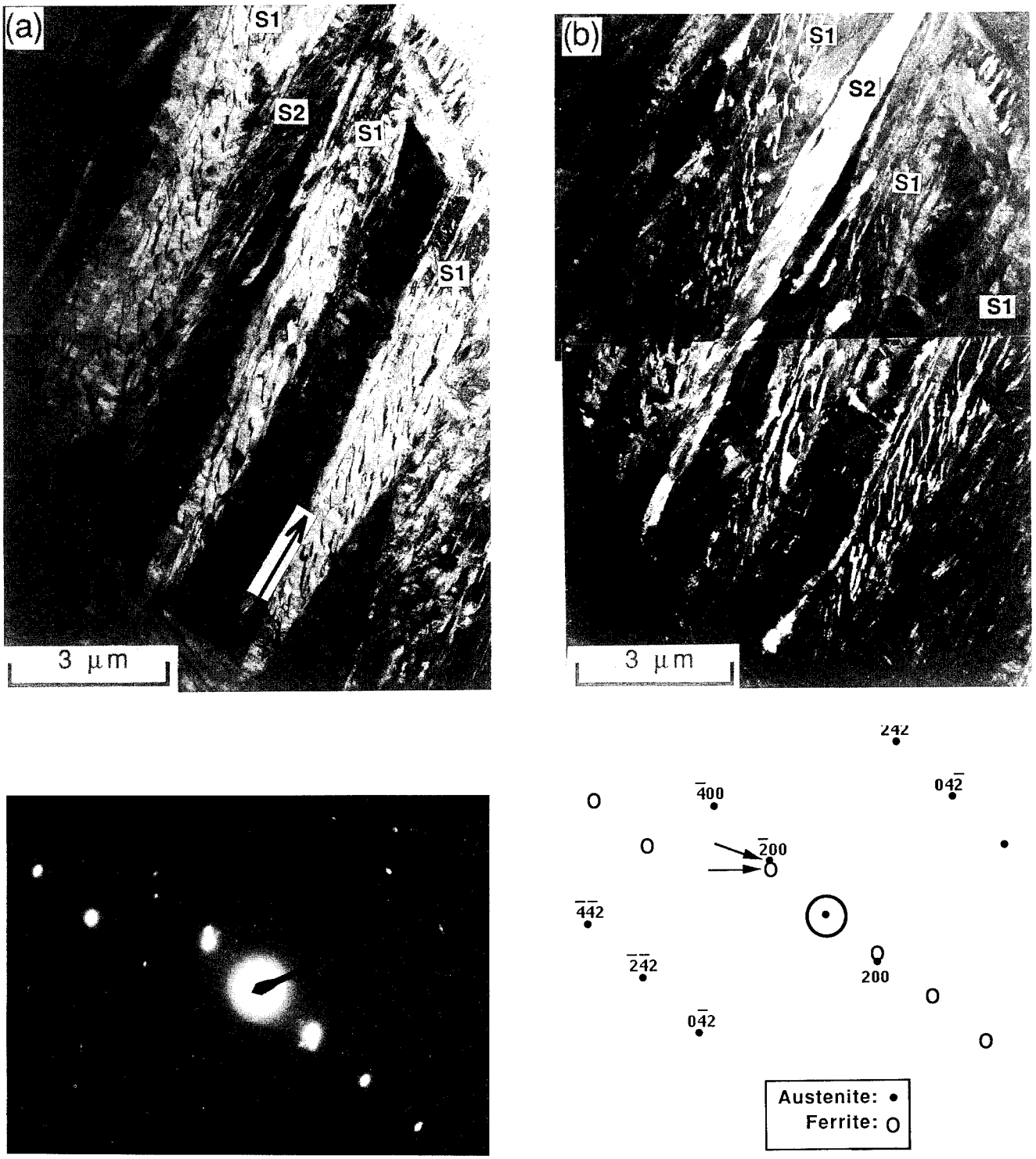
Optical micrographs of lower bainite in the higher carbon alloy reacted at 200 °C for two different transformation times are presented in Figure 8. The typical platelike morphology of lower bainite is again obvious when viewed at this level of magnification. Higher magnification TEM bright-field and dark-field micrographs from a group of sheaves are presented in Figure 9. Long films of retained austenite lying roughly parallel to the sheaf axis are again evident between lower bainite sheaves; typical examples (labeled A1) are shown in the dark-field micrograph presented in Figure 9(b). There are also a number of the somewhat shorter gaps continuous with these films but lying at a larger apparent angle to the sheaf axis. These are thinner and often very faint; some of the brighter ones are again labeled A2 in Figure 9(b). The microstructural features shown in Figure 9 are thus quite similar to those in lower bainite formed in the 0.41 pct C alloy and presented in Figures 4 and 6. Figure 9(b) also demonstrates that, just as in the 0.41 pct C alloy, a single reflection lights up gaps lying along different directions. Sheaves of more than one crystallographic variant of ferrite are apparently present in Figures 9(a) and (b), since only one variant of the sheaves shown in these micrographs is imaged in dark field from the ferrite reflection used in Figure 9(c). Additionally, at least some of the other bright areas shown in Figure 9(b) could have been due to ferrite and/or martensite spots overlapping the austenite reflection used in this dark-field micrograph.

An isolated lower bainite sheaf is presented in Figure 10. The fine structure of the sheaf is again demonstrated in these micrographs. The angular gap morphologies are pointed out by short arrowheads in Figures 10(a) and (b). In these micrographs of the higher carbon alloy, although the image contrast suggests the possibility that carbide precipitation occurred within the gaps, the diffraction patterns obtainable could not be unequivocally indexed as either epsilon carbide or cementite. It is possible that a large austenite film between sheaves may have been present in this case also, but any austenite in these regions apparently fell out of the foil during thinning, as evidenced by the large region void of material, running across the top of the sheaf and pointed out by a large horizontal arrow in Figures 10(a) and (b).

IV. DISCUSSION

A. *Comparison with Morphological Observations of Some Other Investigators*

Many previous studies in other alloy systems have revealed morphological features similar to those presently



(*c*)

Fig. 4—Lower bainite sheaves formed in an Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy reacted at 275 °C for 40,000 s. (*a*) Bright-field TEM micrograph, (*b*) corresponding dark-field micrograph taken from an austenite reflection, and (*c*) corresponding diffraction pattern. S1 and S2 designate sheaves corresponding to different crystallographic variants of ferrite.

observed, suggesting that the mechanism of lower bainite formation proposed earlier[11] and supported here may also be operative in systems other than just Fe-C-4 pct Mn-2 pct Si and Fe-C-2 pct Mn.[11] Some of these prior observations will now be briefly reviewed in the context of the present mechanism.

In a 1967 TEM investigation of an Fe-0.6 pct C-6.2 pct Ni alloy, Yada and Ooka[26] showed that the tips of lower bainite plates had a serrated edge, with "lathlike defects"[27] extending back from the serrations; carbide precipitation occurred in association with these defects. As pointed out previously,[28] the lathlike defects

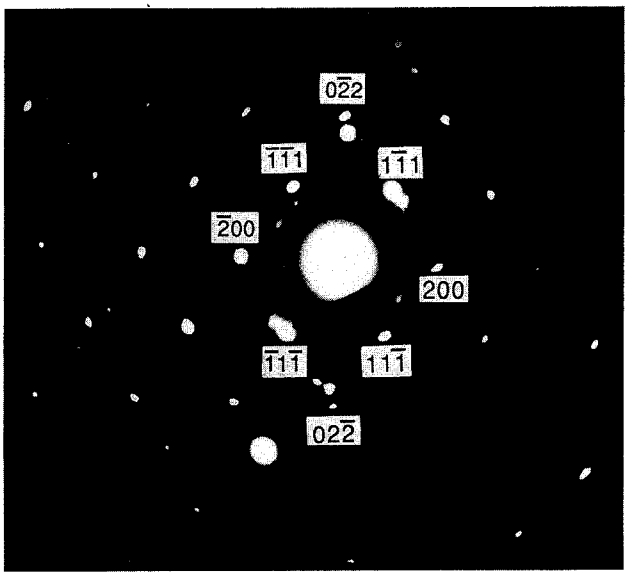


Fig. 5—Diffraction pattern corresponding to a $(\bar{2}1\bar{1})$ austenite zone axis. Taken from a different region (and different tilt conditions) of the same specimen presented in Fig. 4.

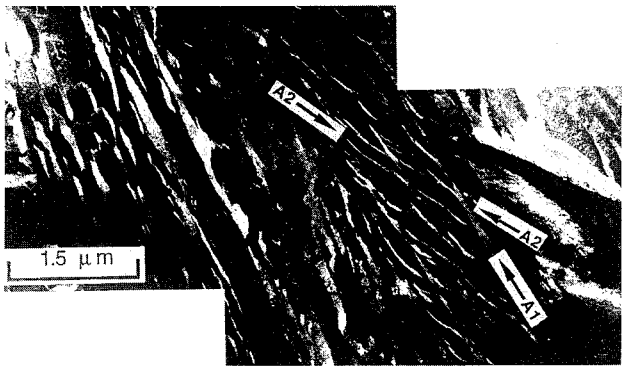


Fig. 6—Dark-field TEM micrograph taken from an austenite reflection in an Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy reacted at 275 °C for 100,000 s. Austenite gaps contained within multiple sheaves of lower bainite (within a single austenite grain) appear bright in this micrograph. A1 and A2 arrows point out austenite gaps (roughly) parallel and at an angle to the corresponding sheaf axis, respectively.



Fig. 7—Bright-field micrographs of lower bainite formed in (*a*) an Fe-0.34 pct C-2 pct Mn reacted at 350 °C for 40 s (reproduced from Ref. 11) and (*b*) an Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy reacted at 275 °C for 100,000 s, for comparison. (*c*) Dark-field TEM micrograph corresponding to (b), taken from a $(\bar{1}1\bar{1})_\gamma$ reflection.

observed by Yada and Ooka[26] may, in the context of the Spanos *et al.* mechanism,[11] correspond to thin austenite gaps which have been rapidly filled in with carbides and subsequently with ferrite.

In 1970, Hehemann reported that "Lower bainite plates generally exhibit one straight edge and thicken from one side only," as demonstrated in a 4360 (0.6 pct C) steel reacted at 345 °C.[5] In light of the new model, it would appear that the thickening side, or "rough" side, corresponds to the serrated side of the sheaf resulting from the presence of the gaps, and that the smooth side corresponds to the largely carbide-free ferrite spine.[11]

In an Fe-0.54 pct C-1.87 pct Si-0.79 pct Mn-0.3Cr alloy reacted at 315 °C, Huang and Thomas[8] observed "S-shaped" untransformed austenite regions in bainite formed at 315 °C. These regions resemble the austenite gaps observed here, though the gaps in the present study appear to be more uniformly oriented at a distinct angle to the sheaf axis. This may be related to the lower isothermal transformation temperatures (275 °C and 200 °C) employed here. Additionally, Huang and Thomas used crystallographic and morphological evidence to deduce that the carbides in bainite formed in the temperature range of 275 °C to 400 °C precipitated from within the austenite phase (in both the alloy just noted and in an Fe-0.4 pct C-1.73 pct Si alloy).[8]

In a more recent TEM study of a binary Fe-0.7 pct C alloy, Ohmori proposed a mechanism in which "Thickening of a bainite plate occurs by the side-by-side nucleation of thin bainite plates." Cementite nucleation was suggested to occur at ferrite-austenite boundaries bordering enriched austenite layers located at the tips of the thin ferrite crystals comprising the composite plate.[29] Thus, Ohmori's conclusion that lower bainite plates in an Fe-0.7 pct C alloy are actually a composite

of fine ferrite crystals at whose ferrite-austenite interfaces carbide nucleation occurs is in complete agreement with the observations presented in this article and can be understood in terms of the mechanism of Spanos *et al.*[11] On the other hand, Ohmori's proposed mechanism differs critically from the one supported here in two important ways. First, he did not consider (or observe) thin austenite gaps or films extending back into the lower bainite sheaves. This was most likely because in his alloy there was no silicon to inhibit carbide precipitation; therefore, once carbides promptly formed in the (presumed) gaps, supersaturation in the austenite should have quickly returned, allowing continued growth of ferrite to consume rapidly any remaining austenite in the gaps. Additionally, Ohmori was not able to retain any austenite at room temperature in the Fe-0.7 pct C alloy studied.[29] Consequently, even if a small amount of austenite were left in gaps after completion of isothermal reaction, it would have transformed to martensite during the quench to room temperature. These are presumably the same reasons why Spanos *et al.* were not able to retain the austenite gaps in their Fe-C-2 pct Mn alloys.[11] A second major difference in Ohmori's mechanism is that he suggested that the formation of the small ferrite crystals comprising a sheaf occurs by a displacive rather than a diffusional mechanism,[11] with some carbon supersaturation remaining in the ferrite.[29]

Very recently, Sun *et al.*[30] studied bainite formation in Fe-C-Mn-Si-Cr alloys reacted at 310 °C. They also observed retained austenite films between bainitic ferrite crystals, which they termed A_R^I (parallel to the overall sheaf) and A_R^{II} films (bending into the sheaf at a relatively small angle). Their A_R^I films closely resemble the large A1 austenite regions observed here and their A_R^{II} films somewhat resemble the A2 gaps currently observed, but their films typically lie at a smaller angle to the sheaf axis and are much fewer in number than those presently observed. These differences may be due in part to differences in composition between the alloys used in the two investigations (compare Table I here and Table I of Reference 30), since it is well known that substitutional alloying elements strongly affect the formation and morphology of ferrite in steels.[31,32,33] In contrast to the present study, Sun *et al.* repeatedly observed midribs in their bainite, most likely due to the fact that their isothermal reaction temperatures were very near the M_s temperatures (within 20 °C to 29 °C) in their alloys. Thus, the bainite they observed probably nucleated atop isothermal martensite midribs formed just above the M_s temperature; this type of phenomena has been characterized in detail by Okamoto and Oka[23,24,25] and was also observed earlier by Smith *et al.*[22]

In a study of an Fe-0.4 pct C-2 pct Si-3 pct Mn alloy, Bhadeshia observed ferrite-plus-austenite microstructures similar to the lower bainite reported by Sun *et al.*,[30] but he referred to these microstructures as upper bainite, even when they formed below 300 °C. At an isothermal reaction temperature of 286 °C, he reported the coexistence of this microstructure and carbide-containing bainite, the latter which he referred to as lower bainite. In the present investigation, one type of reaction product was observed for a series of isothermal transformation times studied in the lower bainite temperature range. This product was identified here as lower bainite (in which carbide precipitation was intentionally inhibited in order to study the austenite gaps) due to its morphological resemblance with lower bainite formed in many other alloys.[3,4,11,26,30,34]

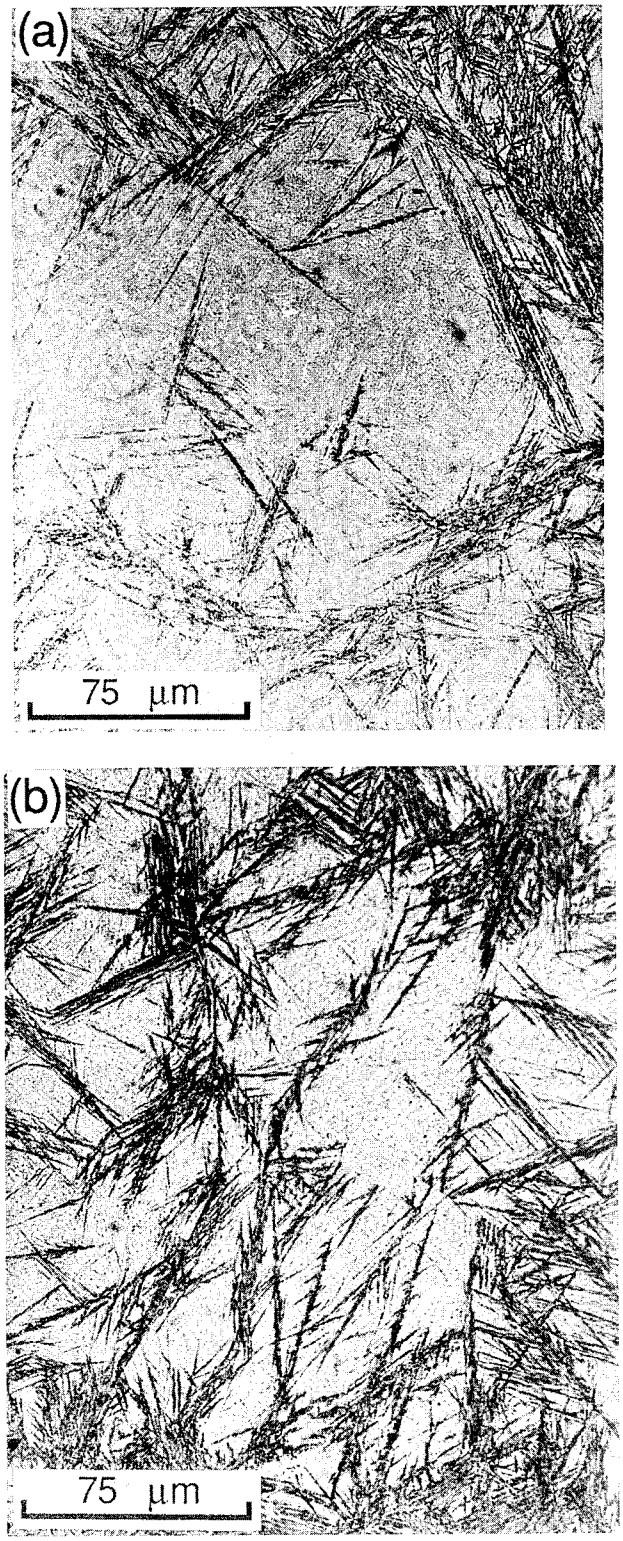


Fig. 8—Optical micrographs of bainite formed in an Fe-0.60 pct C-4.02 pct Mn-2.02 pct Si alloy reacted at 200 °C for (*a*) 858,240 s and (*b*) 1,096,860 s.

In a Ti-4.9 at. pct Cr alloy reacted at 540 °C, Enomoto

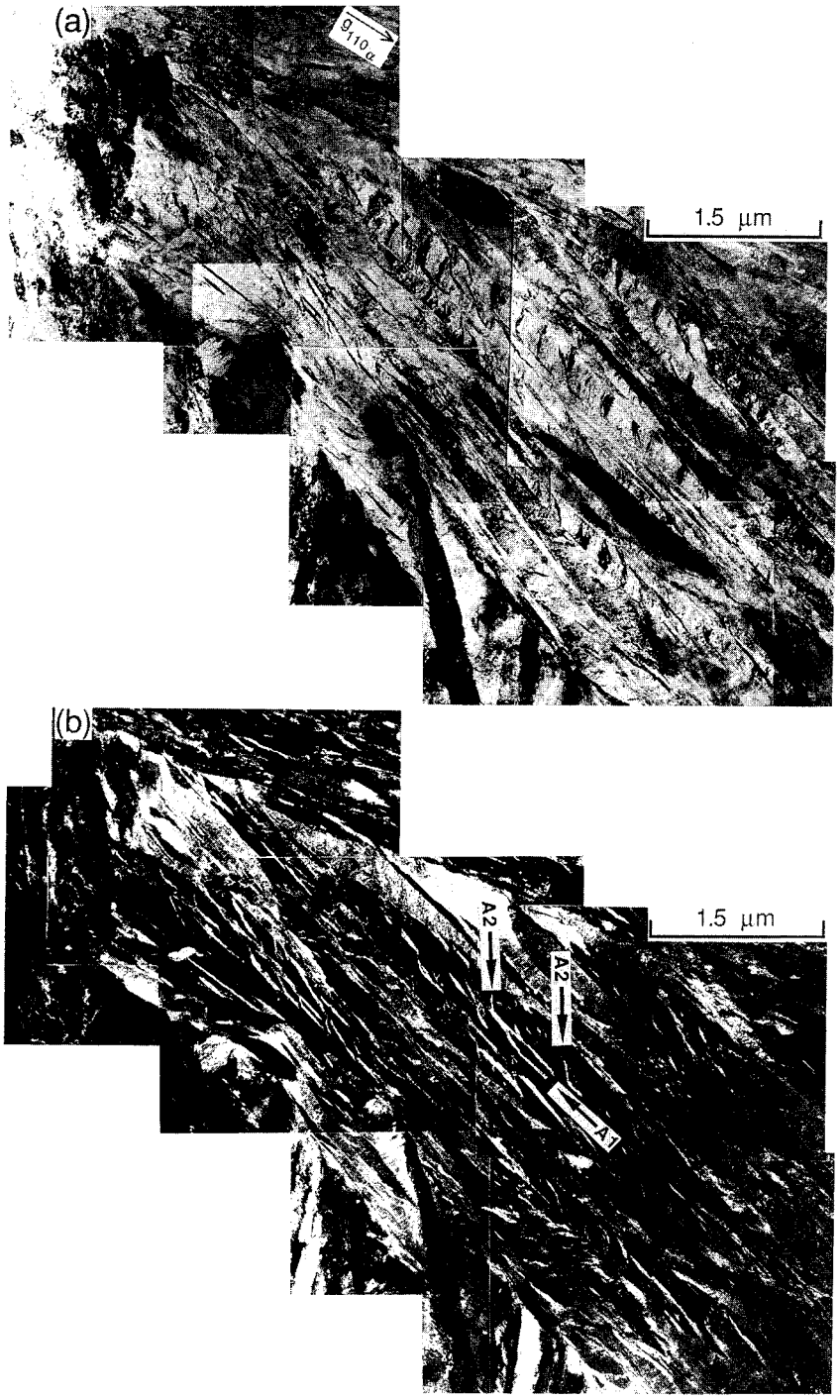


Fig. 9—Transmission electron microscopy micrographs of lower bainite sheaves formed in an Fe-0.60 pct C-4.02 pct Mn-2.02 pct Si alloy reacted at 200 °C for 2,439,060 s. (*a*) Bright-field TEM micrograph, (*b*) corresponding dark-field micrograph taken from an austenite reflection, and (*c*) corresponding dark-field micrograph taken from a ferrite reflection. A1 and A2 gap morphologies are indicated in (b).

and Fujita[35] studied lower bainite formation resulting from the eutectoid reaction: $\beta \rightarrow \alpha + TiCr_2$. They stated that a

> ". . . sheaf of thin plates is formed and thicken with time initially. . .As time is increased to an hour or so, the constituent plates begin to coalesce and finally form one large plate. The retained β matrix is engulfed by the α plate to from globular $TiCr_2$ compounds which appear just as if they were precipitated from the supersaturated large α plate. However a closer inspection soon reveals that these compound particles are forming a row which indicates the location of prior α/β interfaces at which thin plates were amalgamated."[35]

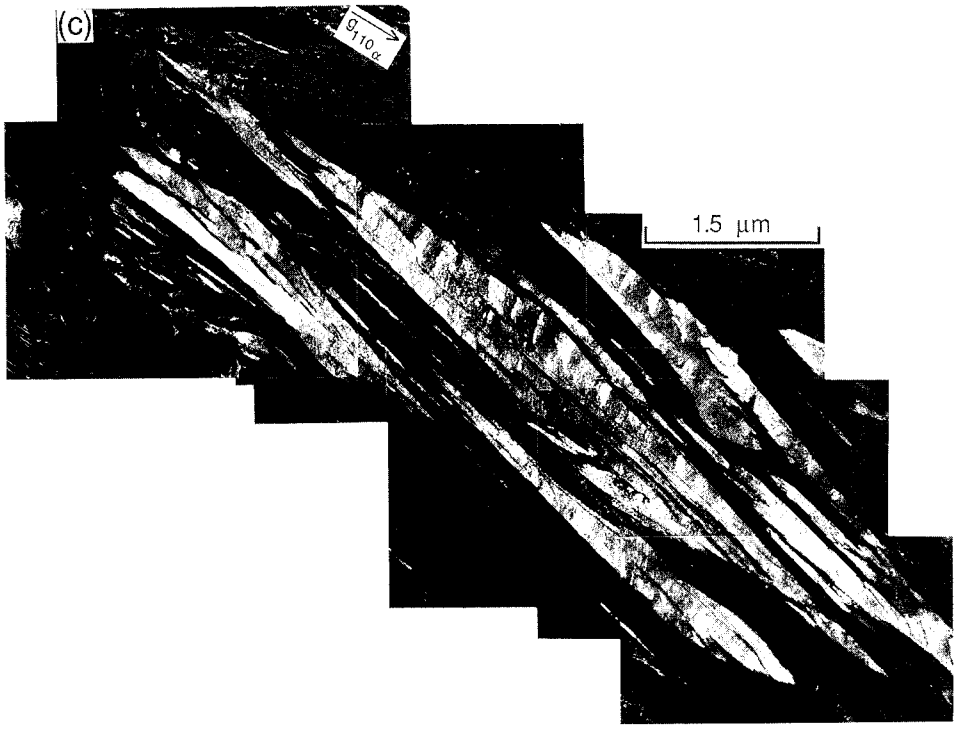


Fig. 9 Cont.—Transmission electron microscopy micrographs of lower bainite sheaves formed in an Fe-0.60 pct C-4.02 pct Mn-2.02 pct Si alloy reacted at 200 °C for 2,439,060 s. (*a*) Bright-field TEM micrograph, (*b*) corresponding dark-field micrograph taken from an austenite reflection, and (*c*) corresponding dark-field micrograph taken from a ferrite reflection. A1 and A2 gap morphologies are indicated in (b).

This mechanism is essentially identical to the one presented here. It thus appears that the mechanism of the lower bainite formation originally suggested by Spanos *et al.*[11] and supported by the present work may also be applicable to at least some nonferrous alloy systems.

It is not clear why discrete ferrite-ferrite boundaries sometimes observed by Spanos *et al.* between ferrite spines and the fine ferrite crystals comprising lower bainite sheaves in Fe-C-2 pct Mn alloys[11] were not observed in the present study. Instead, the long, thin ferrite regions occasionally observed here were continuous with, and imaged from the same dark-field reflections as, the fine ferrite crystals comprising the majority of the sheaf; the former regions were thus apparently of the same crystallographic variant as the fine ferrite crystals.

B. *Crystallography*

The crystallographic relationships among carbide, austenite, and ferrite in lower bainite are sometimes used to gain insight into whether the carbides precipitate from within a highly supersaturated ferrite or from within austenite. In this vein, unique orientation relationships (ORs) which provide good atomic matching between the precipitate and matrix lattices are generally obtained during solid-solid phase transformations in order to minimize the barrier to precipitate nucleation.[36] In the present context, if the carbides nucleate in austenite gaps (presumably at ferrite-austenite interfaces),[11] an austenite-carbide OR which allows for good atomic matching would be expected. Hence, a review and analysis of the crystallography of lower bainite will now be presented.

The ferrite-austenite (α-γ) OR observed in this study was usually within a few degrees of the Kurdjumov-Sachs[37] OR as demonstrated in the diffraction pattern presented in Figure 11. A common representation of this OR is

$$(110)_\alpha // (111)_\gamma$$
$$[1\bar{1}1]_\alpha // [0\bar{1}1]_\gamma \qquad [1]$$

This OR has been reported previously for lower bainite.[43] Occasionally, the Nishiyama-Wasserman[39,40] OR, which corresponds to a rotation of 5.26 deg from the Kurdjumov-Sachs OR about the normal to the parallel close packed planes, has also been reported for lower bainite.[41]

In the alloys investigate here, diffraction patterns from carbides were not obtained, mainly since carbide precipitation was avoided in order to retain and study the austenite gaps. Nevertheless, the Bagaryatskii[42] OR between ferrite and cementite, *c*, has been repeatedly observed for lower bainite.[38,41,43] This OR may be represented as

$$(110)_\alpha // (100)_c$$
$$[1\bar{1}1]_\alpha // [010]_c \qquad [2]$$

In order to study the three-phase crystallography among ferrite, cementite, and austenite in bainite, one can combine the operative α-γ and α-*c* ORs. Thus, using the specific variants of the ferrite, austenite, and cementite crystallographic relationships given by Eqs. [1] and [2], respectively, the three-phase crystallography

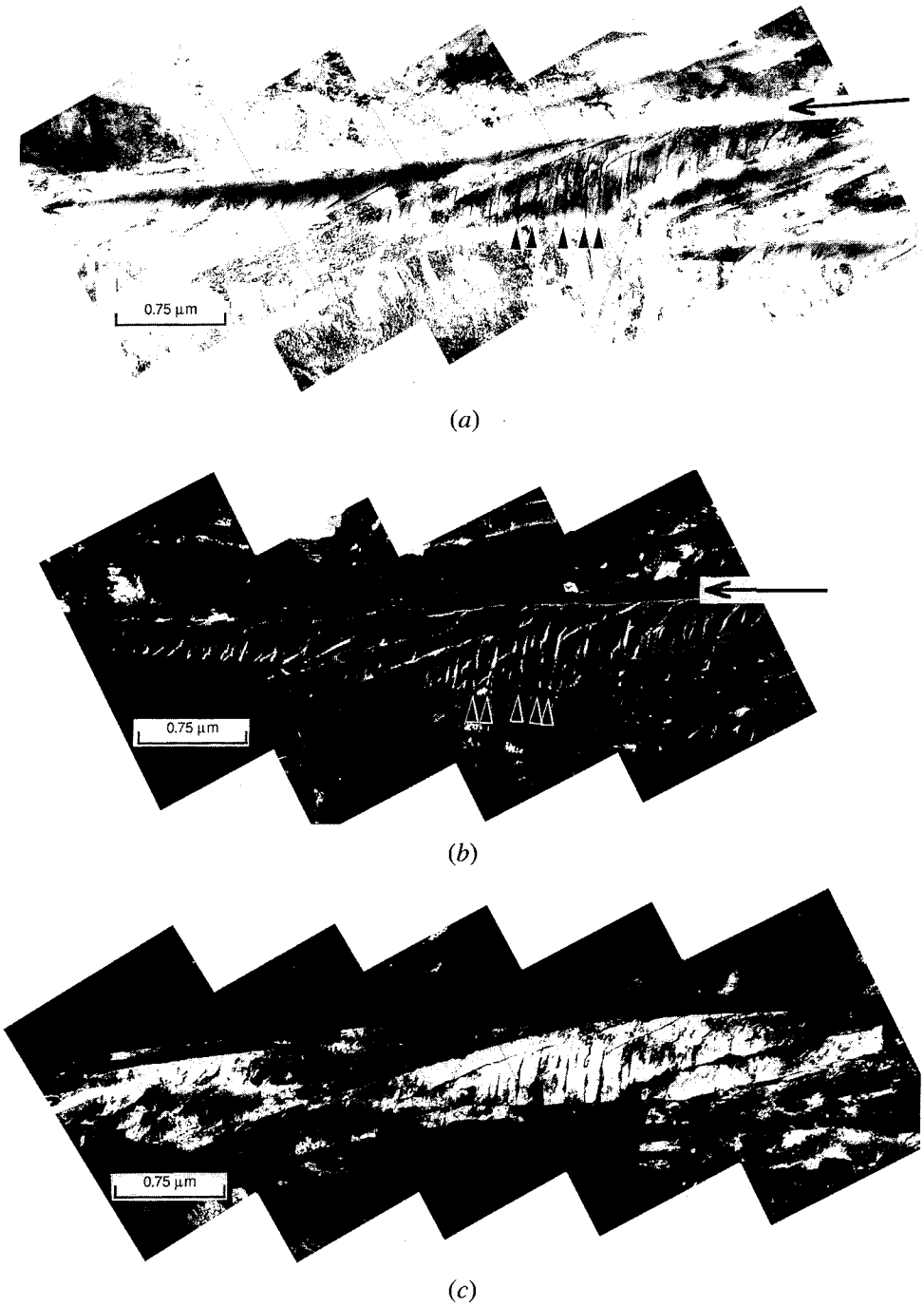


Fig. 10—Transmission electron microscopy micrographs of lower bainite sheaves formed in an Fe-0.60 pct C-4.02 pct Mn-2.02 pct Si alloy reacted at 200 °C for 1,096,860 s. (*a*) Bright-field TEM micrograph, (*b*) corresponding dark-field micrograph of the gaps, and (*c*) corresponding dark-field micrograph taken from a ferrite reflection. In (a) and (b), angular gap morphologies are pointed out by short arrowheads, while the large horizontal arrow indicates an area void of material running along the top of the sheaf.

represented in Figure 12 was generated on a computer. It can be seen that low index, closed-packed planes, and directions are parallel in all three crystals structures, *i.e.*, that the cementite exhibits reasonable matching with both the ferrite lattice and the austenite lattice. In particular, in a study of proeutectoid cementite plates, Spanos and Aaronson[44] used atom matching computer studies to show that the $[010]_c//\langle 110\rangle_\gamma$ parallelism, which is present in Figure 12, represents the direction of best fit between the cementite and austenite lattices. The cementite-austenite OR generated in the present analysis corresponds to a rotation from the appropriate variant of the Pitsch OR,[45,46] which is usually observed for proeutectoid cementite plates,[44,45,46] of about 5.8 deg about the $(010)_c//\langle 110\rangle_\gamma$ "good matching" direction. This rotation is represented in Figure 12 by the 5.8 deg angle between the $(455)_\gamma$ and the $(100)_c$ poles and between the $(\bar{5}\bar{2}\bar{2})_\gamma$ and the $(001)_c$ poles.

Whether the specific crystallographic variants described previously (and represented by the combination

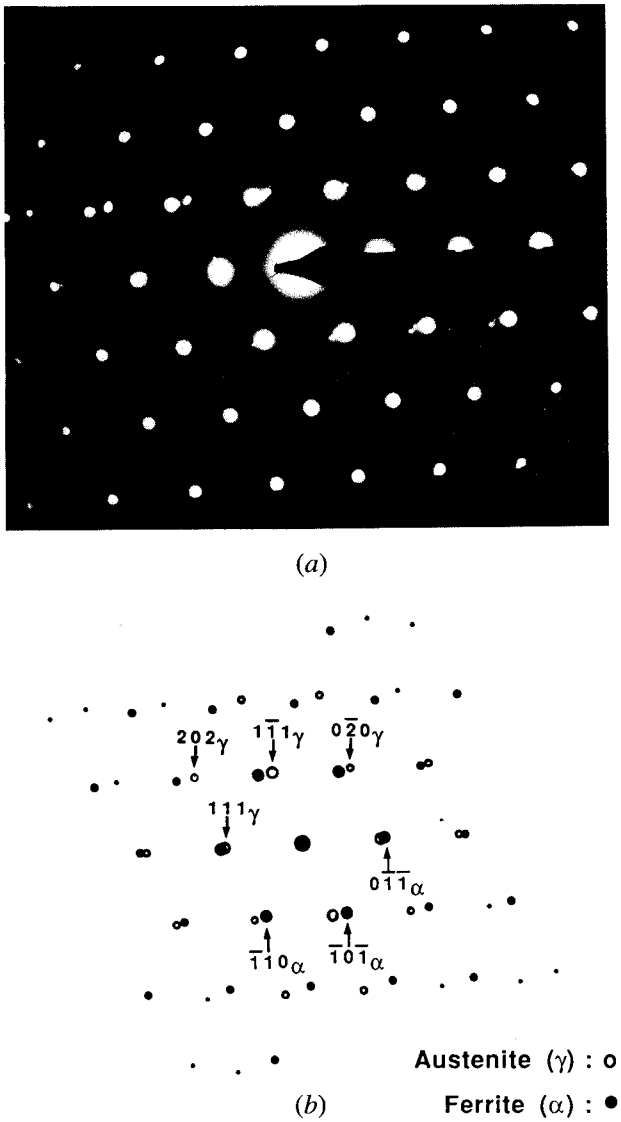


Fig. 11—(*a*) Diffraction pattern corresponding to the Kurdjumov-Sachs OR[37] between bainitic ferrite and austenite in the Fe-0.41 pct C-4.01 pct Mn-1.99 pct Si alloy reacted at 275 °C for 100,000 s. (*b*) Diffraction pattern generated by a computer assuming the Kurdjumov-Sachs OR and corresponding to the zone axis in (a).

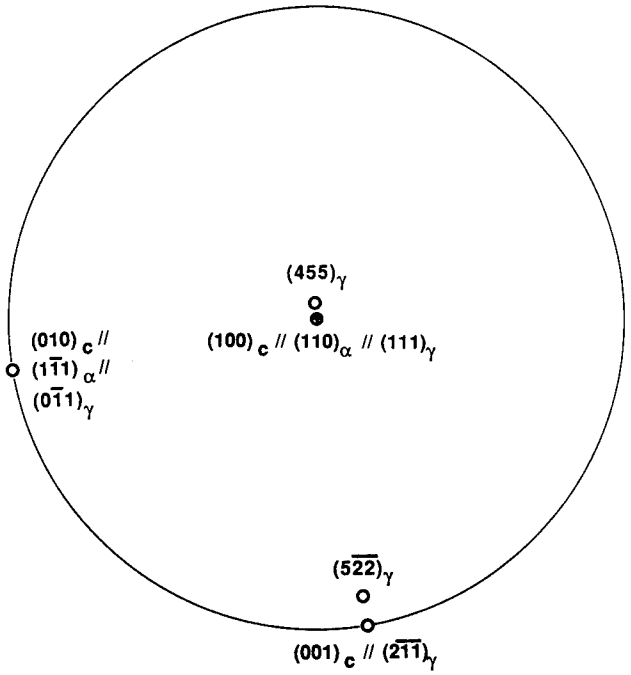


Fig. 12—[100]-cementite stereographic projection demonstrating the cementite, ferrite, and austenite crystallography for a specific ferrite variant in which the Bagaryatskii[42] ferrite-cementite O.R. is obeyed and the Kurdjumov-Sachs[37] ferrite-austenite O.R. is obeyed.

of Eqs. [1] and [2]) are actually operative must be determined experimentally in an alloy system in which all three phases are present during observation. This was done by Srinivisan and Wayman[43] for lower bainite formed in an Fe-1.1 pct C-7.9 pct Cr alloy; they determined experimentally that the particular variants observed resulted in the three-phase crystallographic matching represented in Figure 12. The present author also analyzed a published diffraction pattern from a study by Huang and Thomas (Figure 9 of Reference [8]) and found that for bainite formed at 315 °C in an Fe-0.54 pct C-1.87 pct Si-0.79 pct Mn-0.30 pct Cr alloy, the precise three-phase crystallography demonstrated in Figure 12 was again obeyed. Alternatively, Bhadeshia[41] has reported that for bainite formed at 435 °C in an Fe-0.3 pct C-4.08 pct Cr alloy, which he identified as lower bainite, the Nishiyama-Wasserman α-γ OR and the Bagaryatskii α-c OR are operative, and the actual variant obeyed does not provide any reasonable matching between the cementite and austenite lattices.[41,15]

Probably the most complete set of data on ferrite-cementite crystallography in bainite is that obtained by Shackleton and Kelly.[38] They showed that the Bagaryatskii O.R. was almost always obeyed between the ferrite and cementite lattices in lower bainite formed in Fe-C-Mn-Si-Cr-Mo-B steels.* Citing the fact that this O.R. is also observed for tempered martensite, they suggested that, by analogy, the cementite in lower bainite precipitates from within ferrite. In the upper bainite regime, they also observed a number of "new" relationships *in addition to the Bagaryatskii OR*. They demonstrated that by assuming the Pitsch O.R. between cementite and austenite[45,46] and the Kurdjumov-Sachs O.R. between ferrite and austenite, one could produce a number of ferrite-cementite ORs which fell into two categories: (1) those which obeyed (or were very close to) all of the observed new relationships and (2) those corresponding to the Bagaryatskii OR. Since both the new and the Bagaryatskii ferrite-cementite ORs observed for upper bainite could be produced in this way, they took this as evidence that the cementite in upper bainite precipitates in austenite, since the Pitsch cementite-austenite OR employed in their analysis is commonly observed for the precipitation of proeutectoid cementite in austenite.[44,45,46] It is important to note, though, that even in the upper bainite range, 23 of the 36 crystals

*The Pitsch ferrite-cementite OR[47] was also observed, but very rarely.[38]

which they analyzed obeyed the Bagaryatskii ferrite-cementite OR. Even with this predominance of the Bagaryatskii OR in upper bainite, these authors concluded, based on their crystallographic analysis, that upper bainitic cementite precipitates in austenite. This interpretation is directly opposite from their deduction that the observed Bagaryatskii O.R. supports carbide precipitation in ferrite in the case of lower bainite. It thus appears that observation of the Bagaryatskii α-c O.R. in lower bainite[38] could have been explained equally well in terms of precipitation in austenite.

The Isaichev ferrite-cementite OR[48] has also been reported occasionally, though less frequently, for lower bainite[34] and can be represented as follows:

$$(011)_\alpha // (103)_c$$
$$[11\bar{1}]_\alpha // [010]_c \qquad [3]$$

Combining the Kurdjumov-Sachs α-γ O.R. with the proper variant of the Isaichev α-c O.R., Huang and Thomas[8] produced a three-phase crystallographic correspondence similar to that generated by the computer here and presented in Figure 13. In this figure, the good fit $[010]_c // \langle 110 \rangle_\gamma$ parallelism[44] is again obtained. The austenite and cementite lattices are within about 1 deg of directly obeying the relevant variant of the Pitsch OR,[45,46] as demonstrated by the close proximity of the $(5\bar{4}5)_\gamma$ and the $(100)_c$ poles and the $(252)_\gamma$ and the $(001)_c$ poles in Figure 13. That is, the cementite exhibits the same lattice matching with austenite as that commonly observed for proeutectoid cementite plates.[44,45,46] Insofar as the present author is aware, no experimental observation in lower bainite of cementite, ferrite, and austenite coexisting and following the Isaichev α-c O.R. and the Kurdjumov-Sachs α-γ O.R. has been made. Thus, direct experimental determination of whether the proper combination of crystallographic variants just described is operative is not yet available.

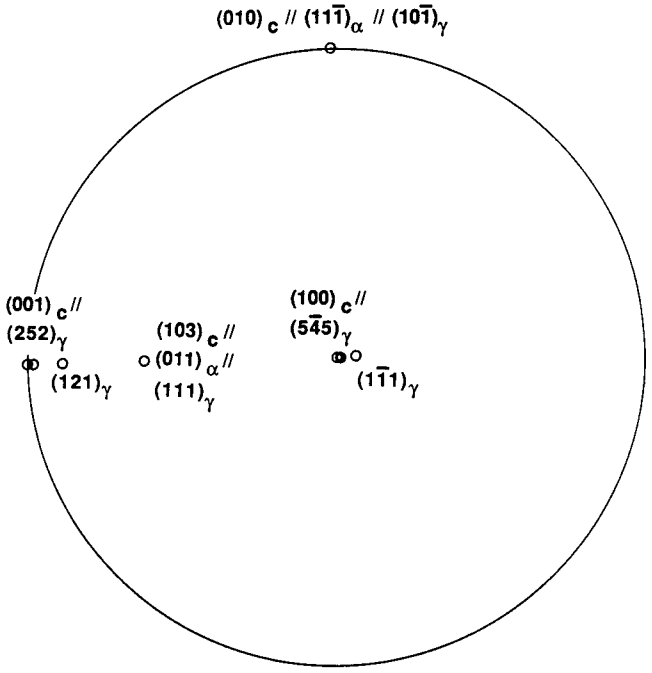


Fig. 13—[100]-cementite stereographic projection demonstrating the cementite, ferrite, and austenite crystallography for a specific ferrite variant in which the Isaichev[48] ferrite-cementite O.R. is obeyed and the Kurdjumov-Sachs[37] ferrite-austenite O.R. is obeyed.

Epsilon carbide (ε) has also sometimes been observed in association with lower bainite formed at low temperatures and often in alloys containing silicon. Huang and Thomas[8] observed the Jack[49] α-ε OR in lower bainite formed at 275 °C in an Fe-1.87Si-0.54C-0.79Mn-0.3Cr alloy; this O.R. is also observed for tempered martensite in Si-containing steels.[49] Huang and Thomas showed that by using the Jack α-ε O.R. and assuming the Kurdjumov-Sachs OR between the bainitic ferrite and austenite, the following ε-γ OR could be obtained:

$$(0001)_\alpha // (111)_\gamma$$
$$(1\bar{2}10)_\alpha // (1\bar{1}0)_\gamma \qquad [4]$$

They reported that this OR exhibited good matching between the epsilon carbide and austenite lattices; thus, observation of the Jack OR could equally be explained by precipitation of epsilon carbide in austenite. Unfortunately, they were not able to retain enough austenite in order to verify this suggestion.

In all cases analyzed earlier, in which good matching austenite-carbide ORs were either observed[8,38] or generated by computer, good matching ferrite-carbide O.R.s were also present. This is in accord with an analysis of nucleation and growth during interphase precipitation by Aaronson *et al.* in which they concluded that low-energy ORs with respect to both phases (austenite and ferrite in the present context) are required to make nucleation of the third phase (carbide) feasible at interphase interfaces (ferrite-austenite).[50] As far as the present author is aware, a direct TEM observation of carbides in the critical nucleus size range in lower bainite has not been made, *e.g.*, the carbides observed by Spanos *et al.*[11] had grown to sizes which had essentially filled the austenite gaps and had often impinged with one another. It is suspected, though, that the carbides precipitate heterogeneously in the austenite gaps at ferrite-austenite interfaces, in order (1) to decrease the barrier to nucleation afforded by the removal of the appropriate portion of the ferrite-austenite interface[50] and (2) to take advantage of any increased supersaturation (driving force for nucleation) in austenite at the interface; on the diffusional growth model,[11,13,18–20] this supersaturation results from rejection of carbon by ferrite into austenite during growth. Even in the event that some carbides might have nucleated wholly within the austenite in the gaps, though, one might still expect a three-phase ferrite-austenite-cementite OR due to the mutual ORs between austenite and ferrite and austenite and carbide. Nevertheless, Aaronson *et al.*[50] have concluded that attempts to determine from direct observation precisely where the critical nucleus formed may be a moot point, since embryos formed at an interphase interface may quickly grow about equally into both phases.*

*Concerning the potential effects of Si segregation on the carbide nucleation site, it has been demonstrated that no appreciable partitioning of substitutional alloying elements (such as Si) occurs between

austenite and ferrite and/or bainite at the low temperatures used here.[51,52] If there were *equilibrium* segregation of Si in austenite over even a few atom layers at the interface, it should not affect the driving force for carbide nucleation because it would not change the local chemical potential[53] (even though Si in general raises the activity of carbon in austenite[54]). If there were any nonequilibrium segregation[55] of Si to the interface, it might locally decrease the driving force for cementite precipitation due to the insolubility of Si in cementite;[56] it is not clear what effect this might have on the nucleation site of epsilon carbide.

Nearly all of the available evidence on the crystallography of carbide precipitation in lower bainite[8,34,38,43] can thus be explained equally well in terms of carbide precipitation in austenite (presumably at ferrite-austenite boundaries), rather than from within highly supersaturated ferrite. From observations in which the carbide, ferrite, and austenite phases were all simultaneously retained, the results of Srinivisan and Wayman[43] in an Fe-1.1 pct C-7.9 pct Cr alloy reacted at 285 °C and Huang and Thomas for bainite formed at 315 °C in an Fe-0.54C-1.87Si-0.79Mn-0.30Cr alloy[8] indicate that the proper crystallographic variants which allow for good matching between all three phases are observed. Alternatively, Bhadeshia concluded that no good matching obtains between the cementite and austenite phases for bainite formed at 435 °C in an Fe-0.3 pct C-4.08 pct Cr alloy.[41] It is presently unknown whether differences in alloy composition and/or formation temperature could be responsible for the differences between these results[41] and the morphological and crystallographic evidence presented (and reviewed) here.

In addition to the morphological studies presented in the previous section, the majority of the crystallographic findings just noted[8,34,38,43] can thus also be explained by the mechanism of bainite formation described earlier[11] and supported here, in which carbide precipitation occurs in small austenite gaps between fine ferrite crystals comprising lower bainite sheaves. This model differs significantly from the mechanism(s) adopted by the authors of all but one[8] of these previous crystallographic studies.[8,34,38,43]

V. SUMMARY AND CONCLUSIONS

Based upon observations in Fe-C-2 pct Mn alloys, Spanos *et al.*[11] proposed a mechanism for lower bainite formation in which carbides precipitate in austenite gaps between an assembly of small ferrite crystals that quickly grow together and resemble a large monolithic plate. The present investigation was performed in Fe-C-2 pct Si-4 pct Mn alloys containing 0.41 pct and 0.60 pct C in order to retain the austenite matrix at room temperature and directly test some important features of this model. A review was made of a number of previous investigations in other alloy systems in which morphological observations exhibiting similarities to those presented in this work were reported. An analysis and brief review of the crystallography of lower bainite were also undertaken. The following conclusions were reached:

1. The austenite gaps (or films) proposed by Spanos *et al.*[11] were directly imaged with bright-field and dark-field TEM techniques in lower bainite formed in Fe-C-4 pct Mn-2 pct Si alloys, providing direct support for the model that lower bainite is not composed of monolithic plates, but instead consists of an agglomeration of a number of fine ferrite crystals.
2. Two general types of austenite gap were observed: (A1) films of austenite between lower bainite sheaves and lying roughly parallel to the overall sheaf axis, and (A2) a large number of thin austenite gaps lying at a large angle to the sheaf axis and often continuous with the A1 gaps. Although these features are somewhat similar to the A_R^I and A_R^{II} films observed by Sun *et al.* in Fe-C-Mn-Si-Cr alloys reacted at 310 °C,[30] a much higher density of gaps lying at a large angle to the sheaf axis was observed in the alloys studied in the present investigation.
3. The A2 gaps separate small ferrite crystals comprising the large lower bainite sheaf. On the model of Spanos *et al.*,[11] these gaps account for the large volume fraction of carbides oriented at an angle to the sheaf axis repeatedly observed in lower bainite.[5–9]
4. The mechanism[11] supported by the current results is inconsistent with models based on the precipitation of a large volume fraction of carbides within highly supersaturated ferrite plates.[5,14,15]
5. Observations of a number of other researchers were shown to contain morphological features which can be explained by the present model of lower bainite formation.
6. The ORs typically observed between α, γ, and carbide in lower bainite are shown in most cases, to be consistent with carbide precipitation in austenite gaps, at α-γ boundaries.

ACKNOWLEDGMENTS

Financial support from the Office of Naval Research is gratefully acknowledged. Appreciation is expressed to Mr. M.L. Schofield for assistance with the isothermal heat treatments and the optical metallography. The author is grateful to Dr. J.F. Butler, Mr. E.L. Steib, and Mr. B. Yerek of LTV Steel Company, Independence, OH, for supplying the alloys. Sincere thanks are expressed to Dr. R.A. Vandermeer, Professor W.T. Reynolds, Jr., and Dr. H.I. Aaronson for critically reviewing the manuscript. The author wishes to dedicate this paper to the late Mr. W.E. King, Jr., who performed the dilatometry and the photographic reproduction for the manuscript. His quality workmanship and friendship will be greatly missed.

REFERENCES

1. E.S. Davenport and E.C. Bain: *Trans. AIME*, 1930, vol. 90, pp. 117-54.
2. R.F. Mehl: *Hardenability of Alloy Steels*, ASM, Metals Park, OH, 1933, p. 1.
3. F.B. Pickering: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 109-32.
4. J.M. Oblak and R.F. Hehemann: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 15-38.
5. R.F. Hehemann: *Phase Transformations*, ASM, Metals Park, OH, 1970, pp. 397-432.
6. Second Progress Report of Subcommittee XI of Committee E4: *Trans. ASTM*, 1950, vol. 50, pp. 444-92.

7. H. Modin and S. Modin: *Jernkontorets Ann.*, 1955, vol. 139, pp. 480-515.
8. D.H. Huang and G. Thomas: *Metall. Trans. A*, 1977, vol. 8A, pp. 1661-74.
9. G.Y. Lai: *Metall. Trans. A*, 1975, vol. 6A, pp. 1469-71.
10. R.K.W. Honeycombe: *Steels: Microstructure and Properties*, Edward Arnold (Publishers) Ltd. London and ASM, Metals Park, OH, 1981, p. 102.
11. G. Spanos, H.S. Fang, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1381-90.
12. D.A. Porter and K.E. Easterling: *Phase Transformations in Metals and Alloys*, Van Nostrand Reinhold Co., New York, NY, 1981, p. 339.
13. R.F. Hehemann, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1972, vol. 3, pp. 1077-94.
14. H.K.D.H. Bhadeshia and D.V. Edmonds: *Metall. Trans. A*, 1979, vol. 10A, pp. 895-907.
15. H.K.D.H. Bhadeshia and J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 767-97.
16. K.W. Andrews: *J. Iron Steel Inst.*, 1965, vol. 203, pp. 721-26.
17. J.R. Bradley, T. Abe, and H.I. Aaronson: *Rev. Sci. Instrum.*, 1982, vol. 53, pp. 98-99.
18. H.I. Aaronson: *The Decomposition of Austenite by Diffusional Processes*, V.F. Zackay and H.I. Aaronson, eds., Interscience, NY, 1962, pp. 387-546.
19. H.I. Aaronson, W.T. Reynolds, Jr., G.J. Shiflet, and G. Spanos: *Metall. Trans. A*, 1990, vol. 21A, pp. 1343-80.
20. W.T. Reynolds, Jr., H.I. Aaronson, and G. Spanos: *Mater. Trans. Jpn. Inst. Met.*, 1991, vol. 32, pp. 737-46.
21. F.S. Lepera: *Metallography*, 1979, vol. 12, pp. 263-68.
22. M.F. Smith, G.R. Speich, and M. Cohen: *Trans. AIME*, 1959, vol. 215, pp. 528-30.
23. H. Okamoto and M. Oka: *Metall. Trans. A*, 1985, vol. 16A, pp. 2257-62.
24. H. Okamoto and M. Oka: *Metall. Trans. A*, 1986, vol. 17A, pp. 1113-20.
25. M. Oka and H. Okamoto: *Metall. Trans. A*, 1988, vol. 19A, pp. 447-52.
26. H. Yada and T. Ooka: *J. Metall. Soc. Jpn.*, 1967, vol. 31, pp. 771-76.
27. H. Yada: *Metall. Trans. A*, 1991, vol. 22A, pp. 1674-76.
28. G. Spanos, H.S. Fang, and H.I. Aaronson: *Metall. Trans. A*, 1991, vol. 22A, pp. 1676-78.
29. Y. Ohmori: *Mater. Trans. Jpn. Inst. Met.*, 1989, vol. 30, pp. 487-97.
30. J. Sun, H. Lu, and M. Kang: *Metall. Trans. A*, 1992, vol. 23A, pp. 2483-90.
31. W.T. Reynolds, Jr., F.Z. Li, C.K. Shui, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1433-64.
32. W.T. Reynolds, Jr., F.Z. Li, S.K. Liu, S. Hartfield, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1479-92.
33. S.K. Liu, W.T. Reynolds, Jr., H. Hu, G.J. Shiflet, and H.I. Aaronson: *Metall. Trans. A*, 1985, vol. 16A, pp. 457-66.
34. Y. Ohmori: *Trans. Iron Steel Inst. Jpn.*, 1971, vol. 11, pp. 95-101.
35. M. Enomoto and M. Fujita: *Metall. Trans. A*, 1990, vol. 21A, pp. 1547-56.
36. W.C. Johnson, C.L. White, P.E. Marth, P.K. Ruf, S.M. Tuominen, K.D. Wade, K.C. Russell, and H.I. Aaronson: *Metall. Trans. A*, 1975, vol. 6A, pp. 911-19.
37. G.V. Kurdjumov and G. Sachs: *Z. Phys.*, 1933, vol. 64, p. 647.
38. D.N. Shackleton and P.M. Kelly: *Physical Properties of Martensite and Bainite*, Special Report No. 93, Iron and Steel Institute, London, 1965, pp. 126-34.
39. Z. Nishiyama: *Sci. Rep. Tohoku Univ.*, 1934, vol. 23, p. 368.
40. G. Wasserman: *Arch. Eisenhüttenwes.*, 1933, vol. 16, p. 647.
41. H.K.D.H. Bhadeshia: *Acta Metall.*, 1980, vol. 28, pp. 1103-14.
42. Y.A. Bagaryatskii: *Dokl. Akad. Nauk SSSR*, 1950, vol. 73, p. 1161.
43. G.R. Srinivisan and C.M. Wayman: *Acta Metall.*, 1968, vol. 16, pp. 609-20.
44. G. Spanos and H.I. Aaronson: *Acta Metall. Mater.*, 1990, vol. 38, pp. 2721-32.
45. W. Pitsch: *Acta Metall.*, 1962, vol. 10, pp. 897-900.
46. W. Pitsch: *Arch. Eisenhüttenwes*, 1963, vol. 34, pp. 381-90.
47. W. Pitsch: *Acta Metall.*, 1962, vol. 10, pp. 79-80.
48. I.V. Isaichev: *Zh. Tekh., Fiz.*, 1947, vol. 17, p. 835.
49. K.H. Jack: *J. Iron Steel Inst.*, 1951, vol. 169, p. 26.
50. H.I. Aaronson, M.R. Plichta, G.W. Franti, and K.C. Russel: *Metall. Trans. A*, 1978, vol. 9A, pp. 363-71.
51. H.I. Aaronson and H.A. Domian: *Trans. AIME*, 1966, vol. 236, pp. 781-96.
52. W.T. Reynolds, Jr., M. Enomoto, and H.I. Aaronson: *Phase Transformations in Ferrous Alloys*, A.R. Marder and J. Goldenstein, eds., TMS-AIME, Warrendale, PA, 1984, pp. 155-200.
53. K.C. Russell: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 1933-39.
54. J.S. Kirkaldy, B.A. Thomson, and E.A. Baganis: *Hardenability Concepts with Applications to Steel*, D.V. Doane and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, pp. 82-125.
55. K.R. Kinsman and H.I. Aaronson: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 39-53.
56. W.S. Owen: *J. Iron Steel Inst.*, 1951, vol. 167, pp. 117-20.

Morphology of Bainite and Widmanstätten Ferrite

YASUYA OHMORI, HIROYUKI OHTSUBO, YUN CHUL JUNG,
SHUJI OKAGUCHI, and HIROO OHTANI

Morphology of bainite and Widmanstätten ferrite in various steels has been investigated by means of microstructural and surface relief observations. It was shown that upper and lower bainite should be classified by ferrite morphology, *i.e.*, lathlike or platelike, and that the morphology of cementite precipitation cannot be the index for the classification. Widmanstätten ferrite formed in the upper C-nose where ferrite grain-boundary allotriomorphs nucleate exhibits quite similar appearance with bainitic ferrite that forms in the lower C-nose of bainitic reaction. The only difference between them exists in the fact that Widmanstätten ferrite laths grow in the temperature range where primary ferrite forms and often terminate at a grain boundary ferrite but that bainitic ferrite has its own C-curve at temperatures below B_s and nucleates directly at an austenite grain boundary. The mechanisms for their formations are discussed.

I. INTRODUCTION

PLATELIKE or lathlike ferrite transformation products formed at temperatures above M_s are often referred to as Widmanstätten ferrite, acicular ferrite, bainitic ferrite, and bainite. Since such terminology creates confusion[1] in the discussion of their formation mechanisms, in the present study, Widmanstätten ferrite and bainite are defined as follows.

(1) Widmanstätten ferrite is the platelike or the lathlike ferrite grown at temperatures between (A_{e3}) and B_s where diffusional transformations are thought to occur.
(2) Bainite is a similar transformation product, but it forms between B_s and M_s temperatures. Thus, bainitic ferrite is the carbide-free bainite formed at relatively higher temperatures in this range.

Although Widmanstätten ferrite and bainitic ferrite have thus been classified, it has been well recognized that both structures are quite similar in morphology and crystallography;[2] *i.e.*, both are lathlike ferrite with the habit plane lying close to the region between {112} and {111} austenite planes[3–6] and accompany surface reliefs when they form.[6–10] It is also of interest that the Widmanstätten ferrite laths can grow continuously from upper bainite laths by raising the temperature during isothermal transformation.[2]

The aims of this article are, therefore, to clarify the fundamental difference between both structures and to elucidate their formation mechanisms.

YASUYA OHMORI, Professor, HIROYUKI OHTSUBO, Research Associate, and YUN CHUL JUNG, Research Student, are with the Department of Materials Science and Engineering, Ehime University, Matsuyama, 790, Japan. SHUJI OKAGUCHI, Research Engineer, and HIROO OHTANI, General Manager, are with Iron and Steel Research Laboratories, Sumitomo Metal Industries, Amagasaki 660, Japan.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

II. EXPERIMENTAL PROCEDURE

Materials used in the present study are shown in Table I. Steel A is an extremely low-carbon steel containing 3 pct Mn and 1 pct Cr. Steels B and C are 0.2 pct C-1.37 pct Cr-0.59 pct Mo and 0.23 pct C-1.86 pct Mn steels, respectively. Steels D through F are commercial Si-Mn steels containing 0.17 to 0.58 pct C. Steel G is a high-carbon steel containing 1 pct C and 1 pct Cr. They were received in as-forged or as-rolled condition and were machined into 1.5-mm thick, 5-mm wide, and 15-mm long platelets.

Heat treatments were carried out as follows. The specimens were austenitized at temperatures from 1100 °C to 1250 °C for 10 minutes in a dynamic nitrogen atmosphere and were quenched into either a lead or an alloy bath kept isothermally at temperatures from 700 °C to 260 °C. After holding for various times, the specimens were quenched into an iced brine to suppress further isothermal decomposition. Microstructures of these specimens were observed by means of optical, scanning electron, and transmission electron microscopy. In order to construct T-T-T diagrams, the amounts of transformation products were measured by an image analyzer (linear analysis method) using 5 micrographs with magnification 400 times taken from the regions where the effects of decarburization can be ignored for each specimen. Statistical errors have not been calculated, because the results also involve errors due to experimental problems such as very short heat-treatment time. Some of the specimens were polished and encapsuled in evacuated silica tubes prior to the heat treatments, and the surface reliefs formed by the heat treatments were examined by optical microscopy.

Another type of heat treatment was also done: *i.e.*, the isothermal-decomposition temperature was changed discretely from lower to higher temperature after partial transformation at the lower temperature, as has been done previously by Purdy and Hillert.[2] By this treatment, the growth of the lower temperature decomposition products at higher temperature was examined.

Table I. Chemical Compositions of Materials Used (Weight Percent)

Steel	C	Si	Mn	Cr	Mo	Nb	Ti	B	Al
A	0.0041	0.19	2.98	1.00	—	0.010	0.010	—	0.018
B	0.20	0.30	0.97	1.37	0.59	—	0.015	0.0010	0.061
C	0.23	0.25	1.86	—	—	—	0.018	—	0.042
D	0.17	0.25	0.78	—	—	—	—	—	—
E	0.34	0.32	0.60	—	—	—	—	—	—
F	0.58	0.28	0.59	—	—	—	—	—	—
G	1.03	0.31	0.52	1.01	—	—	—	—	—

III. RESULTS

Figures 1 through 3 illustrate the T-T-T diagrams for the commercial Si-Mn steels E and F, the 0.23 pct C-1.86 pct Mn steel C, and the 1 pct C-1 pct Cr steel G, respectively. The Si-Mn steels exhibit single C-curves as far as the starting points of transformation are concerned. The curves connecting the final points of transformation, however, are not simple C-curves and can be interpreted as consisting of two different reactions separated at about 520 °C. In the upper temperature range, the ferrite/pearlite structure forms. The bainitic reaction is recognized in the lower temperature range. In the case of the alloy steels C and G, on the other hand, two separate C-noses can be observed. In the upper C-nose, the grain-boundary ferrite allotriomorphs or the pearlite colonies nucleate at the austenite grain boundaries in steel C or G, respectively. Within the lower C-nose, bainite is observed to form.

In the hypoeutectoid steels, Widmanstätten ferrite laths often grow from the grain-boundary ferrite allotriomorphs within the upper C-nose. Although Widmanstätten ferrite laths and the bainitic ferrite laths formed in the lower C-nose are very similar in appearance, the only distinct difference recognized between them is in the nucleation characteristics, as described later in detail: *i.e.*, Widmanstätten ferrite laths are frequently terminated at the grain-boundary, ferrite allotriomorphs, while those of bainitic ferrite nucleate specifically at the austenite grain boundaries.

Although some other differences exist, such as dislocation density in ferrite, lath thickness, and sharpness of the ferrite/austenite interfaces, they vary quite continuously with transformation temperature, and no discrete change between the two C-noses can be observed.

Microstructures of the isothermally transformed steel C are shown in Figure 4. T_0 temperature for the Fe-0.23 pct C-1.86 pct Mn alloy is calculated to be about 670 °C by the Hillert-Staffanson model[11] using the data of Uhrenius.[12] At 650 °C, which is significantly lower than

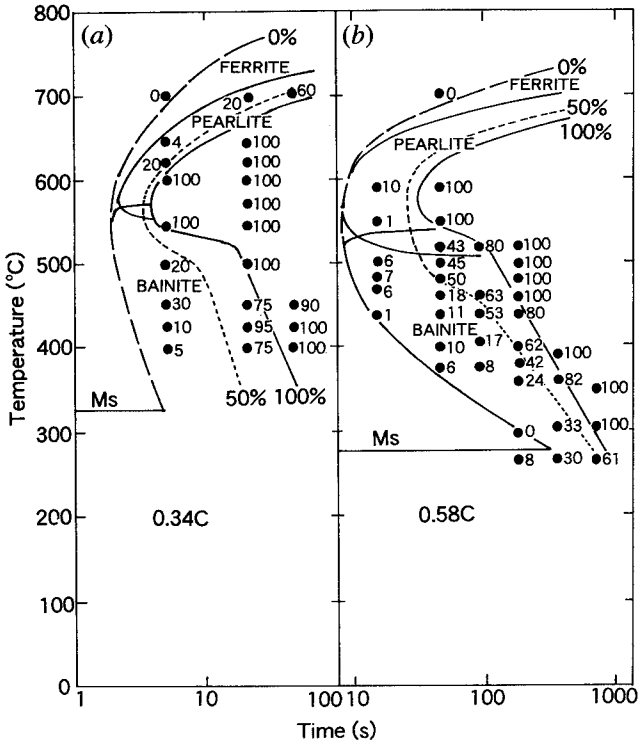


Fig. 1—T-T-T diagrams for the commercial Si-Mn steels, austenitized at 1100 °C for 10 min: (*a*) steel E (0.34 pct C) and (*b*) steel F (0.58 pct C).

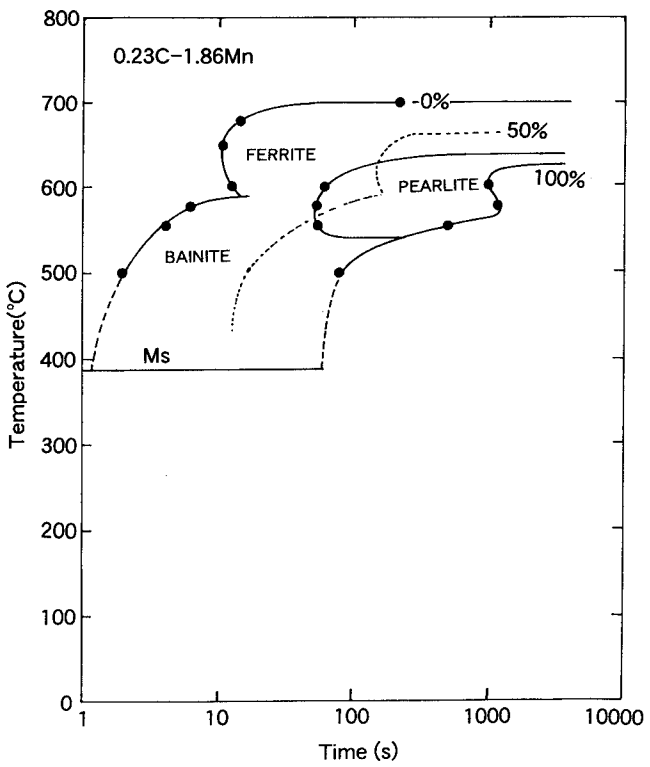


Fig. 2—T-T-T diagram for steel C (0.23 pct C-1.86 pct Mn), austenitized at 1200 °C for 10 min.

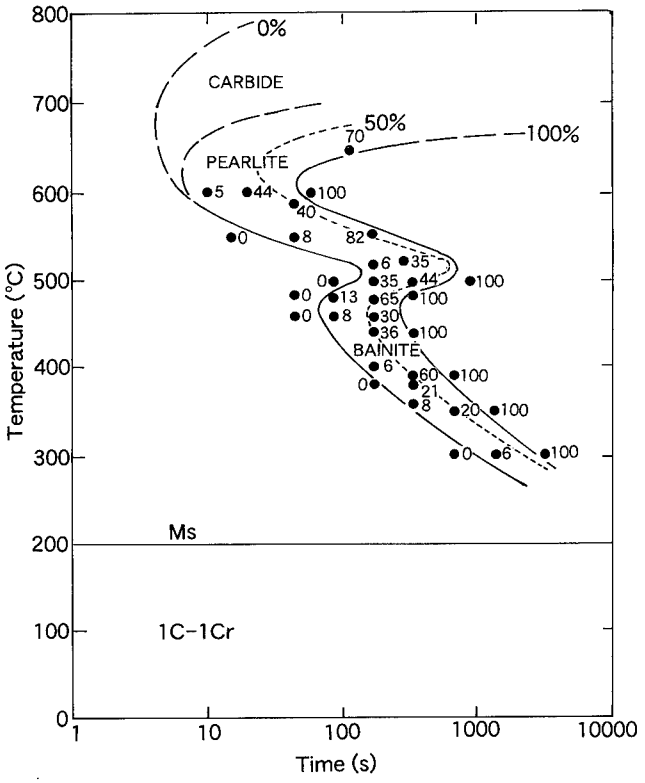


Fig. 3—T-T-T diagram for steel G (1.03 pct C-1 pct Cr), austenitized at 1100 °C for 10 min.

the T_0 temperature, ferrite particles nucleate at an austenite grain boundary, as shown in Figure 4(a), and grow into the grains separated by the boundary as grain-boundary ferrite allotriomorphs during isothermal holding, as in Figure 4(b), the austenite grain boundaries being enveloped by ferrite networks. At 600 °C above B_s temperature, Widmanstätten ferrite laths often grow from the grain-boundary ferrite allotriomorphs as in Figure 4(c), and the coalescence of them leads to the formation of much thicker ferrite plates in a later stage of transformation, as in Figure 4(d). In these structures, the prior austenite grain boundaries cannot be preserved after a complete transformation. On the other hand, at temperatures lower than B_s, 500 °C, the nucleation of bainitic ferrite laths occurs directly at the austenite grain boundaries, as in Figure 4(e). It can be seen that very sharp bainitic ferrite laths nucleate at an austenite grain boundary and that their growth is ended at the annealing twin boundary existing within the austenite grain. With isothermal holding at this temperature for a longer time, cementite particles precipitate between the ferrite laths, and the interfaces of the ferrite laths lose sharpness, suggesting that the interfaces are also migrated slightly by atomic diffusion.

Figure 5 shows the upper and lower bainite formed in the commercial Si-Mn steels containing 0.58 pct C, steel F. At 520 °C, typical upper bainite consisting ferrite laths with cementite particles between them are formed in a side-by-side fashion, and pearlitic reactions, also subsequently occur, as in Figure 5(a). Lowering the decomposition temperature to 280 °C yields platelike lower bainite, as can be seen in Figure 5(b).

Figure 6 shows the surface reliefs for the ultra low-carbon steel A transformed at 600 °C (Figure 6(a)) and the corresponding microstructure revealed by slight polishing and etching in nital (Figure 6(b)). These surface reliefs are due to the Widmanstätten ferrite formation, as can be seen in Figure 6(b), but are fine and sharp as compared with the corresponding microstructure: *i.e.*, a Widmanstätten ferrite lath comprises several parallel surface reliefs, implying that coalescence of the initially formed ferrite sublaths occurs. It is also interesting to note that the region where a polygonal ferrite grain formed consists of two different sets of lathlike surface reliefs, as shown by the arrow *A*.

Figure 7[13] shows the Widmanstätten ferrite lath formed in steel C transformed at 600 °C. Longitudinal morphology of the lath almost normal to the habit plane is illustrated in Figure 7(a). A small-angle grain boundary can be observed to lie parallel to the growth direction of the lath, as indicated by the arrows. The cross section of such a Widmanstätten ferrite lath is illustrated in Figure 7(b). The shape of it is rather complicated and is thought to be formed by the coalescence of several simple sublaths produced in a sympathetic fashion. A similar process of bainite formation in steel B can be seen in Figure 8.[14] Figure 8(a) shows the dark-field image of the cross sections of the sublaths of which coalescence produces bainite laths. The growth directions of these sublaths are nearly normal in this figure. If cementite platelets nucleate between such sublaths lying on the planes parallel to a specific plane, the upper bainite laths containing cementite platelets of a specific orientation will form. Further coalescence of bainite laths with the same habit will result in thicker bainite laths, as in Figure 8(b).

The early stage of lower bainite formation in steel *F* at 280 °C is shown in Figure 9. It has been well established that lower bainite is platelike with an irrational habit plane. Each plate, however, consists of very thin parallel plates, as can be seen in this figure. Such a thin plate of about 100-nm thickness does not increase its individual thickness, and the coalescence of these parallel plates leads to a lower bainite plate.

Figure 10 shows the schematic representation of the heat treatment and the resulting microstructure formed in steel F, the 0.58C steel. The darkly etched regions in lathlike structures are the cementite containing bainite formed at 360 °C, and the white regions surrounding them are the bainitic ferrite grown at 510 °C in the later stage of heat treatment. Although the etching characteristics between them are different because of the difference in cementite precipitation, the bainitic ferrite laths seem to grow continuously from the bainite laths containing cementite platelets. Such a continuous transition can be more clearly observed by scanning electron microscopy, as in Figure 11.[15] However, bainitic ferrite laths cannot grow continuously from bainite plates formed at temperatures below 350 °C, and the separate nucleation at the platelike bainite/austenite interfaces is observed, as reported previously.[15]

IV. DISCUSSION

A. *Morphology of Widmanstätten Ferrite*

In the present study, it has been shown that the morphology of Widmanstätten ferrite laths is quite similar

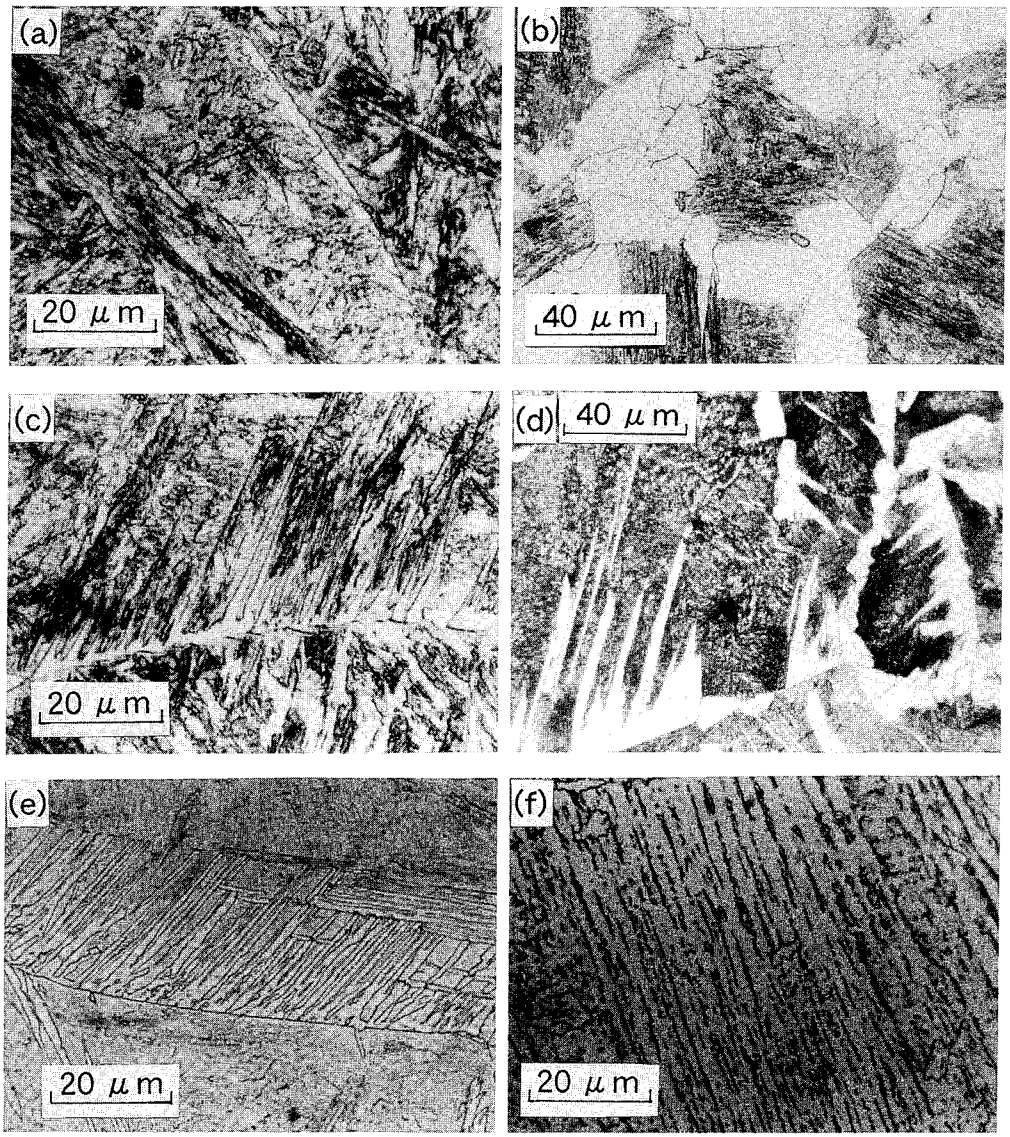


Fig. 4—Microstructures of the isothermally transformed 0.23C-1.86Mn steel (steel C): (*a*) 650 °C, 25 s; (*b*) 650 °C, 600 s; (*c*) 600 °C, 21 s; (*d*) 600 °C, 600 s; (*e*) 500 °C, 7 s; and (*f*) 500 °C, 600 s.

to that of bainitic ferrite, as has been reported to date,[2–6] and that the formation of laths accompanies very sharp surface reliefs, as has been reported by Kinsman *et al.*[7–10] A Widmanstätten ferrite lath revealed by etching, however, is much coarser than an individual surface relief and corresponds to several sets of parallel surface reliefs. The interfaces between Widmanstätten ferrite and untransformed regions are also slightly wavy compared with surface-relief interfaces. It seems likely, therefore, that growth in thickness of an individual sublath of Widmanstätten ferrite does not occur and that the coalescence of sublaths and the diffusional migration of the coalesced-lath interface result in growth in thickness. Such a growth mechanism is in fact supported by the observation that subboundaries often lie parallel to the ferrite growth direction within the ferrite lath, as shown by Figure 7(a).

In order to produce a Widmanstätten ferrite lath, several sublaths in a common growth direction but with the shape strains in the opposite directions will coalesce in a side-by-side fashion, reducing the total elastic strain that should be accommodated both in the austenite and the ferrite.[18] Such a situation has already been observed by Kinsman *et al.* as a tentlike surface relief,[8,9,10] although the interpretation for it is different from that described previously.[18]

B. *Relationship among T_0, B_s, and W_s*

It is also important to know whether Widmanstätten sublaths are formed by a displacive or a diffusional mechanism. The model that the substitutional lattice is transformed by a displacive mechanism with the partition of carbon atoms has been proposed.[16,17] In this case, the localized austenite region where a displacive transformation is to take place should, be depleted of carbon atoms beforehand by diffusion, since the simultaneous occurrence of a displacive transformation of substitutional lattice and the diffusion of carbon seems difficult to imagine. This implies that the displacive

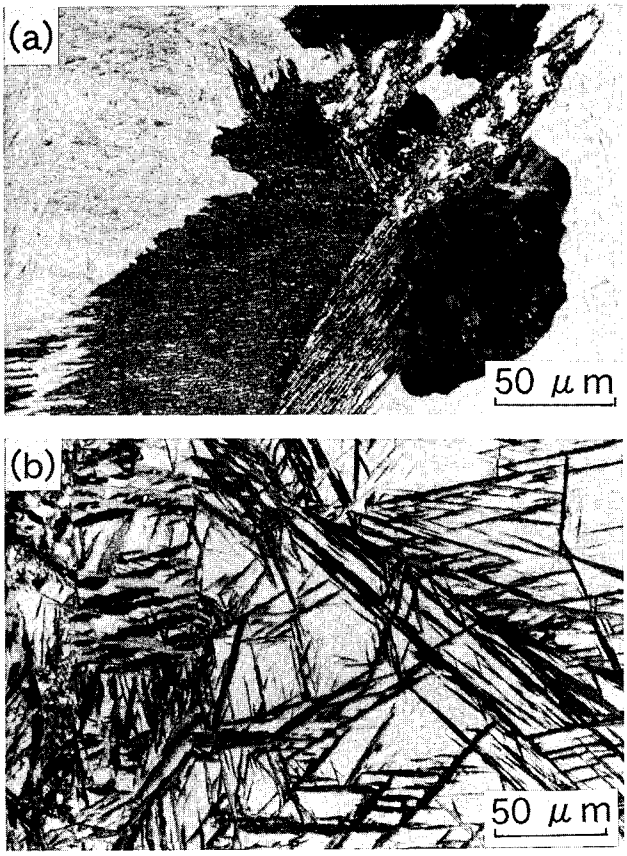


Fig. 5—Microstructures of the isothermally transformed 0.58C steel (steel F): (*a*) 520 °C, 45 s and (*b*) 280 °C, 720 s.

transformation occurs more easily in the localized region where the T_0 temperature is higher than the decomposition temperature. It should also be noted that if the Widmanstätten sublaths form *via* a displacive mechanism, they should accompany the invariant strain-type surface relief. Figure 12 shows the relationship between carbon content and the temperatures where various microstructures form. Calculated T_0 temperatures are also illustrated within the figure. The B_s temperatures decrease almost linearly from 590 °C to 520 °C with increasing carbon content from 0 to 1 wt pct. Such variation of B_s temperature is quite different from that obtained by Aaronson *et al.*[19] This conflict is thought to arise from both the present definition of bainite due to the morphology of ferrite[20] and the classification of bainite and Widmanstätten ferrite, as described in Section I. The B_s temperature for steel G, 1 pct C-1 pct Cr, however, is much higher than the T_0. In this case, however, bainite is formed by alternative stacking of fine ferrite needles and cementite layers. In the hypereutectoid steels, cementite particles will precipitate prior to the ferrite nucleation. The austenite regions adjacent to cementite particles will be depleted in carbon and the local T_0 temperatures in these regions will be higher than the B_s temperature.

W_s temperatures, the upper limits of Widmanstätten ferrite forming temperatures in hypoeutectoid steels, on the other hand, are significantly higher those of B_s but appear to be still lower than T_0 temperatures in the present study. This is more clearly seen in Figure 13 where

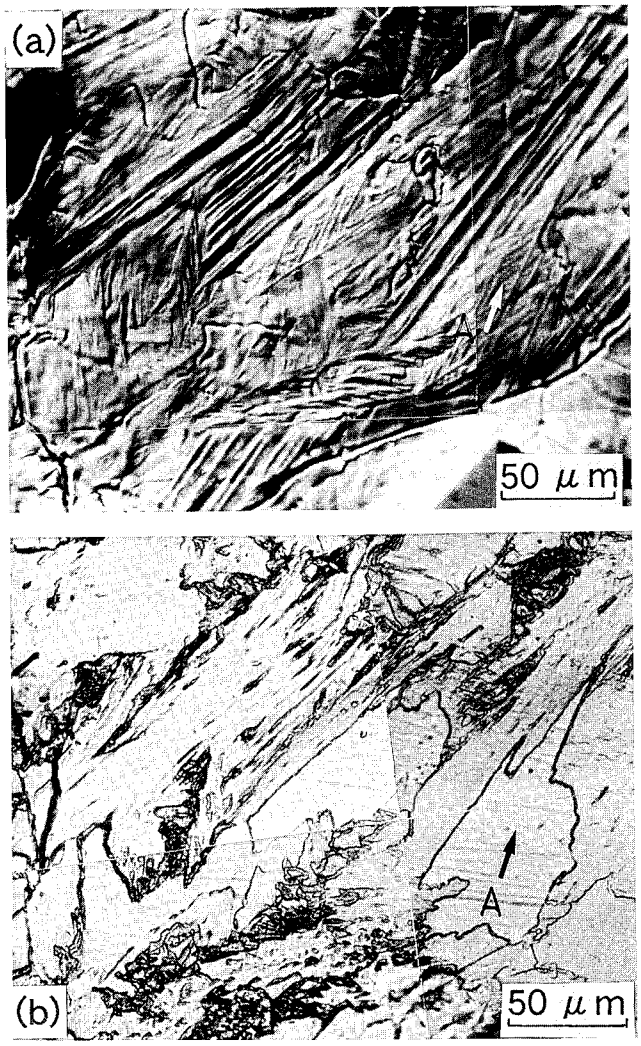


Fig. 6—Surface reliefs for Widmanstätten ferrite and the corresponding microstructures in 0.0041 pct C-2.98 pct Mn-1 pct Cr steel (steel A), decomposed at 600 °C for 600 s; (*a*) surface reliefs and (*b*) microstructures.

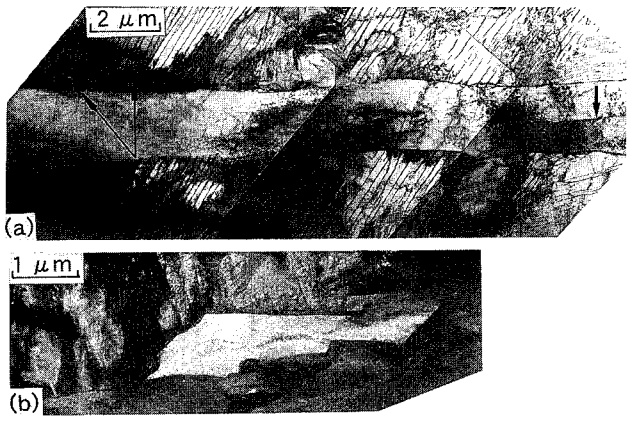


Fig. 7—Widmanstätten ferrite lath observed in steel C transformed at 600 °C for 59 s; (*a*) longitudinal direction and (*b*) cross section.

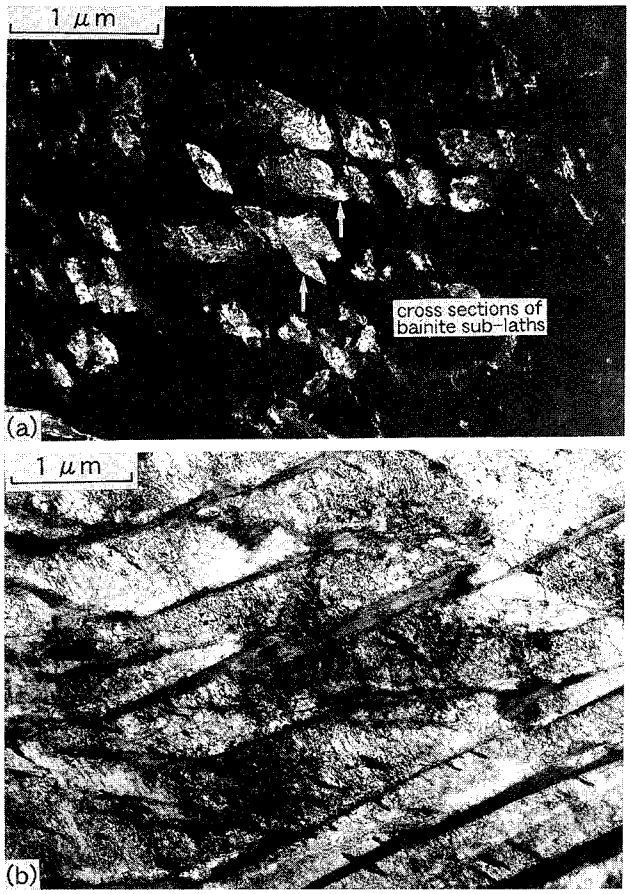


Fig. 8—Upper bainite laths formed in isothermally transformed steel B: (*a*) dark-field image of the ferrite sublath cross sections and (*b*) finally formed bainite.

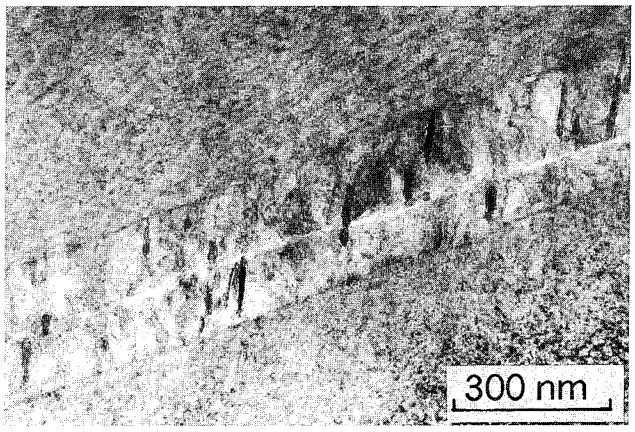


Fig. 9—Early stage of lower bainite formation in steel F transformed at 280 °C for 90 s.

the relationship between $(T_0\text{-}W_s)$ and T_0 is shown. It is interesting to note that $(T_0\text{-}W_s)$ increases with the increase of T_0, *i.e.*, with lowering carbon content. Although W_s seems to be slightly lower than T_0 at the lower limit of T_0, this is within the experimental error. Therefore, the W_s temperatures determined in the present study are thought to be lower than those of T_0.

It has been reported that the W_s temperatures are always in the range between A_{e3} and T_0 except in the cases

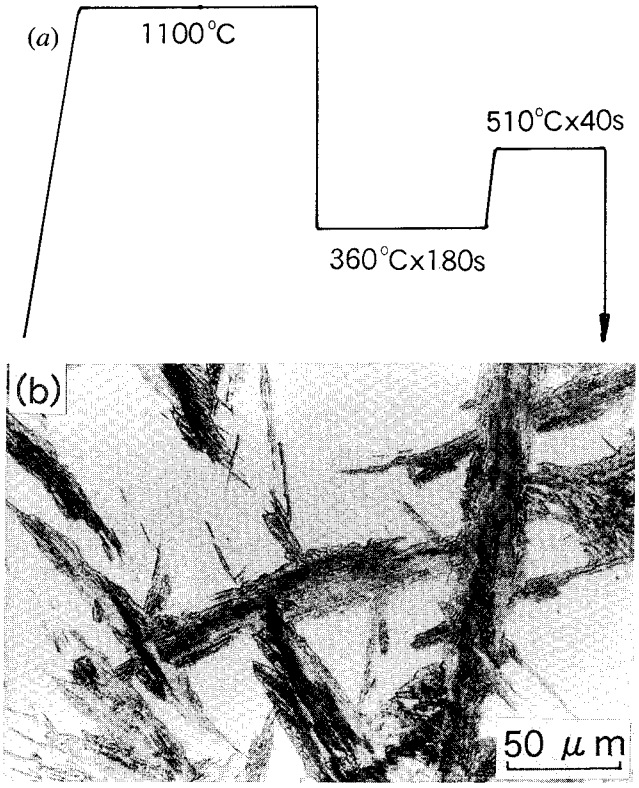


Fig. 10—Growth of typical upper bainite from bainite containing cementite platelets intragranularly in steel F by the abrupt change of decomposition temperature: (*a*) schematic representation of heat treatment and (*b*) optical micrograph.

of Fe-Ni-C and Fe-Mn-C alloys.[8,10,19–21] Since the materials used in the present study contain considerable amounts of Mn, the results described in Section III are mostly in keeping with them. If the mechanism of Widmanstätten ferrite formation is diffusional, there will be no problems concerning T_0 temperature. However, if it forms *via* a displacive mechanism, the W_s temperature is preferred to be lower than T_0 and the following three possibilities should be considered.

(1) Widmanstätten ferrite often grows from the grain-boundary ferrite, allotriomorph during cooling from the isothermal decomposition temperature. But Ni and Mn can increase hardenability of the alloys examined and suppress the Widmanstätten ferrite formation from the grain-boundary proeutectoid ferrite particles during the subsequent cooling.

(2) Another possibility is the existence of two types of Widmanstätten ferrite. One is formed by a diffusional mechanism, such as the migration of ledges.[10] The other is displacive at lower temperatures.

(3) The last one is the mechanism where the substitutional lattice is transformed displacively with simultaneous carbon diffusion.[16,17]

Considering the three possibilities with the experimental results obtained so far, it appears plausible that the Widmanstätten ferrite forms *via* a diffusional mechanism. It seems, however, still difficult to explain the very sharp surface reliefs in terms of a diffusional model. In order to clarify these points, further detailed study will be needed.

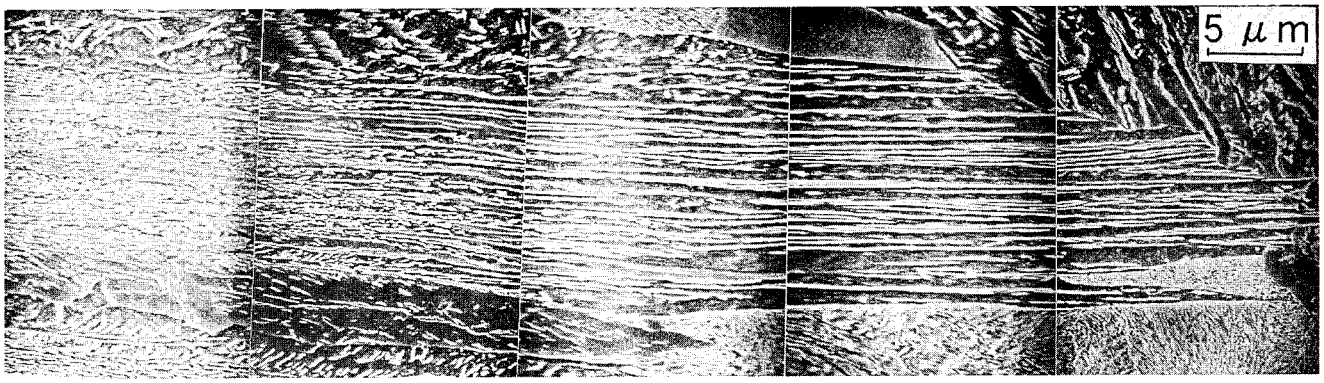


Fig. 11—Scanning electron micrograph of the specimen shown in Figure 10.

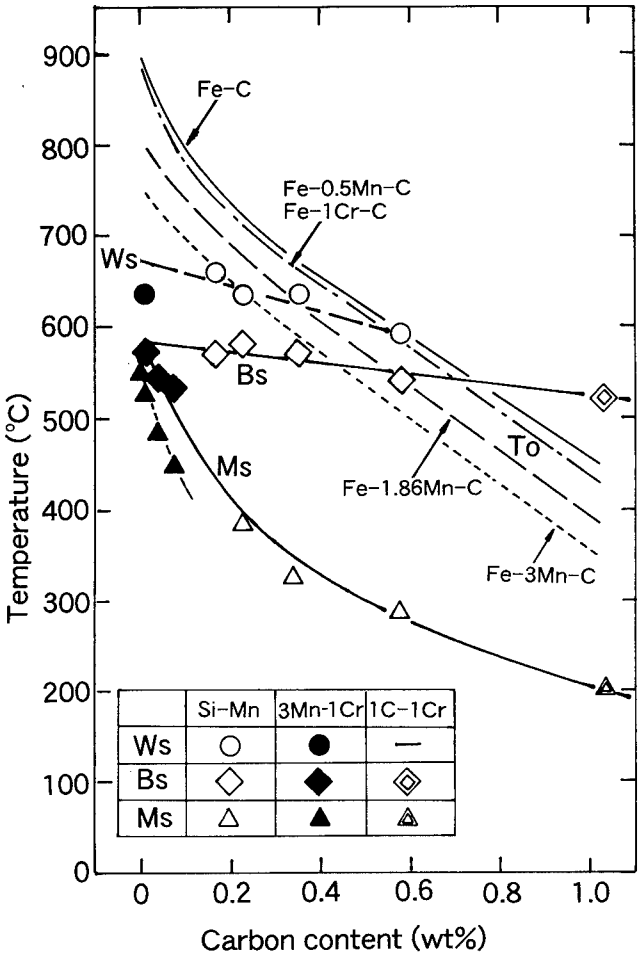


Fig. 12—Relationships among T_0, W_s, B_s, and M_s.

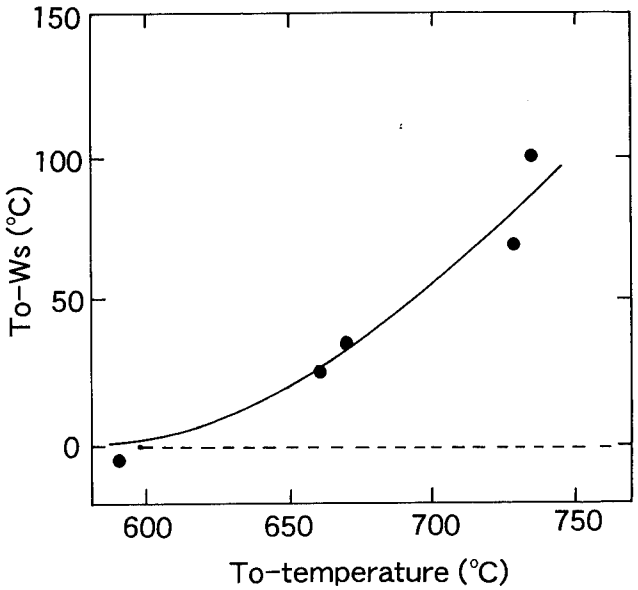


Fig. 13—Relationship between (T_0 − W_s) and T_0 in the materials examined in the present study.

C. *Mechanisms of Bainite and Widmanstätten Ferrite Formation*

The formation of surface reliefs should be expected as the shape change arising from the lattice invariant shear, the Bain correspondence, and the rigid-body rotation according to the crystallography of diffusionless transformation.[22–25] It has, however, been reported that surface relief itself cannot be evidence of a displacive mechanism and that the ledge mechanism can also provide a similar effect.[8,10] Such a situation might be expected if the growth of a Widmanstätten ferrite lath is proceeded by the migration of a single ledge on a growing broad-lath surface. If many ledges nucleate independently on a Widmanstätten ferrite lath surface, the shape change normal to the invariant line on the lath surface will be much reduced by the introduction of numerous misfit dislocations and the surface reliefs will arise mostly from the volume change due to the austenite/ferrite transformation. Thus, it seems more realistic to assume that only a displacive transformation leads to the formation of invariant-plane strain-type surface reliefs, as has been reported.[7,26,27]

Thus, the longitudinal growth of a Widmanstätten ferrite lath is assumed to occur *via* displacive mechanism in the present study. Then, what is the difference between bainitic ferrite and Widmanstätten ferrite? It should be noted that Widmanstätten ferrite laths are

often terminated at a grain-boundary ferrite allotriomorph but that the bainitic ferrite laths formed at temperatures below B_s nucleate directly at an austenite grain boundary. This suggests that a Widmanstätten ferrite nucleates in the orthoequilibrium condition at either perturbed grain-boundary ferrite/austenite interface or a grain boundary *via* a diffusional manner with a specific orientation relationship such as that of Kurdjumov-Sachs, as depicted in Figure 14. Thus, a large supercooling down to B_s will not be necessary, and a Widmanstätten ferrite can grow displacively at lower temperatures within the C-curve for a primary ferrite formation. Similar situations can also be provided at the nonmetallic inclusion/austenite interfaces such as those in the weld heat-affected zone and the weld metal where nonmetallic particles[28,29] such as nitride and oxide particles are artificially distributed to reduce the grain size. On the other hand, it is well established that bainitic ferrite nucleates in the paraequilibrium condition.[30] If the nucleation occurs *via* the diffusional mechanism of this condition, the bainitic reaction will have its own C-curve different from that of Widmanstätten ferrite. Therefore, the difference between a Widmanstätten ferrite and a bainitic ferrite is thought to exist in the nucleation events, but they will grow in the edgewise direction *via* exactly the same displacive mechanism. It should again be mentioned that both Widmanstätten ferrite and bainite laths cannot grow in thickness directions by displacive mechanism. This is because the dislocation arrangements of the lath surfaces are rapidly modified at temperatures where they form. Thus, the thickness can be increased by the coalescence of the sublaths and the diffusional migration of ferrite/austenite interfaces with reducing the sharpness of the lath interfaces. This process might induce the rather complicated situations such that the kinetics of Widmanstätten ferrite growth is controlled by a diffusional mechanism.

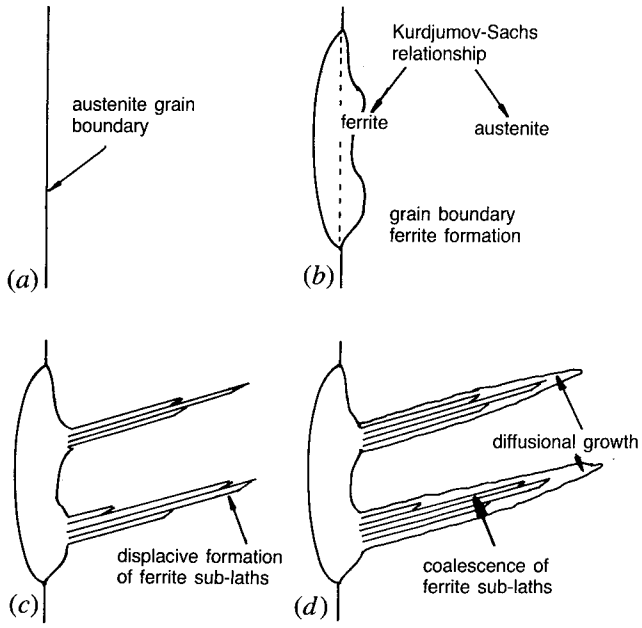


Fig. 14—Schematic illustration of the Widmanstätten ferrite formation: (*a*) austenite grain boundary, (*b*) grain boundary ferrite formation, (*c*) displacive formation of ferrite sublaths, and (*d*) final form of Widmanstätten ferrite.

D. *Classification of Upper and Lower Bainite*

As shown in Figures 10 and 11, typical upper bainite consisting ferrite laths with cementite layers between them can grow quite continuously from the lathlike bainite in which cementite platelets of a specific orientation disperse (this bainite has been referred to as B-III type bainite by the present authors.[20]) According to the definition of bainite due to cementite dispersion, the latter bainite should be classified as lower bainite[31,32,33] or B1 bainite.[34] It has also be confirmed that lathlike bainite cannot grow continuously from a platelike lower bainite.[15] This indicates that bainite should be classified by ferrite morphology but not cementite dispersion: *i.e.*, the shape of ferrite is lathlike or platelike, and that this morphological change occurs at 350 °C in the case of both Fe-C binary alloys and commercial Si-Mn steels. It is interesting to note that martensite also changes its morphology from lathlike to platelike with increasing carbon content,[35,36] and the M_s temperature corresponding to this change is about 350 °C. It seems suggestive that such a correspondence is not accidental and that a transformation mechanism similar to that of martensite operates in the case of bainite.

V. CONCLUSIONS

1. Typical upper bainite laths can grow from the bainite laths containing cementite platelets in a specific orientation (B-III type bainite[20]) quite continuously but do not from the platelike lower bainite.
2. Thus, in view of crystallographic definition of bainite, the classification of bainite should be based on ferrite morphology.
3. A Widmanstätten ferrite sublath is thought to grow from the ferrite nucleus formed in the orthoequilibrium condition.
4. A Widmanstätten ferrite lath is formed by the coalescence of the parallel ferrite sublaths elongated in the same direction. Each sublath formation accompanies a surface relief. Widmanstätten-ferrite-lath/austenite interface can also migrate significantly by a diffusional mechanism, resulting in slightly curved interfaces.
5. Differences between Widmanstätten ferrite and bainitic ferrite may only exist in the nucleation events; *i.e.*, the nucleus for a bainitic ferrite lath forms in the paraequilibrium condition at an austenite grain boundary but the nucleation of a Widmanstätten ferrite lath occurs in the orthoequilibrium condition. It seems likely that the longitudinal growth for both structures occurs *via* the same displacive mechanism.

REFERENCES

1. T. Araki, K. Shibata, and M. Enomoto: *Proc. ICOMAT-89*, Sydney, appeared in *Mater. Sci. Forum*, 1990, vol. 56, pp. 275-80.
2. G.R. Purdy and M. Hillert: *Acta Metall.*, 1984, vol. 32, pp. 823-28.

3. R.F. Mehl, C.S. Barret, and D.W. Smith: *Trans. AIME*, 1933, vol. 105, p. 215.
4. S. Bowles and N.F. Kennon: *J. Aust. Inst. Met.*, 1960, vol. 5, pp. 106-13.
5. M.G. Hall, H.I. Aaronson, and G.W. Lorimer: *Scripta Metall.*, 1975, vol. 9, pp. 533-42.
6. J.D. Watson and P.G. McDougall: *Acta Metall.*, 1973, vol. 21, pp. 961-73.
7. T. Ko and S.A. Cottrell: *J. Iron Steel Inst.*, 1952, vol. 172, pp. 307-13.
8. K.R. Kinsman and H.I. Aaronson: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Greenwich, CT, 1967, pp. 39-53.
9. Y.C. Liu and H.I. Aaronson: *Acta Metall.*, 1970, vol. 18, p. 845.
10. K.R. Kinsman, E. Eichen, and H.I. Aaronson: *Metall. Trans. A*, 1975, vol. 6A, pp. 303-17.
11. M. Hillert and L.I. Staffanson: *Acta Chem. Scand.*, 1970, vol. 24, p. 3618.
12. B. Uhrenius: *Hardenability Concepts with Applications to Steel*, V.D. Doane and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, p. 28.
13. S. Okaguchi, H. Ohtani, and Y. Ohmori: *Mater. Trans., JIM*, 1991, vol. 32, pp. 697-704.
14. H. Ohtani, S. Okaguchi, Y. Fujishiro, and Y. Ohmori: *Metall. Trans. A*, 1990, vol. 21, pp. 877-88.
15. Y. Ohmori, H. Ohtsubo, K. Georgima, and N. Maruyama: *Mater. Trans.*, JIM, 1993, vol. 34, pp. 216-23.
16. L. Kaufman, S.V. Radcliffe, and M. Cohen: *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, pp. 131-352.
17. J.W. Christian: in *Decomposition of Austenite by Diffusional Processes*, Interscience, New York, NY, 1962, p. 480 (referred by H.I. Aaronson).
18. H.K.D.H. Bhadeshia: *Acta Metall.*, 1981, vol. 29, pp. 1117-30.
19. H.I. Aaronson, H.A. Domain, and G.M. Pond: *Trans. TMS-AIME*, 1966, vol. 236, pp. 753-67.
20. Y. Ohmori, H. Ohtani, and T. Kunitake: *Trans. Iron Steel Inst. Jpn.*, 1971, vol. 11, pp. 250-59.
21. H.I. Aaronson, P.G. Boswell, and K.R. Kinsman: *Mechanical Properties and Phase Transformations in Engineering Materials*, S.D. Antolovich, R.O. Ritchie, and W.W. Gerberich, eds., TMS-AIME, Warrendale, PA, 1986, p. 467.
22. M.S. Wechsler, D.S. Lieberman, and T.A. Read: *Trans. AIME*, 1953, vol. 197, pp. 1503-15.
23. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 129-37.
24. J.K. Mackenzie and J.S. Bowles: *Acta Metall.*, 1954, vol. 2, pp. 138-47.
25. J.S. Bowles and J.K. Mackenzie: *Acta Metall.*, 1954, vol. 2, pp. 224-34.
26. J.M. Oblak and R.F. Hehemann: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Greenwich, CT, 1967, pp. 15-30.
27. J.W. Christian: *Metall. Trans. A*, 1990, vol. 21, pp. 799-803.
28. S.S. Babu and H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1990, vol. 6, pp. 1005-20.
29. H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1992, vol. 8, pp. 123-33.
30. H.I. Aaronson, H.A. Domain, and G.M. Pond: *Trans. TMS-AIME*, 1966, vol. 236, pp. 768-80.
31. F.B. Pickering: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Greenwich, CT, 1967, pp. 109-29.
32. R.W.K. Honeycombe and F.B. Pickering: *Metall. Trans.*, 1972, vol. 3, pp. 1099-1112.
33. M. Takahashi and H.K.D.H. Bhadeshia: *Mater. Sci. Technol.*, 1990, vol. 6, pp. 592-603.
34. B.L. Bramfitt and J.G. Speer: *Metall. Trans. A*, 1990, vol. 21, pp. 817-29.
35. A.R. Marder and G. Krauss: *Trans. ASM*, 1967, vol. 60, pp. 651-60.
36. G.R. Speich: *Metall. Trans.*, 1972, vol. 3, p. 1045.

The Influence of the Alloying Elements upon the Transformation Kinetics and Morphologies of Ferrite Plates in Alloy Steels

S.K. LIU, L. YANG, D.G. ZHU, and J. ZHANG

A series of Fe-C-X and Fe-C-X_1-X_2 alloys in which X, X_1, and X_2 either raise or depress the activity of C in γ were investigated by autodilatometer, optical microscopy, and transmission electron microscopy (TEM) to reveal the relations among the chemical composition, transformation kinetics, and morphology of ferrite plates. The incubation time of austenite decomposition at the nose temperature in the time-temperature-transformation (TTT) diagrams, the concentration of C in γ in contact with the α/γ boundary, and the growth rate of ferrite were evaluated to estimate the magnitude of the solute drag-like effect (SDLE) for the different alloying elements used. All the results are consistent qualitatively with the SDLE hypothesis.

I. INTRODUCTION

SINCE there are significant discrepancies between the paraequilibrium model for ferrite growth and the experiments in Fe-C-X systems, where X is Mo, Cr, or Mn, these alloying elements must have an additional influence upon ferrite growth. This influence has been suggested to be a solute drag-like effect (SDLE).[1,2] The SDLE concept has undergone a number of changes since it was introduced in 1967. In its present form, the SDLE is proposed to result from a nonequilibrium segregation of the substitutional solute element and C to moving α/γ boundaries. Segregated solutes which lower the activity of C in austenite in contact with the α/γ boundary, in turn, reduce the growth kinetics of ferrite. The SDLE is also used to explain the incomplete transformation phenomenon below the B_s temperature and the appearance of a bay in the time-temperature-transformation (TTT) diagrams of these alloys.

The essential requirement for the operation of a SDLE is the segregation of substitutional alloying elements, X, at the α/γ boundary. Numerous experiments to check X segregation at α/γ boundaries have been made by field-ion microscopy (FIM)/atom probe (AP) and scanning transmission electron microscopy (STEM). Reynolds *et al.*[3] have reported an FIM/AP analysis of interfaces in an Fe-0.19C-1.81Mo alloy reacted at 585 °C for 30 seconds. The results indicate that Mo enriches to α/γ boundaries by at least a factor of 2. Stark and Smith[4] reported STEM X-ray analysis of interfaces in Fe-0.32C-2.05Mo alloy reacted at 595 °C for 33 hours. The result also indicates that approximately twofold segregation appears to have occurred at α/γ interfaces in this alloy. But contradictory results have been reported for Mn-containing alloys.[5] Since X atoms which have a strong affinity for C are likely to segregate to the disordered portions of α/γ boundaries, *i.e.*, to the kinks on the mobile ledge risers, only the concentration information which comes from the risers can demonstrate whether the X does segregate at α/γ interfaces. Unfortunately, the current level of the preparation of samples and the experimental methods cannot ensure this requirement. Discrepancies among experiments are inevitable.

Primary calculations of $X_C^{\gamma/\alpha}$, the concentration of C in austenite in contact with the α/γ boundary, as a function of the concentrations of C and Mo in the boundaries X_C^{ϕ} and X_{Mo}^{ϕ}, respectively, have been given by Reynolds *et al.*[2] The results show that the smaller the value of $X_C^{\gamma/\alpha}$ for a given alloy, the smaller the C flux and the slower the ferrite growth. Because of the lack of accurate information on solute adsorption at α/γ boundaries in Fe-C-X alloys and solution parameters for the α/γ boundary responsible for growth, these calculations are only a first step toward quantifying the SDLE. Therefore, more calculations are needed to demonstrate the SDLE proposal quantitatively.

In view of the aforementioned facts and difficulties, the authors expect to obtain more data to test the qualitative predictions of the SDLE proposal through the investigation of the relations among the chemical compositions, transformation kinetics, and morphologies of ferrite in a series of Fe-C-X and Fe-C-X_1-X_2 alloys that contain solutes that raise the C activity in γ and solutes that depress the C activity in γ. The authors also intend to further explore the relationship between the segregation of solutes in the α/γ interface, X_X^{ϕ}, the concentration of C in γ in contact with the α/γ boundary, $X_C^{\gamma/\alpha}$, and the relationship between X_X^{ϕ} and the growth rate of ferrite. For this purpose, four groups of alloys were chosen for study.

(1) Ternary Fe-C-X alloys in which X represents Si and Ni. These raise the activity of C in γ, but Si is a ferrite stabilizer and Ni stabilizes austenite.
(2) Ternary Fe-C-X alloys in which X represents V and Mn. These depress the activity of C in γ; V stabilizes ferrite and Mn stabilizes austenite.
(3) Quaternary Fe-C-Si-Mn and Fe-C-Si-V alloys which

S.K. LIU and L. YANG, Professors, and D.G. ZHU and J. ZHANG, Lecturers, are with the Department of Materials Engineering, Southwest Jiaotong University, Chengdu Sichuan, 610031, People's Republic of China.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

consist of one solute that raises and one that depresses the activity of C in γ. These alloys can be used to investigate potential synergistic interactions of alloying elements on the growth of ferrite.

(4) The Fe-Mn alloy which contains a low concentration of C (0.01 wt pct).

Besides the fourth alloy, all the alloys employed contain approximately 0.4 wt pct C to facilitate comparisons among the alloys.

All the samples were reacted isothermally at temperatures ranging from M_s to slightly higher than B_s. Emphasis is placed upon temperatures near the nose of the TTT diagram.

II. EXPERIMENTAL PROCEDURE

Alloys used in this investigation were melted in a vacuum induction furnace from commercial-purity Fe and high-purity Si, Ni, Mn (99.9 pct), and ferrovanadium. The chemical compositions are given in Table I. Ingots were forged and then homogenized at 1200 °C for 72 hours. The TTT curves for these alloys were determined using an autodilatometer combined with optical microscopy. Samples of 3 mm in diameter and 10 mm-long were austenitized in vacuum at 1200 °C for 10 minutes and quenched in a stream of N or He to the isothermal reaction temperatures. The initiation of transformation was determined by the inflection in the dilatation *vs* time curve and checked with optical microscopy. The fraction transformed and the fraction of ferrite located at prior austenite grain boundaries and within grains were determined by point-counting optical micrographs. The initial and terminal lines in TTT curves correspond to 1 and 95 pct of austenite transformed, respectively.

Specimens from representative alloys reacted for selected transformation temperatures and times were examined with transmission electron microscopy (TEM) to determine the morphology characteristics, the fine structures of ferrite, and the presence of carbides. Foil specimens were prepared by double-jet polishing in a 10 pct perchloric acid, 90 pct ethanol solution at −30 °C and 50 V. The TEM analyses were carried out in a Hitachi H-700H microscope at an operating voltage of 200 kV.

The Auger experiment was carried out on a PERKIN-ELMER* PHI 600 instrument to determine the enrichment of C and Si to γ/γ boundaries in the Fe-C-Si-Mn alloy. Specimens were broken within the chamber in a vacuum of 10^{-7} Pa. Then, Ar^+ ions were used to sputter the fractured surface of the specimens at an operating voltage of 3 kV and a 0.07 μA-beam current.

*PERKIN-ELMER is a trademark of Perkin-Elmer Physical Electronics, Eden Prairie, MN.

Table I. Chemical Composition of Experimental Alloys (Weight Percent)

Alloy	C	Ni	Si	Mn	V
Fe-C-Ni	0.46	2.32	—	—	—
Fe-C-Si	0.38	—	1.73	—	—
Fe-C-Mn	0.38	—	—	3.11	—
Fe-C-V	0.42	—	—	—	1.0
Fe-C-Si-Mn	0.37	—	2.23	2.29	—
Fe-C-Si-V	0.42	—	1.96	—	0.96
Fe-Mn	0.01	—	—	3.24	—

The shear elasticity, E_s, for each alloy was measured by a so-called "microshear testing technique" to estimate the effect of volume-strain energy developing within the matrix during decomposition of austenite upon the incubation time of the transformation in the alloys.[6] The average value of six tests for each alloy is accurate enough for the purpose of this investigation.

III. RESULTS

A. *Fe-C-Ni and Fe-C-Si*

The TTT curves of Fe-C-Ni and Fe-C-Si alloys are shown in Figure 1. The overall kinetics of Fe-C-Si are fast, and the decomposition of austenite proceeds without stasis. There is no bay in the TTT curves. Although a shallow bay is present in the TTT curve of the Fe-C-Ni alloy, the transformation proceeds rapidly and without stasis. Thus, transformation does not exhibit the incomplete phenomenon at the temperatures investigated in the Fe-C-Ni and Fe-C-Si alloys. Although both Si and Ni are elements which raise the activity of C in γ, Si is a ferrite stabilizer and Ni is an austenite stabilizer; the former raises the A_{e_3} temperature, and the latter depresses it. Thus, compared with the Fe-0.4C alloy, the Fe-C-Si alloy has slightly faster and the Fe-C-Ni alloy has slower transformation kinetics.

The morphology of the ferrite component of bainite in the Fe-C-Ni alloy reacted at a temperature range from 700 °C down to 400 °C, and the Fe-C-Si alloy reacted at a temperature range from 560 °C to 400 °C, similar to that in Fe-0.4C alloys. Once the allotriomorph ferrites are replaced by Widmanstätten ferrite plates, the Widmanstätten structure remains as the dominant morphology down to the M_s temperature. Figures 2(a) and (b) show the optical micrographs of the Fe-C-Ni and Fe-C-Si reacted at 550 °C for 10 seconds and 450 °C for 3 seconds, respectively. It can be seen from these micrographs that most of the ferrite plates initiate from the prior γ grain boundaries and appear as parallel, slightly tapering plates. Intragranular plates are rare in both specimens, treated at 450 °C and above 450 °C for a short

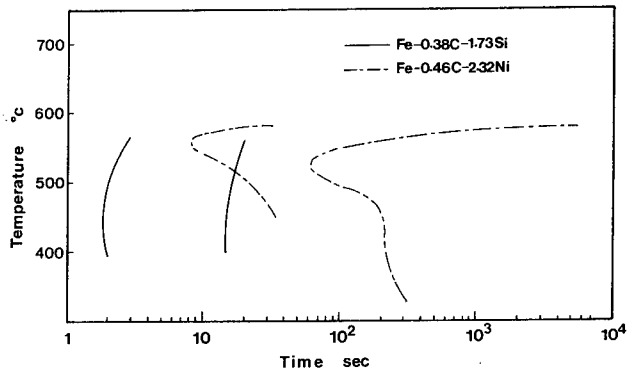


Fig. 1—TTT diagrams of Fe-0.46C-2.32Ni and Fe-0.38C-1.73Si alloys.

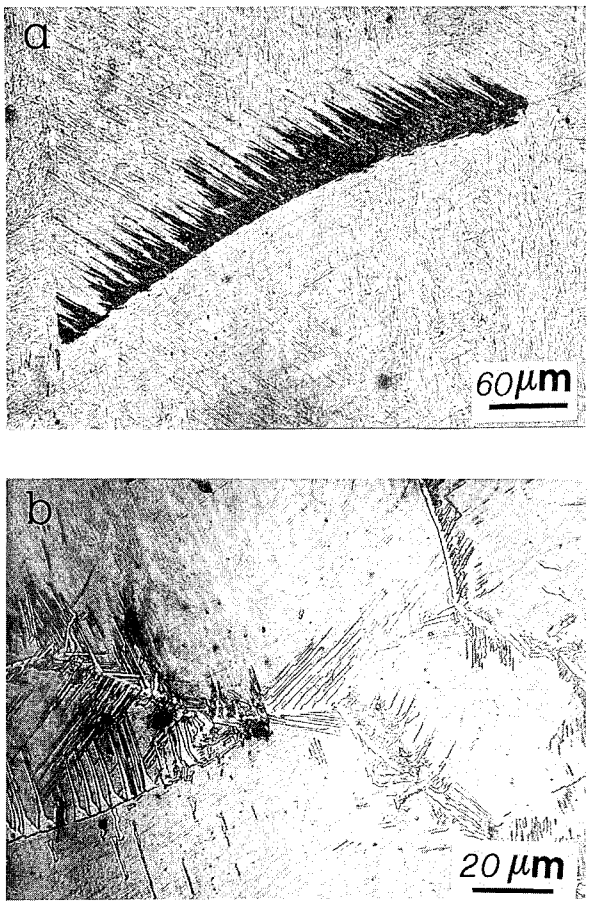


Fig. 2—Optical microstructures of (*a*) Fe-0.46C-2.32Ni alloy treated isothermally at 550 °C for 10 s and (*b*) Fe-0.38C-1.73Si alloy treated isothermally at 450 °C for 3 s.

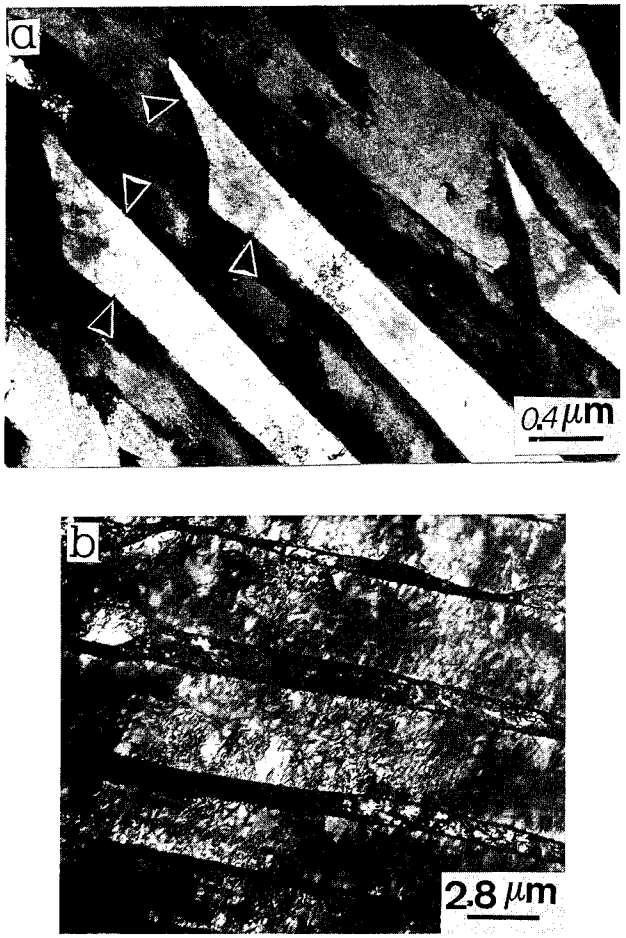


Fig. 3—TEM microstructure of Fe-0.38C-1.73Si treated at 450 °C for (*a*) 3 s and (*b*) 10 s.

time. The dark etching characteristic of the ferrite sideplates in the Fe-C-Ni alloy (Figure 2(a)) suggests a high density of unresolvable carbides. No evidence was found of carbide formation in association with ferrite sideplates in Fe-C-Si owing to the inhibiting effect of Si on the precipitation of Fe_3C. The TEM micrographs of carbide-free ferrite sideplates in the Fe-C-Si alloy isothermally treated at 450 °C for 2 and 10 seconds are shown in Figure 3. The steps indicated by arrows are likely to be growth ledges. In this case, the growth of ferrite plates takes place mainly by ledge growth. However, edge-to-face sympathetic nucleation can be observed on the branching of plates at later stages of transformation.

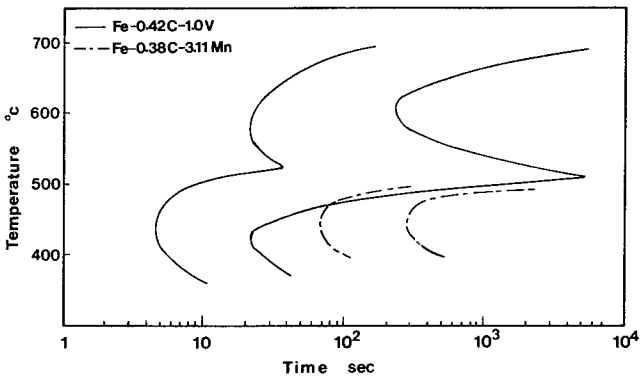


Fig. 4—TTT diagrams of Fe-0.42C-1.0V and Fe-0.38C-3.11Mn alloys.

B. *Fe-C-V and Fe-C-Mn*

The TTT curves of the Fe-C-V and Fe-C-Mn alloys are shown in Figure 4. Compared with the TTT curves of Fe-C-Si and Fe-C-Ni alloys, it can be seen that the incubation time for the decomposition of austenite in the Fe-C-V alloy treated at intermediate temperatures is longer, and a deep bay is associated with its TTT diagram. A deep bay in a TTT diagram does not guarantee the existence of incomplete transformation in an alloy treated at kinetics—B_s. It is evident that almost a horizontal stage on the curve of fraction transformed *vs* time means cessation of the transformation. But, this is not the case with this Fe-C-V alloy.

The initiation of transformation in Fe-C-Mn also needs a longer time than in Fe-C-Si and Fe-C-Ni alloys. At 495 °C, bainite and pearlite precipitate from austenite and the volume fractions of both the bainite and pearlite increase with time simultaneously. No indication of transformation stasis was detected at temperatures from 400 °C to 495 °C in the Fe-C-Mn alloy.

The typical morphology of the ferrite component of bainite in the F-C-V alloy isothermally reacted at 550 °C for 70 seconds is shown in Figure 5. Five conspicuous differences between the morphologies of ferrite in the Fe-C-V and some Fe-C alloys are noteworthy. First, the morphology of grain-boundary ferrite allotriomorphs in the Fe-C-V has a very small aspect ratio and extends farther along the grain boundaries as a net. This has been termed the "snake" morphology[7] to distinguish it from the conventional allotriomorph ferrite which has an aspect ratio of approximately one-third in many Fe-C alloys. Second, above 400 °C, ferrite plates originate from prior γ grain boundaries and do not often develop as parallel plates, unlike ferrite in Fe-C, Fe-C-Si, and Fe-C-Ni alloys (compare Figure 5 with Figure 2). Third, unlike Fe-C, Fe-C-Si, and Fe-C-Ni alloys in which ferrite plates originate from γ grain boundaries at higher temperatures and develop at grain boundaries and within the grain simultaneously at lower temperatures, once the plates appear in the microstructure of the Fe-C-V specimen, they develop at the grain boundaries and within the grain simultaneously. Fourth, there is no carbide precipitation accompanying ferrite growth at the early stage of transformation at 510 °C. It takes approximately 200 to 800 seconds for carbide to precipitate adjacent to prior γ grain boundaries and within the γ grains. Fifth, within the temperature range from approximately 50 °C above to 100 °C below the B_s, the predominant morphology of ferrite is a degenerate form of the Widmanstätten structure.

The morphology characteristics of ferrite in Fe-C-Mn alloy treated at different temperatures are similar to those of ferrite in the Fe-C-V alloy, except there is lack of snake allotriomorphs in the microstructure of the Fe-C-Mn alloy (Figure 6). From Figures 5 and 6 and the results of TEM examination, it can be observed that carbides precipitate accompanying ferrite plates in the Fe-C-Mn alloy but not in the Fe-C-V alloy at an early stage of the transformation.

C. *Fe-C-Si-Mn and Fe-C-Si-V*

Figure 7 shows TTT diagrams of the Fe-C-Si-Mn and Fe-C-Si-V alloys. These two TTT diagrams shift to slightly longer times compared with TTT diagrams of the Fe-C-Si, Fe-C-Mn, and Fe-C-V alloys. Introducing Si (which raises the activity of C in γ) into a ternary alloy containing Mn or V (which depress the activity of C in γ) does not weaken the drag-like effect of Mn or V on the growth of ferrite; rather it promotes their inhibiting effect on ferrite growth. At a temperature close to the B_s, the initiation and especially the completion of austenite decomposition is exceedingly slow. There is a zero slope on the curve of fraction transformed *vs* time for the Fe-C-Si-Mn (Figure 8) and the Fe-C-Si-V alloys. Thus, the incomplete-transformation phenomenon is observed in these alloys.

The microstructure of Fe-C-Si-Mn alloy reacted isothermally at 450 °C for 2300 seconds is shown in Figure 9. It is evident that slightly tapering degenerate ferrite plate appears to initiate within prior γ grains and ferrite sideplates are rare. As the isothermal temperature decreases, the slightly degenerate ferrite plates are replaced gradually by sheaves of ferrite. Each sheaf consists of almost parallel ferrite "subunits," presumably

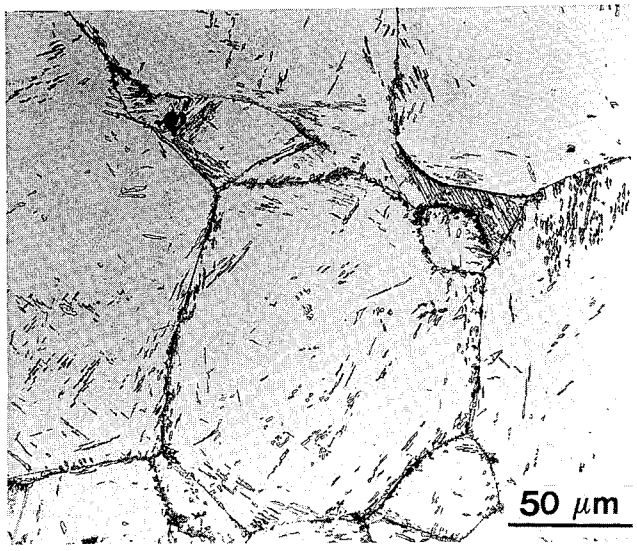


Fig. 6—Optical microstructure of Fe-0.38C-3.11Mn alloy reacted at 495 °C for 420 s.

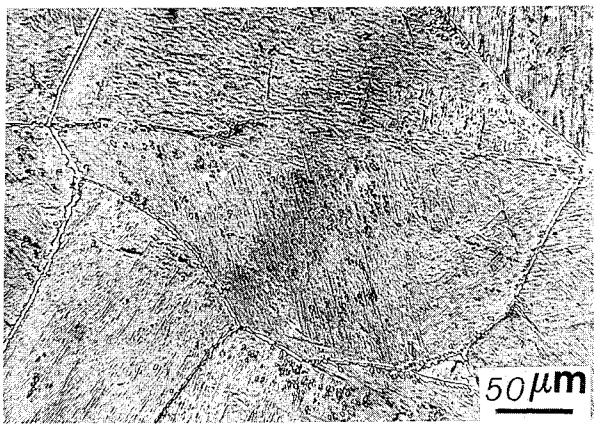


Fig. 5—Optical microstructure of Fe-0.42C-1.0V alloy reacted at 550 °C for 70 s.

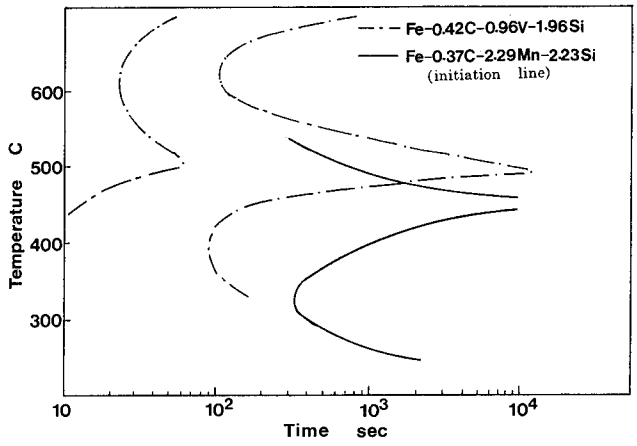


Fig. 7—TTT diagrams of Fe-0.37C-2.23Si-2.29Mn and Fe-0.42C-1.96Si-0.96V alloys.

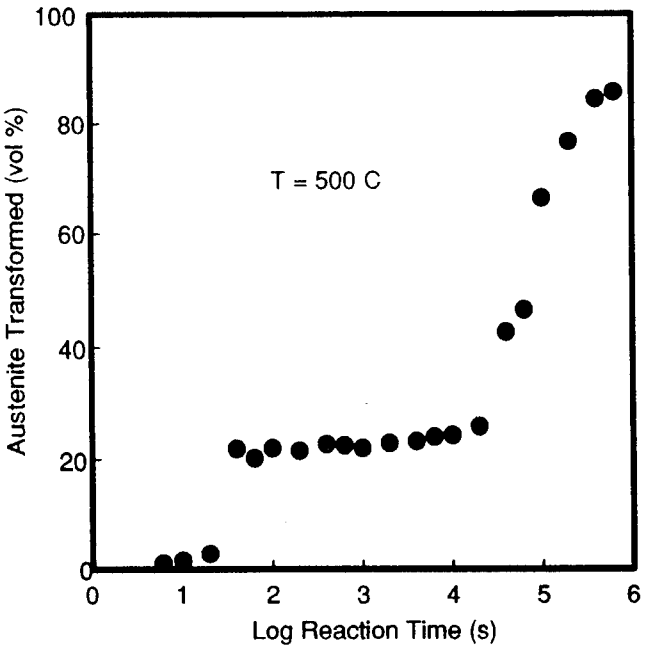


Fig. 8—Isothermal transformation curve for Fe-0.37C-2.23Si-2.29Mn transformed at 500 °C.

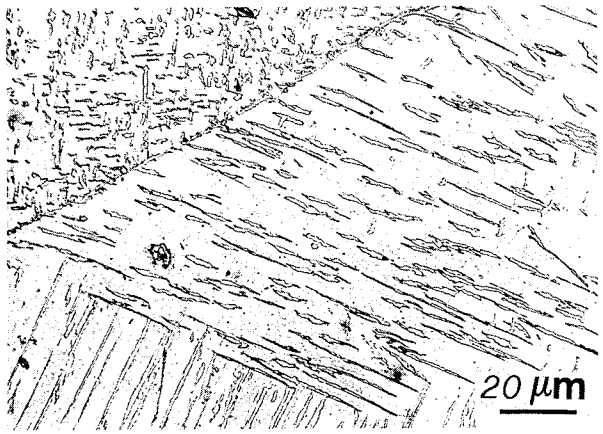


Fig. 9—Optical microstructure of Fe-0.37C-2.23Si-2.29Mn reacted at 450 °C for 2300 s.

formed by face-to-face and edge-to-edge sympathetic nucleation (see the TEM micrograph shown in Figure 10). There was no carbide precipitation accompanying the formation of sheaves after holding at 450 °C for less than 3 to 4 hours. Retained-austenite laths exist between the subunits. But after holding at 450 °C for more than 4 to 5 hours, Fe_3C precipitates can be detected.

Figures 11(a) through (c) show the microstructures of the Fe-C-Si-V alloy reacted at 500 °C for 100, 2000, and 7000 seconds, respectively. The morphology characteristics of ferrite in this alloy are the same as those in the Fe-C-V alloy. Because of the inhibiting effect of Si on carbide precipitation, it takes almost 10 times longer in comparison with the Fe-C-V alloy for the carbide to precipitate along the prior γ grain boundaries (Figure 11(b)).

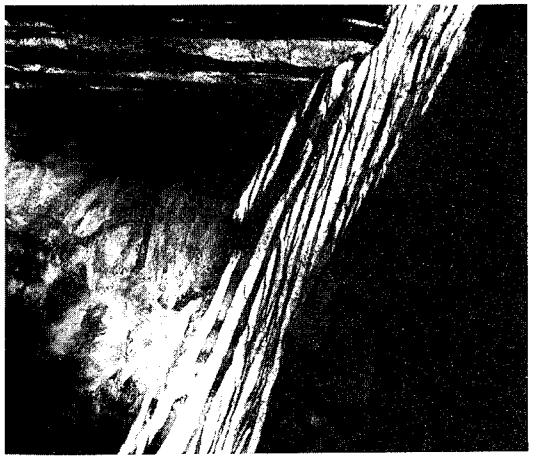


Fig. 10—TEM microstructure of Fe-0.37C-2.23Si-2.29Mn reacted at 375 °C for 100 s.

D. *Fe-Mn*

Figure 12 shows the TTT diagram of Fe-3.24Mn alloy. At temperatures near the nose of the TTT diagram, transformation is too rapid for the early stage to be observed with the method of quenching the specimens into ice brine.

E. *Solute Segregation at Prior γ Grain Boundaries*

From the derivative Auger spectra from the cleavage and the prior γ grain boundary in the Fe-C-Si-Mn alloy isothermally reacted at 475 °C for 20 seconds it can be seen that C is segregated to the prior γ grain boundary. In addition, depletion of Si was found at the prior γ grain boundary in the same sample. The concentration profile of Si and C examined by sputtering the intergranular facets with Ar^+ ions is shown in Figure 13.

IV. INTER-RELATED EVALUATIONS

A. *Evaluation of Incubation Time*

Since more contemporary studies of austenite decomposition have favored models involving "instantaneous" nucleation-site saturation followed by growth-determined initiation of the TTT curve,[8,9] it is possible to ascertain qualitatively the extent of the SDLE on the incubation time of austenite decomposition in the alloy systems investigated by comparing the calculated relative incubation time, τ_{cal}, with the experimental incubation time, τ_{rel}.

According to Kirkaldy,[9] if the isothermal growth velocity of a product is v, then the time τ_{TTT} to any desired volume percent of transformation is given by

$$\tau_{TTT} = \text{const}/v\,(T)$$

and

$$V \propto (\Delta T)^2 \exp\,(-Q/RT)$$

where ΔT is the undercooling, and Q is the activation energy for diffusion.

For small undercooling, ΔT, it can be shown that

$$\Delta G_v \propto \Delta T$$

Thus, we have

$$\tau \propto 1/\{\Delta G_v^2 \exp(-Q/RT)\}$$

A similar dependence is reached if one assumes the incubation time on the TTT diagram is determined by the nucleation rate rather than the growth rate of the product.[10]

The ΔG_v values for all the alloys were calculated by a regular solution model and supercomponent model.[11] Taking the $1/\Delta G_v^2$ value of the Fe-C-Si alloy at the nose temperature as a unit, the ratios of $1/\Delta G_v^2$ values of the other Fe-C-X alloys to the $1/\Delta G_v^2$ value of the Fe-C-Si alloy at the nose temperature are the calculated relative incubation times, τ_{cal}, of the alloys and are shown in Table II. The experimental incubation time, τ_{exp}, corresponding to the nose temperature in the TTT diagrams, which were measured from the TTT diagrams shown in Figures 1, 4, 7, and 12, is also shown in Table II. The ratios of τ_{exp} of the other Fe-C-X alloys to the τ_{exp} of the Fe-C-Si are taken as the real relative incubation times, τ_{rel}. According to the preceding equations, the τ_{cal} value should contain an exponential factor $\exp(-Q/RT_{nose,Si} + Q/RT_{nose})$. Therefore, the τ_{cal} values calculated simply by the ratio of ΔG_v neglect the effect of alloying elements on the diffusivity of C and, thus, the τ values. Since the same exponential factor should appear in the values of τ_{rel}, this factor can be canceled if we measure the relative potence of SDLE by computing τ_{rel}/τ_{cal} for each alloy at the nose temperature of the TTT diagram. Then, the relative incubation time (an approximate measure of the ferrite growth rate) can be predicted using only the thermodynamic driving force. When it is significantly larger than one, growth is much slower than expected and an SDLE must be operating. The τ_{rel}/τ_{cal} ratios are also shown in Table II. Since the τ_{rel}/τ_{cal} ratios for Fe-C-Mn, Fe-C-Mn-Si, Fe-C-V, and Fe-C-V-Si alloys are significantly larger than one, there must be SDLE operating in these alloys during the transformation occurring at the temperature near the nose of the TTT diagram.

B. *Evaluation of $X_C^{\gamma/\alpha}$ and L^2/t*

To estimate the magnitude of the SDLE exerted by the alloying elements used in this investigation on the growth kinetics of ferrite, the authors calculate the concentration of C in austenite in contact with the α/γ boundary, $X_C^{\gamma/\alpha}$, and the values L^2/t (where L is the location of the α/γ interface relative to the γ grain boundary at which it nucleated and t is reaction time) as functions of the Mn, V, Ni, and Si concentrations within the α/γ interface, X_X^ϕ, and the reaction temperatures, T, by a paraequilibrium model and growth of a planar interface model, respectively.

In paraequilibrium, the distribution of C and X can be shown schematically in Figure 14. According to the proposal of SDLE, some of these solutes may segregate to the α/γ interface. If we take the interface as a two-dimensional phase (ϕ), the partition of X among α, γ, and ϕ at a given temperature might be shown schematically (Figure 15). Suppose there is a series of Fe-C-Mn alloys with different concentrations of Mn and X_{Mn}^0. For a given concentration of Mn, the corresponding $X_C^{\gamma/\alpha}$ and $X_C^{\alpha/\gamma}$ can be calculated by equating the chemical potential of C in γ and α. When segregation of X to the α/γ boundary occurs ($X_{Mn}^\phi \neq X_{Mn}^0$), the values of $X_C^{\gamma/\alpha}$ and $X_C^{\alpha/\gamma}$ are taken to be set by the chemical potential of C as determined by X_{Mn}^ϕ. In the absence of detailed information on the thermodynamics of the α/γ boundary, it is assumed that $X_C^{\gamma/\alpha}$ and $X_C^{\alpha/\gamma}$ can be calculated using X_{Mn}^ϕ instead of X_{Mn}^0. Therefore, a functional relationship between the concentration of Mn segregation to the α/γ boundary and the concentration of C in γ in contact with the α/γ boundary at a given temperature can be obtained by varying the parameter X_{Mn}^ϕ. Figure 16 shows the calculated results of the $X_C^{\gamma/\alpha} \sim X_X^\phi$ relation for the investigated alloys by using the computer program provided by Reynolds.[12]

For simplicity, migration of a ferrite riser is regarded as growth behind a planar incoherent interface. Following Zener,[13] we have

$$\frac{(U_C^{\gamma/\alpha} - U_C^{\gamma/0})}{(X_C^{\gamma/\alpha} - X_C^{\alpha/\gamma})}\left(\frac{\delta U_C^\gamma}{\delta X_C^\gamma}\right)^{-1}\left(\frac{V_M^\alpha}{V_M^\gamma}\right)\sqrt{\pi} = \sqrt{\frac{k}{4}}\exp\left(\frac{k}{4}\right)\operatorname{erfc}\sqrt{\frac{k}{4}} \quad [1]$$

$$L^2/t = KD_c \quad [2]$$

where t = time; D_c = diffusion coefficient of C in austenite; $u_C^{\gamma/\alpha}$ = chemical potential of C in γ in contact with the α/γ interface; $u_C^{\gamma/0}$ = chemical potential of C in γ as $t = 0$; V_M^α and V_M^γ = mole vol of α and γ, respectively; and k = constant.

Given the parameters,[12,14] we can obtain the numerical solution of L^2/t with respect to a series of X_X^ϕ. The calculated results are also shown in Figure 16.

V. DISCUSSION

A. *Kinetics*

Explaining the significant discrepancy between the calculated (τ_{cal}) and measured (τ_{rel}) incubation times for the alloys that contain solutes depressing a_C^γ may be a key point to understanding the characteristics of ferrite growth kinetics at intermediate temperatures.

As shown in Table II, the differences in ΔG_v values among the alloys employed in this investigation are small. The large differences in the ratios of τ_{rel}/τ_{cal} between the alloys that contain solutes depressing a_C^γ and the alloys that contain solutes raising a_C^γ are not likely to be caused by such small changes in ΔG_v. The differences among the τ_{cal} for the different alloys (τ_{cal} varies from 0.49 to 4.3) only reflect differences in ΔG_v due to the type and concentration of alloying elements.

The strain energy accompanying the formation of ferrite might be another important factor influencing the transformation kinetics of austenite. It can be seen from

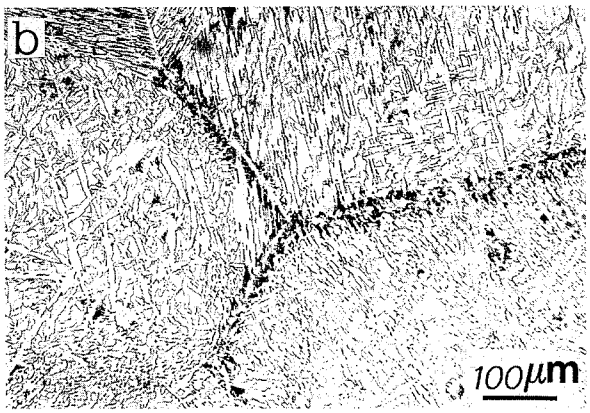
Table II that the ratios of shear elasticity, E_s, for the Fe-C-X and Fe-C-X_1-X_2 alloys to that of the Fe-0.37C-2.20Si-2.29Mn alloy are close to unity, *i.e.*, from 0.77 to 1.12. Therefore, the differences of $(\Delta G_v + \Delta G_s)$ among the alloys which contain elements raising the activity of C in γ and the alloys which contain elements depressing the activity of C in γ cannot be the reasons which make the latter have a remarkably slow decomposition rate of austenite.

In fact, the calculated results of $X_C^{\gamma/\alpha}$ shown in Figure 16 are exactly the paraequilibrium $(\alpha + \gamma)/\gamma$ phase boundary lines for an alloy with variable amounts of X. For alloys which contain solutes stabilizing austenite (Mn and Ni), $X_C^{\gamma/\alpha}$ decreases with an increase in solute X_X^{ϕ}; on the other hand, for alloys which contain solutes stabilizing ferrite (Si and V), $X_C^{\gamma/\alpha}$ increases with an increase in solute concentration (Figure 17). The trends of $X_C^{\gamma/\alpha} \sim X_X^{\phi}$ curves shown in Figure 16 are consistent with these patterns. Therefore, it is evident that an increase of a solute which depresses the activity of C in γ can make the $X_C^{\gamma/\alpha}$ increase or decrease according to which phase (α or γ) it can stabilize. In other words, solutes which reduce the activity of C in γ do not necessarily reduce the $X_C^{\gamma/\alpha}$ value of the alloy.

From Eq. [1], it is clear that the following three factors effect the calculated result of the L^2/t value: $(u_C^{\gamma/\alpha} - u_C^{\gamma/0})$, $(X_C^{\gamma/\alpha} - X_C^{\alpha/\gamma})$, and $(\delta U_C^{\gamma}/\delta X_C^{\gamma})$. Since a solute affects not only the $X_C^{\gamma/\alpha}$ and $X_C^{\alpha/\gamma}$ values but also the $u_C^{\gamma/\alpha}$ and $(\delta U_C^{\gamma}/\delta X_C^{\gamma})$ values, the trends of $L^2/t \sim X_X^{\phi}$ curves shown in Figure 16 are not consistent with those of $X_C^{\gamma/\alpha} \sim X_X^{\phi}$ curves. For alloys containing solute that depresses a_C^{γ} (Mn or V), the L^2/t value decreases with the increase of X_X^{ϕ}. For example, the $X_C^{\gamma/\alpha}$ and $u_C^{\gamma/\alpha}$ of Fe-C-Mn alloy decrease with the increase of X_X^{ϕ}. As a result, L^2/t decreases significantly with the increase of X_X^{ϕ} and reaches zero as X_{Mn}^{ϕ} reaches approximately 11 at. pct. As to the Fe-C-V alloy employed, although the $X_C^{\gamma/\alpha}$ increases with the increase of X_v^{ϕ}, the L^2/t value still decreases and approaches zero as the X_v^{ϕ} reaches approximately 10 at. pct owing to the $(u_C^{\gamma/\alpha} - u_C^{\gamma/0})$ and $(\delta U_C^{\gamma}/\delta X_C^{\gamma})$ decrease with the increase of X_v^{ϕ}.

It is evident that the decrease in activity should lead





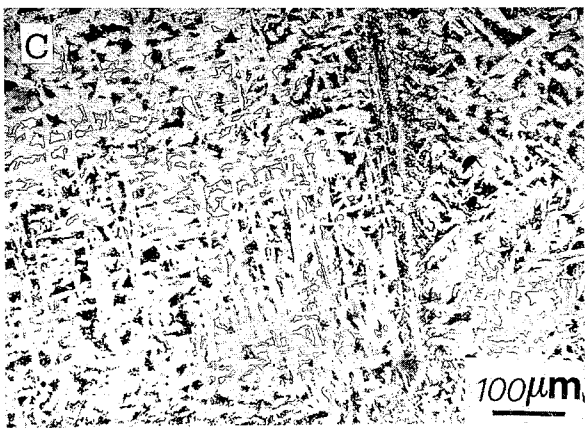


Fig. 11—Optical microstructures of Fe-0.42C-1.96Si-0.96V alloy treated at 500 °C for (*a*) 100 s, (*b*) 2000 s and (*c*) 7000 s.

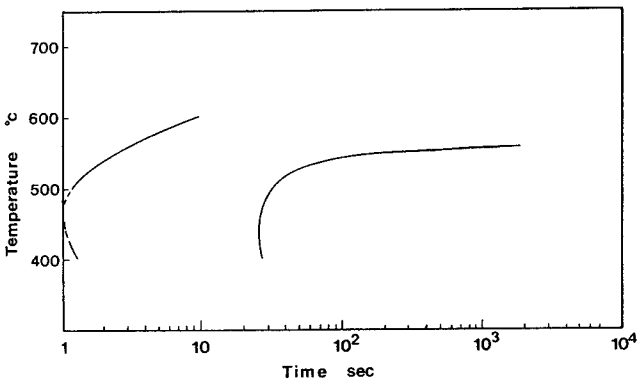


Fig. 12—TTT diagram of Fe-3.24Mn alloy.

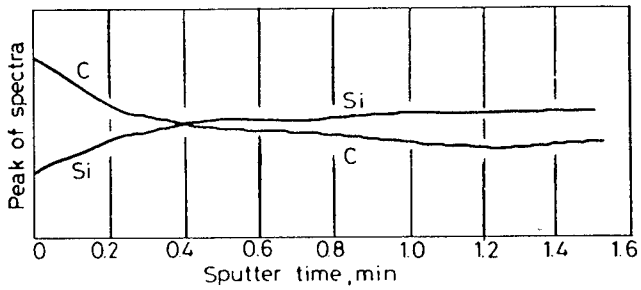


Fig. 13—Concentration profile of species C and Si of Fe-0.37C-2.23Si-2.29Mn alloy reacted at 475 °C for 20 s.

Table II. Calculated ΔG_v and Relative Incubation Time τ of Experimental Alloys

Alloy	Nose Temperature (°C)	ΔG_v (J/mol)	τ_{cal}	τ τ_{exp} (s)	τ^*_{rel}	τ_{rel}/τ_{cal}	$\frac{E_s}{E_s(\text{Fe-C-Si-Mn})}$
Fe-C-Si	450	−1354.9	1	2	1	1	—
Fe-C-Ni	565	− 653.0	4.3	8	4	0.93	0.91
Fe-C-Mn	450	− 707.2	3.58	58	29	8.10	—
Fe-C-V	435	−1949.5	0.49	5	2.5	5.10	0.88
Fe-C-Si-Mn	350	−1463.7	0.84	300	150	178.0	1.0
Fe-C-Si-V	400	−1779.66**	0.58	7	3.5	6.03	1.12
Fe-Mn	470	−1355.0	0.97	1	0.5	0.5	0.77

*$\tau_{rel} = \tau_{exp}/\tau_{exp}$(Fe-C-Si).
**Data provided by J.K. Chen and W.T. Reynolds, Jr. and calculated by the central atoms model.

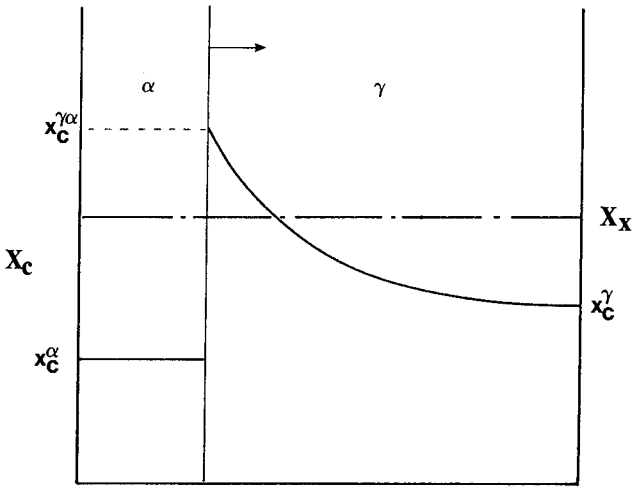


Fig. 14—Concentration profile of C and X according to the para-equilibrium model.

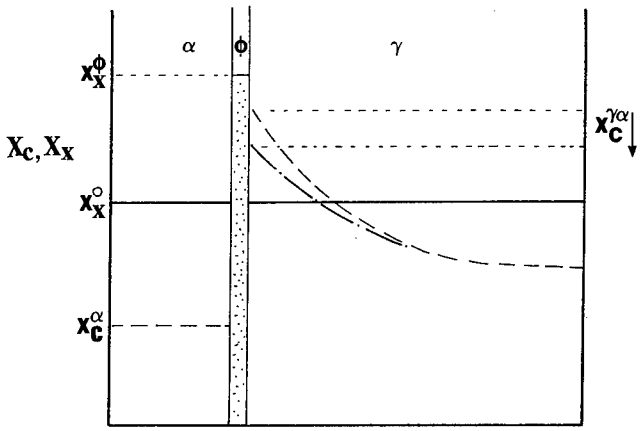


Fig. 15—Concentration profile of C and X according to the para-equilibrium model and considering X segregating at the α/γ boundary.

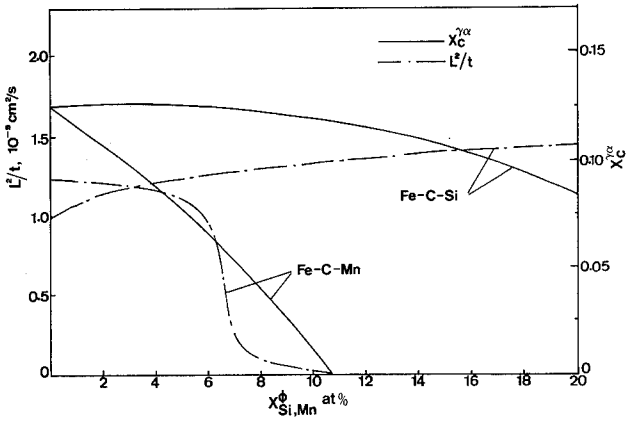
to the reduction of chemical potential. This is the case with the Fe-C-Mn and Fe-C-V alloys in which Mn and V are the solutes reducing the activity of C in γ. Therefore, as compared with an Fe-0.4C alloy, the $u_C^{\gamma/\alpha}$ of both alloys decreases regardless of whether the solute is an α or γ stabilizer. On the other hand, once an alloy contains a solute (Si or Ni) that increases $a_C^{\gamma/\alpha}$, the L^2/t value increases slightly with the increase of X_X^{ϕ} when the solute is either an α or γ stabilizer.

Figure 18 gives the L^2/t *vs* X_{Mn}^{ϕ} curves for alloys isothermally treated at 500 °C containing various concentrations of C and Mn. The $X_C^{\gamma/\alpha}$ *vs* X_{Mn}^{ϕ} curve of the Fe-C-Mn alloy at 500 °C is also drawn in the figure. If there is no segregation of Mn within the α/γ boundary, it is evident from the intersections of the curves with the ordinate that the differences among the L^2/t values for various Fe-C-Mn alloys are less than 0.5×10^{-9} cm^2/s. But to any Fe-C-Mn alloy shown in the figure, the L^2/t value descends rapidly with the increase of the segregation concentration of Mn within the α/γ boundary.

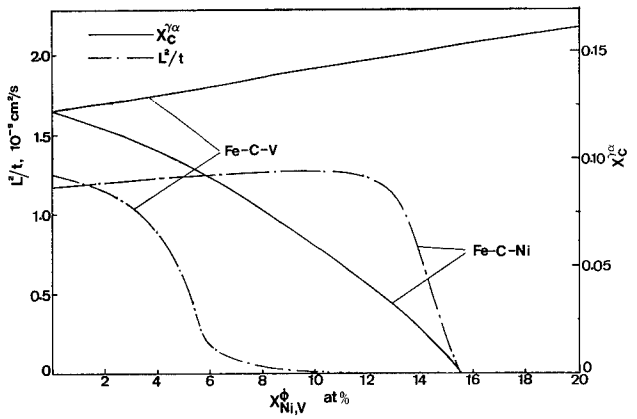
In view of the aforementioned statements, it is reasonable to suggest that the ferrite growth kinetics of alloys that contain solute that depress a_C^{γ} are inhibited by the SDLE, provided that the element segregates to the α/γ boundary. Thus, according to the proposal of SDLE, any Fe-C-X alloy which contains an element X that depresses the activity of C in γ may depress the growth kinetics of ferrite and shift the TTT diagram to the right. That is the case with the F-C-Mn, Fe-C-V, Fe-C-Si-Mn, and Fe-C-Si-V alloys.

B. *Morphology*

From the results given in Section A, we can observe that the morphology of ferrite is also closely related to the composition of the alloys. Those alloys that contain Si or Ni, and therefore raise the activity of C in γ, have normal Widmanstätten ferrite like that in Fe-C alloys. Although sympathetic nucleation occurs in these alloys and is responsible for branching of ferrite plates, the lengthening of the ferrite plates occurs mainly by continuous growth. Thus, most of the ferrite plates are tapered and approximately parallel to each other. As for the Fe-C-V and Fe-C-Mn alloys, a strong SDLE on α/γ boundaries during growth is expected due to the strong interaction between C and the solutes and the tendency of V and Mn to reduce the activity of C in γ. The SDLEs of these elements inhibit the migration of growth risers, and thus, the ferrite plates lengthen slowly. The relative lengthening and thickening rates cause ferrite to have a small aspect ratio; *i.e.*, degenerate plates mainly develop through the agency of repeated edge-to-edge and face-to-face sympathetic nucleation. Therefore, a ferrite plate

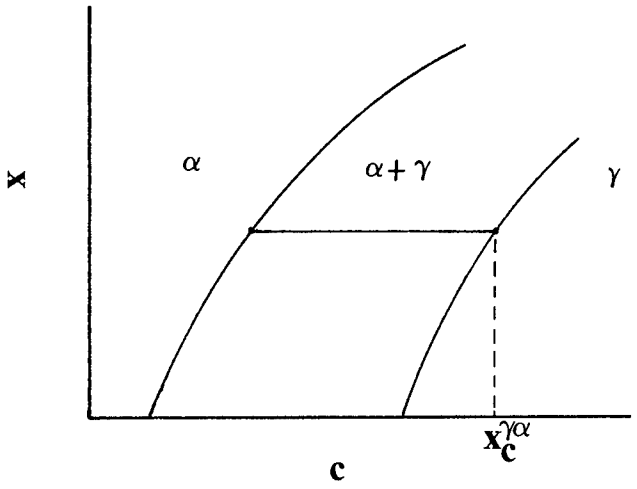


(*a*)

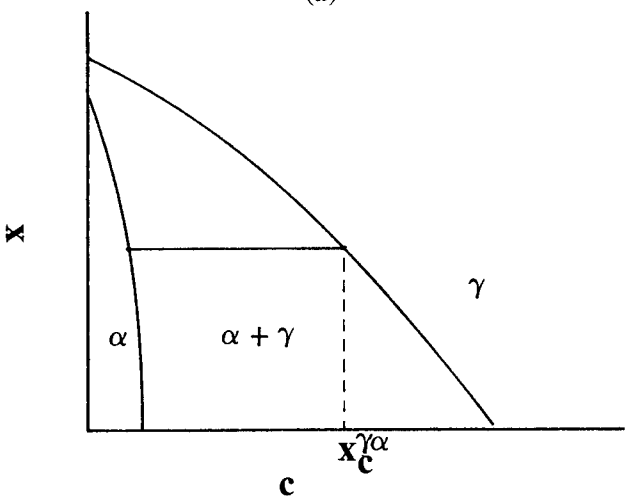


(*b*)

Fig. 16—$X_C^{\gamma/\alpha} \sim X_X^{\phi}$ and $L^2/t \sim X_X^{\phi}$ curves at 500 °C for the ternary alloys employed.



(*a*)



(*b*)

Fig. 17—Schematic Fe-C-X isotherms for which (*a*) X is an α-stabilizer and (*b*) X is a γ-stabilizer.

in the Fe-C-V and Fe-C-Mn alloys consists of individual ferrite crystals, the so-called subunits of bainite. Carbide precipitates in conjunction with the formation of the ferrite subunits in Fe-C-V and Fe-C-Mn alloys employed in this investigation.

The difference in morphology between the Fe-C-Mn, Fe-C-V, and the quaternary Fe-C-Si-Mn and Fe-C-Si-V alloys is due to the introduction of Si, which inhibits the precipitation of carbide from the austenite; thus, when the C concentration in austenite is high enough, a certain amount of retained austenite may remain among the subunits of ferrite sheaves.

C. *The Segregation of Solutes at the γ-Grain Boundary and Its Effect on the Formation of Ferrite Sideplates*

From Figure 13 and the experiments with Auger electron spectroscopy (AES), STEM plus energy-dispersive spectroscopy (EDS), and static secondary ion mass spectroscopy (SIMS) executed by the authors,[15] it is evident that C and Mn segregate to γ-grain boundaries during the isothermal treating of the Fe-C-Si-Mn alloy at intermediate temperatures for a short time. Since the diffusion rate of substitutional solutes along the γ grain boundary is rapid, the concentration of Mn and V at the γ grain boundary may be inherited easily by the newly formed α/γ interface during the formation of the α phase at the prior γ grain boundary. When decreasing the activity of C in γ adjacent to the α/γ interface, the segregation of Mn and V at the α/γ interface will inhibit the growth of ferrite sideplates. This is a reasonable explanation for the fact that the amount of ferrite sideplates is well below that of the intragranular ferrite plates in these alloys (Figures 9 and 10).

VI. SUMMARY

1. The effect of solutes on the growth kinetics of ferrite is consistent with their effect on the activity of C in γ. Solutes (Mn and V) that have a strong interaction with C and depress activity of C in γ depress grain-boundary ferrite growth kinetics significantly by segregating to α/γ boundaries and decreasing the $X_C^{\gamma/\alpha}$ and $u_C^{\gamma/\alpha}$ of the alloys. Therefore, SDLE might be the reasonable explanation for the formation rate of ferrite plate in these alloys.
2. Simplified calculations of $X_C^{\gamma/\alpha}$ and L^2/t are consistent qualitatively with the effects of the investigated

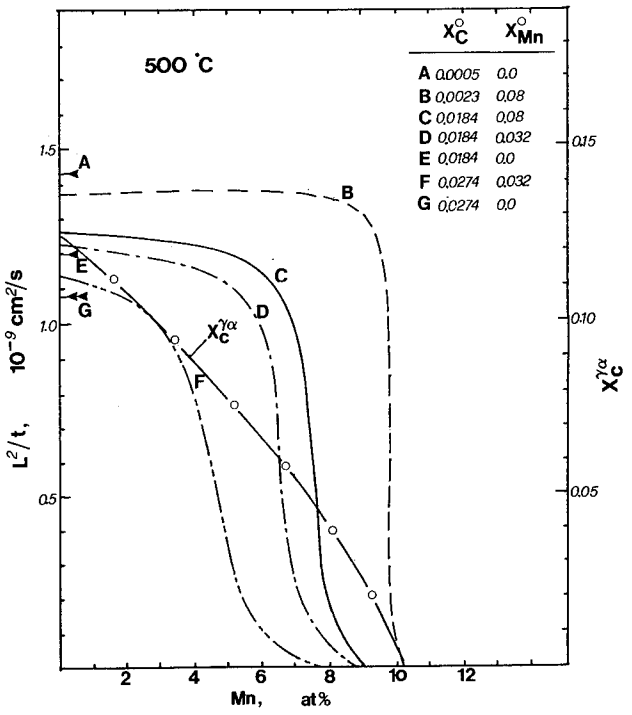


Fig. 18—$L^2/t \sim X_X^\phi$ curves for various Fe-C-Mn alloys marked in the figure.

solutes on ferrite plate-growth kinetics. However, more rigorous treatment of the calculations is needed to establish the quantitative relation between the SDLE and ferrite-growth kinetics.

3. The effects of solutes on the ferrite morphology are related to their effect on growth kinetics. The formation of degenerate ferrite plates in Mn- or V-containing alloys results from the SDLE of Mn and V.
4. Once ferrite nucleates at austenite grain boundaries, the newly formed α/γ boundary inherits the original segregation of Mn of V. The SDLE on them inhibits the growth of ferrite sideplates at prior γ grain boundaries. Therefore, the fraction of sideplates in Mn- and V-containing alloys evidently decreases.

ACKNOWLEDGMENTS

The authors wish to acknowledge financial support from the National Science Foundation of China through Grant No. 5860248, and Ph.D. W.T. Reynolds, Jr. for his helpful comments on the manuscript.

REFERENCES

1. K.R. Kinsman and H.I. Aaronson: *Transformation and Hardenability in Steel*, Climax Molybdenum Co., Ann Arbor, MI, 1967, p. 139.
2. W.T. Reynolds, Jr., F.Z. Li, C.K. Shui, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1433-63.
3. W.T. Reynolds, Jr., S.S. Brenner, and H.I. Aaronson: *Scripta Metall.*, 1988, vol. 22, pp. 1343-48.
4. I. Stark and G.D.W. Smith: *Phase Transformation 87*, Institute of Metals, London, 1988, p. 475.
5. G. Chen, H.I. Aaronson, S.K. Liu, and S.S. Brenner: Carnegie Mellon University and University of Pittsburgh, Pittsburgh, PA, unpublished research, 1989.
6. L. Dorn: *Schweißung und' schneiden*, 1979, vol. 29 (2), pp. 246-49.
7. T. Abe, H.I. Aaronson, and G.J. Shiflet: *Metall. Trans. A*, 1985, vol. 16A, pp. 521-27.
8. J.W. Cahn and W.C. Hagel: *Decomposition of Austenite by Diffusional Processes*, Interscience Publishers, New York, NY, 1962, p. 131.
9. J.S. Kirkaldy, B.A. Thomson, and E.A. Baganis: *Hardenability Concepts with Applications to Steel*, D.V. Downe and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, pp. 82-124.
10. J. Feder, K.C. Russell, J. Lothe, and G.M. Pound: *Adv. Phys. (Phil. Mag. Suppl.)*, 1966, vol. 15(1), p. 111.
11. L. Yang, S.K. Liu, and P. Yang: *Acta Metall.*, 1993, vol. 29, p. A1.
12. W.T. Reynolds, Jr.: Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA, 1988.
13. C. Zener: *J. App. Phys.*, 1949, vol. 20, pp. 950-53.
14. B. Uhrenius: in *Hardenability Concepts with Applications to Steel*, D.V. Downe and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, pp. 28-81.
15. S.K. Liu and J. Zhang: *Metall. Trans. A*, 1990, vol. 21A, pp. 1517-25.

Formation Mechanism of Bainitic Ferrite and Carbide

HONG-SHENG FANG, JIA-JUN WANG, and YAN-KANG ZHENG

Superledges on the broad faces of bainitic ferrite plates have been observed with transmission electron microscope (TEM). The observed superledges, ranging from less than 1 to 24 nm in height, are imaged in three dimension by way of tilt operation under TEM. Also, an array of smaller superledges are observed to pile up in front of a secondary phase. Pileup of superledges in front of a barrier is indicative of the mobility of an individual superledge. The precipitation of carbide associated with bainitic ferrite is also studied. It is observed that a carbide of a wedgelike shape may exist in front of a superledge with its tip (*i.e.*, thinner end) penetrating the austenite and its root (the other end) originating at $\alpha:\gamma$ boundary. This condition indicates that the observed carbides may nucleate at the austenite side of $\alpha:\gamma$ phase boundary and grow toward austenite matrix.

I. INTRODUCTION

DURING an investigation of the various morphologies of proeutectoid ferrite in plain carbon steels, Dubé *et al.*[1] devised a system for morphologies, which is known as the Dubé morphological classification system. The system was then extended by Aaronson[2] and applied to describe the different morphologies of proeutectoid cementite.[3] Many of the precipitates involved in this classification scheme have a plate morphology.[4] In addition to proeutectoid ferrite and cementite, which form at relatively high temperatures in steel, martensite and bainite are also formed in the shape of a plate. After more than 50 years of debate,[5] considerable disagreement remains as to whether a bainite plate forms by shear or diffusion.[6,7]

Note that the term "shear"[8] is used here to define the "military"[9,10,11] process in which the movement of atoms on one side of the interphase boundary is coordinated simultaneously or precisely by the sequential glide-type motion of both substitutional and interstitial atoms across the boundary. This behavior of well-coordinated motion requires glissile interphase boundaries.[9] The term "diffusion" is used in this article to describe a nonmilitary process in which atoms migrate toward or away from the interphase boundary individually, *i.e.*, a diffusion process is accomplished by the biased diffusional random walk of substitutional and interstitial atoms.[12] We still lack the methods needed to study directly the movement of individual atoms across the phase boundary during transformations. Thus, we must rely on the analysis of some experimental results, which provide us with some indirect but very important clues to the atomic mechanism of growth.

This article will deal with the roles of superledges in the formation mechanism of ferrite plates in steel. Although bainitic carbides are not plates, their formation mechanism is also discussed because bainitic ferrite plate precipitation is usually accompanied by that of carbides. The formation mechanism of bainitic carbides is closely associated with that of ferrite plates.

II. EXPERIMENTAL METHODS

Specimens of an Fe-2.17C-12.18Cr-0.26Mn-0.31Si (wt pct) steel were homogenized for 69 hours at 1423 K, austenitized for 10 minutes at 1373 K in evacuated quartz capsules, and isothermally reacted for 22 or 29 hours at 573 K in a deoxidized salt bath after the capsules were broken. The alloy composition and heat-treatment process were designed to avoid martensite formation. The M_s temperature in this alloy is about 223 K. Thin films for the transmission electron microscopy (TEM) study were prepared by chemically thinning samples to 0.04 mm in a solution of 5 pct HF and 95 pct H_2O_2, polishing in a dual-jet unit with 5 pct $HClO_4$, 25 pct glycerol, 70 pct alcohol, and finally, ion milling. The TEM studies on this alloy were performed with an H800 electron microscope operating at 200 KV.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The optical microstructure of the experimental alloy that reacted for 29 hours at 573 K is presented in Figure 1. The bainite is shown by the dark etching and is plate shaped, as indicated by arrow A in Figure 1. The matrix around an individual bainite plate is retained austenite (arrow B). Additionally, $(CrFe)_7C_3$ carbide (arrow C) of equiaxed shape, which is not dissolved during the austenitizing process, is also present.

Note that owing to the presence of $(CrFe)_7C_3$ carbide, the average carbon content in austenite and bainite is approximately 1.2 wt pct,* which is quite different from

HONG-SHENG FANG and YAN-KANG ZHENG, Professors, and JIA-JUN WANG, Ph.D., are with the Department of Materials Science and Engineering, Tsinghua University, Beijing 100084, the People's Republic of China.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

*As to the carbon composition of the austenite in the experimental alloy, the authors have tried to determine it by several methods. First, the carbon concentration was measured to be 1.34 wt pct through wave-dispersed spectroscopy when a $(CrFe)_7C_3$ standard sample was used. Second, the calculated carbon concentration of the austenite was determined to be approximately 1.2 wt pct by the method of quantitative metallography. The area ratio of $(CrFe)_7C_3$ in the alloy was determined by computer image processing systems under scanning electron microscopy (SEM). Finally, the thermodynamics calculation

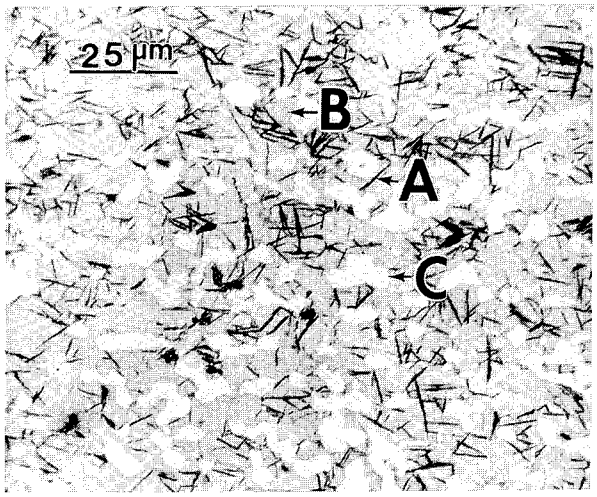


Fig. 1—Optical microstructure of an Fe-2.17C-12.18Cr-0.26Mn-0.31Si alloy reacted 29 h at 573 K.

is also performed with this alloy. It shows that the equilibrium carbon concentration in the austenite at the austenitizing temperature given here (1100 °C) is roughly 0.96 wt pct. On the basis of the previous experimental results, it seems to be reasonable that the actual carbon concentration in the austenite in this experimental alloy is about 1.2 wt pct.

the overall alloy composition, 2.17 wt pct, as given in Section II, determined by quantitative chemical analysis.

The TEM microstructure of the Fe-2.17C-12.18Cr-0.26Mn-0.31Si alloy reacted for 22 hours at 573 K is shown in Figure 2(a). Figure 2(c) is the diffraction pattern of the ferrite plates in Figure 2(a). The indexing results are given in Figure 2(d).* The corresponding dark-field image of Figure 2(a) is presented in Figure 2(b), using the $(0\bar{1}1)_o$ reflection. Apparently, there are three ferrite plates inside the concerned field, which are respectively labeled A, B, and C. Plates B and C are laths, while plate A is a wedge.

*There are two sets of complete patterns in Figure 2(c), which are produced by $[01\bar{1}]$ and $[\bar{1}\bar{3}\bar{3}]$. Apparently, the latter set of patterns is asymmetrical; this indicates that the beam direction is tilted from the given zone axis. Moreover, in addition to these patterns, there are still three unknown diffraction points in Figure 2(c). Based on the X-ray diffraction result which shows that there are only three phases in the heat-treated alloy, ferrite, austenite, and $CrFe_7C_3$, the three points may be arbitrarily indexed as $(\bar{2}\bar{4}4)_\gamma$, $(24\bar{4})_\gamma$, and $(51\bar{3})_\gamma$. They may be generated by $[\bar{4}\bar{7}9]_\gamma$.

As discussed in detail by Aaronson *et al.*,[8] ledges on the broad faces of plates can be of various types. They include structural ledges,[13] misfit compensating ledges, and growth ledges. The latter type of ledge, as observed on the broad faces of plate-shaped phases,[14] is often much greater in height than the ferrite unit cell, ranging from tens of nanometers to about 1 μm.[15] These growth ledges are termed superledges.[16]

After slightly tilting the TEM film from the orientation of Figure 2(a), the ferrite plate A was again photographed in Figure 3(a). As indicated by arrows A through F in Figure 3(a), an array of ledges with apparent heights of 4.6, 22, 3.4, 8.6, 12, and 16 nm, respectively, are observed at the broad face of this ferrite plate. It is obvious that the apparent heights of the superledges given previously are different from the actual heights, since thickness fringes can be seen in front of the riser. But according to the distance between two neighboring thickness fringes the angle of inclination of the riser can be calculated to be greater than 76 deg, *i.e.*, the error owing to the inclination is less than 2.6 pct. Apparently, this has little influence on the measurement of the heights of the superledge risers. Meanwhile, since the observed ledges are 10 times higher than the unit cell, they are clearly superledges, as previously defined. The superledges in front of arrows a and b are also imaged in Figure 3(b) through adjusting the orientation of the TEM film by way of tilt and azimuth. It is clear that the superledges indicated by arrows a and b in Figure 3(a) are composed of several smaller superledges at the heads of arrows a, b, c, and d in Figure 3(b).

Christian and Edmonds[17] have claimed that superledges cannot exist because the Burgers vector associated with the riser of a superledge would be too large. But the experimental observation of arrays of superledges during the present investigation as well as many previous works[18] contradicted the Christian-Edmonds view of superledges.

The observed superledges provide an important evidence supporting diffusion-controlled ledgewise mechanism of bainite plate growth. However, the problem of whether a bainite plate grows by shear or diffusion still remains, since more evidence is needed showing that the thickening of bainite occurs through the lateral migration of the risers of a superledge by diffusional mechanism. The nature and mobility of superledges will be discussed in the following sections of this article.

The plates of lower bainitic ferrite so far discussed contain no carbide. Nevertheless, lower bainite is usually considered to be a mixture of bainitic ferrite and carbides. It is apparent that the formation of lower bainitic carbides is closely associated with the growth of bainitic ferrite and, thus, forms a subject for further investigation in the present context.

A bright-field image of a bainite sheaf consisting of two subplates is shown in Figure 4(a). Superledges are observed at arrow heads a, b, c, d, and e on the broad faces of these plates. Highly magnified micrographs of Figure 4(a) are presented in Figures 4(b) and 4(f). Figure 4(f) was imaged after the foil was tilted several degrees away from the orientation of Figure 4(b).

It can be clearly seen that there are superledges on the broad faces of bainitic subplates in Figures 4(b) and 4(f). Figure 4(d) is the diffraction pattern of the structure in Figure 4(b). Indexing of this pattern in Figure 4(e) indicates that it consists of three simple patterns, respectively reflected by $[1\bar{1}0]_\gamma$, $[01\bar{3}]_\theta$, and $[4\bar{3}3]_{(CrFe)_7C_3}$. A centered dark-field image of $[\bar{1}\bar{1}3]_\theta$ corresponding to Figure 4(b) is presented in Figure 4(c). As indicated by arrows a, b, c, d, and e in Figures 4(a), 4(b), and 4(c), the carbide precipitated from austenite in front of a superledge in the vicinity of the superledge risers (and presumably in contact with both the riser and the "lower" terrace in each case).

The carbide is the shape of a wedge with the root (thick part) at the α/γ interface and the tip (thin part) within the austenite matrix. After tilting the TEM foil

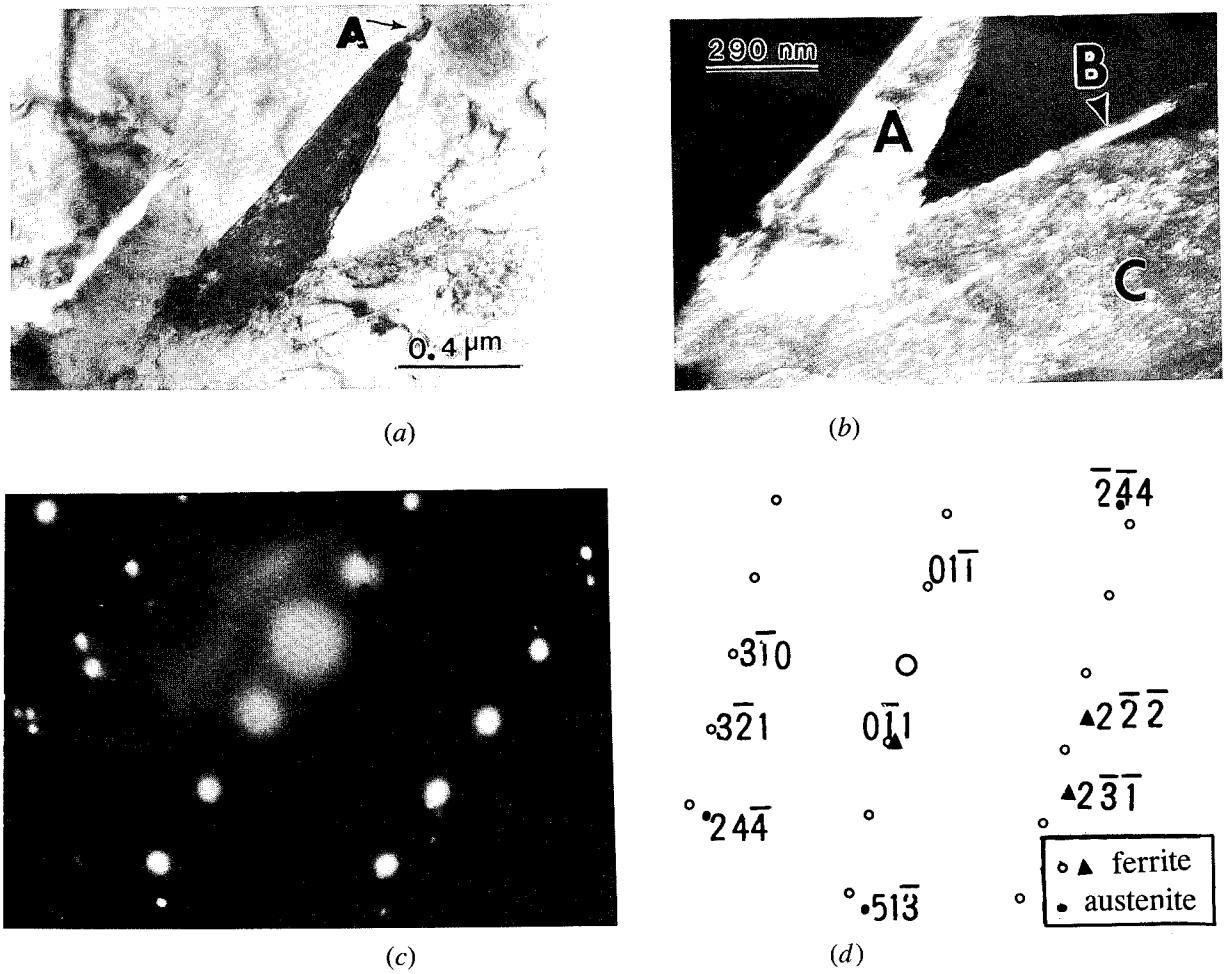


Fig. 2—TEM microstructure of an Fe-2.17C-12.18Cr-0.26Mn-0.31Si alloy reacted 22 h at 573 K: (*a*) bright-field TEM image of lower bainite ferrite; (*b*) dark-field micrograph using $(0\bar{1}1)_\alpha$ reflection; (*c*) diffraction pattern; and (*d*) indexing result, including $[21\bar{1}]_\alpha$, $[\bar{1}3\bar{3}]_\alpha$, and $[\bar{4}\bar{7}\bar{9}]_\gamma$.

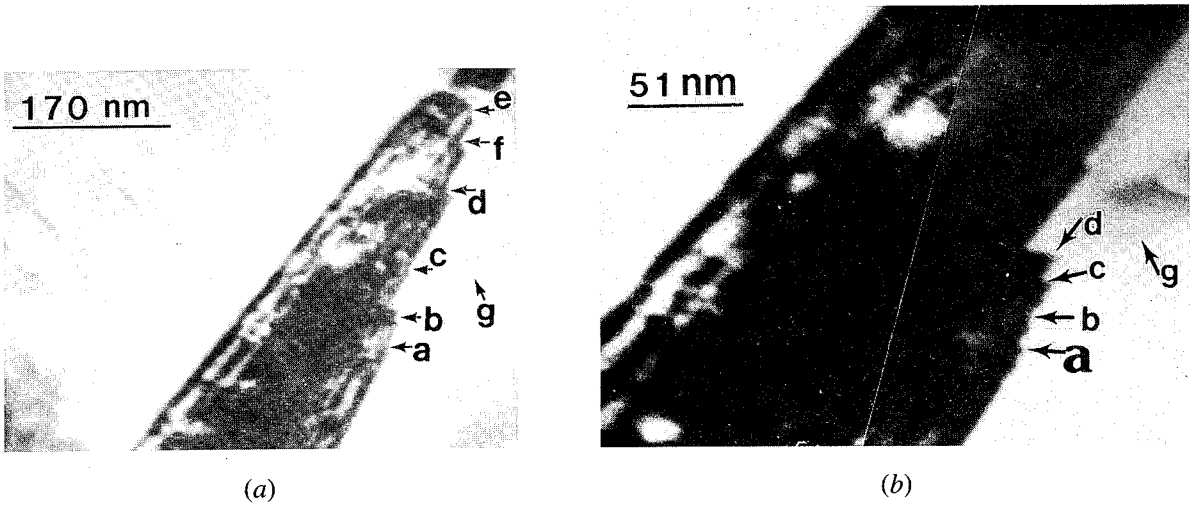
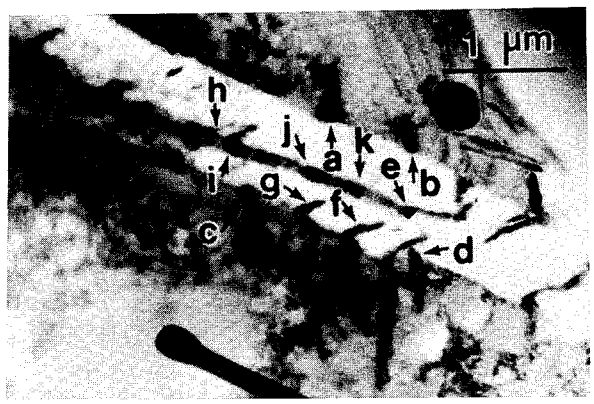
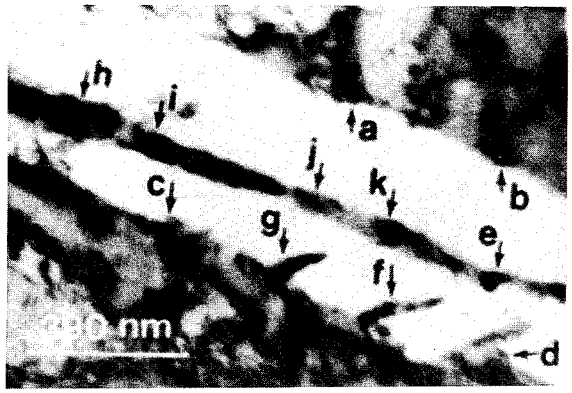


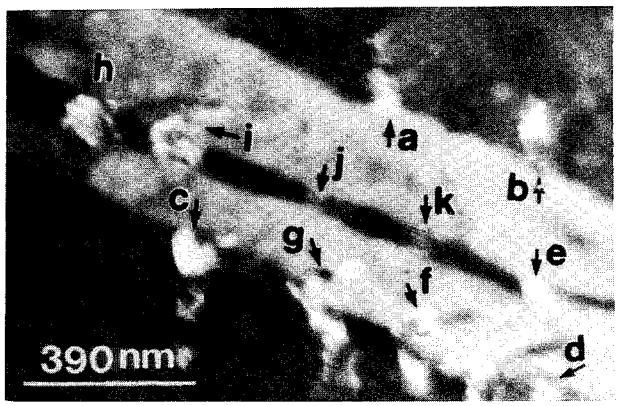
Fig. 3—An array of superledges on the broad faces of bainite ferrite in an Fe-2.17C-12.18Cr-0.26Mn-0.31Si alloy, reacted 22 h at 573 K, shown in (*a*) and (*b*) the same area as (a) but with different magnification and different orientation through adjusting the TEM foil.



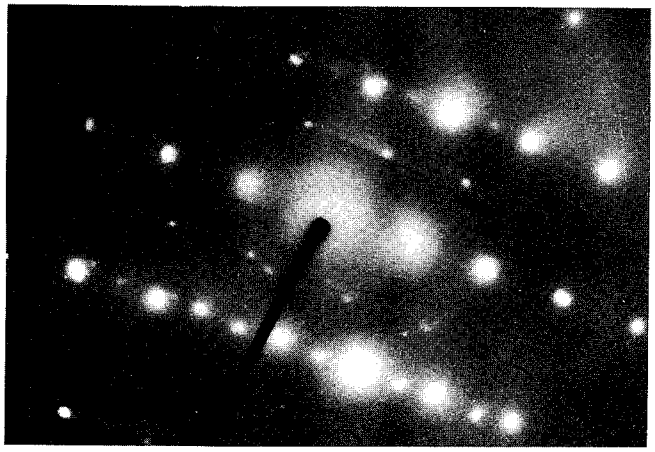
(a)



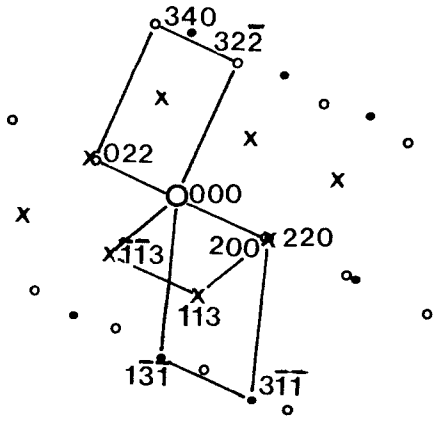
(b)



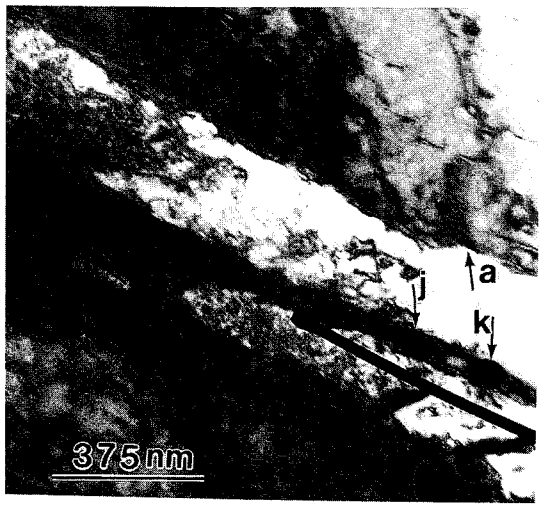
(c)



(d)



(e)



(f)

Fig. 4—Bainitic carbides precipitating in an Fe-2.17C-12.18Cr-0.26Mn-0.31Si alloy, reacted 29 h at 573 K: (*a*) bright-field image; (*b*) enlargement of (a); (*c*) dark-field image of $(113)_\alpha$ corresponding to (b); (*d*) diffraction pattern; (*e*) indexing result; and (*f*) imaged after slightly tilting the foil from (b), again showing the presence of carbides.

away from the orientation of Figure 4(b), the carbides do not disappear, as shown by arrow a in Figure 4(f).

It was reported by Bhadeshia and Christian[19,20] and by many workers earlier[9] that bainitic cementite nucleates and grows within supersaturated ferrite. The cementite is said to be confined to the ferrite and hardly penetrates the austenite; occasional impingements of carbide with the α/γ interface can be regarded as chance events. Contrary to this scenario, the present results show that many carbides are in contact with α/γ interfaces. Thus, the impingements cannot be regarded as chance events. They can only be interpreted as a consequence of carbide precipitation at ferrite:austenite interfaces.

The nucleation and growth of a carbide in contact with the α/γ interface can be visualized schematically, as shown in Figure 5. Supposing a ferrite plate thickens by means of superledges, steep carbon concentration gradients are thus formed around each superledge. The carbon profile in the adjacent austenite in front of a superledge is schematically illustrated in Figure 5(b).[21–24] Obviously, the nearer to the riser, the greater the degree to which the carbon concentration in austenite is enriched. This will provide a higher driving force for carbide nucleation. Moreover, a carbide will tend to nucleate at the riser of a superledge because of the destruction of portions of both the riser and the adjacent (lower) terrace. Thus, a carbide will form in front of a superledge, as illustrated in Figure 5(c).

The wedge morphology of a carbide in Figure 5(d) can be attributed to both crystallography and carbon supersaturation in austenite. First, the carbide must have low-energy nucleus/matrix boundaries and, hence, have partially coherent interfaces with the adjacent austenite.[25,26] The carbide forms in the shape of a plate with partially coherent broad faces at its early stage; *i.e.*, the orientation of a carbide plate is determined by the crystallography factor. Second, the less supersaturation of carbon atoms farther away from the interface may give rise to the wedge shape, while carbide grows toward the austenite matrix.

It should be pointed out that bainitic carbides can also nucleate in the austenite away from superledges. Arrow h in Figures 4(a) and (b) show carbides precipitated on the broad faces of ferrite plates away from superledges. Arrows i, j, and k in Figures 4(a) through (c) show the carbides precipitated between two subplates, and carbides in the austenite between two subunits are indicated by arrows f and g. This is consistent with some previous works.[27,28,29] As reported by Spanos *et al.*[29] lower bainitic carbide can exist inside bainitic ferrite, owing to the growth of ferrite after the carbide forms within austenite regions between ferrite crystals.

Another point is that the model given in Figure 5 is not the only one to describe the process of carbide precipitation. Suppose that a carbide first nucleated on the broad face of a bainite plate and then grew into adjacent austenite; this may also lead to pileup of superledges in front of the carbide. But the model given in Figure 5 is more effective to interpret the very important fact that carbides precipitate in front of the risers of superledges in Figures 4(a) through (c). As previously mentioned in this section, the riser of a superledge is actually a semicoherent interface; in other words, the structure of the riser is the same as that of the terrace (broad face) except that there are many small ledges or kinks on the riser. The risers of these small ledges and kinks are incoherent interfaces; this is required to fulfill a diffusion-controlled process. Apparently, the movement of kinks and ledges will lead to the enrichment of carbon in front of the riser of a superledge. Now that the structures of a riser and a terrace are the same, the carbide will prefer to precipitate in front of a riser of a superledge where more carbons are enriched, especially at the junction of the terrace and the riser. Moreover, it has not been reported in other published articles, and not observed by us, that a wedgelike carbide precipitates, independently, *i.e.*, far away from a riser superledge, on the broad faces of bainitic ferrite plates.

In addition to carbide precipitation at the risers of superledges, the sympathetic nucleation of other ferrite plates at or near these risers has been identified as a significant hindrance to the lateral growth of superledges. It has been pointed out that the ferrite plates in Figure 2(b) are of slight different orientations among each other. Measured directly from the indexing results in Figure 2(d), $(0\bar{1}1)_\alpha$ of plate A is about 3 deg. away from $(0\bar{1}1)_\alpha$ of plate C. Figure 6 is imaged by the reflection of $(0\bar{1}1)_\alpha$ of both plates A and C. Bainitic plates, indicated by arrow g in Figures 6 and 3(a), are observed to be in front of superledges. The orientations of the

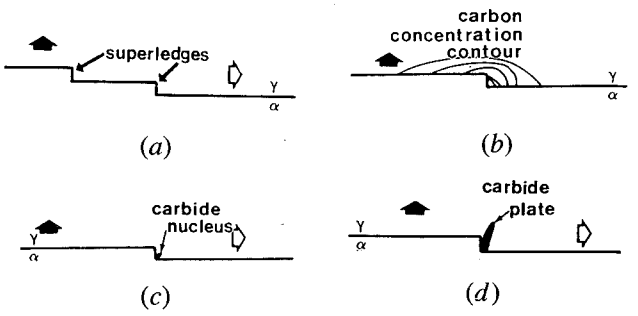


Fig. 5—Schematic formation model of bainitic carbide: (*a*) growth ledges on the broad face of ferrite; (*b*) carbon concentration gradient in front of the riser; (*c*) carbide nucleation in front of the riser; and (*d*) carbide growth toward the γ matrix.

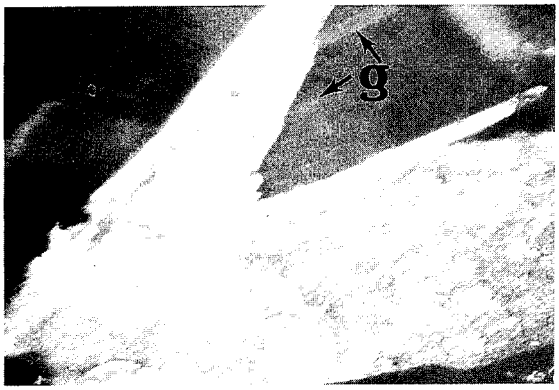


Fig. 6—Pileup of superledges in front of a barrier; imaged by $(\bar{0}11)_\alpha$ reflection.

plates that may sympathetically nucleate in front of the riser, as indicated by arrow g in Figures 3(a) and (b) and Figure 6, are away from that of plate A. Meanwhile, more attention should be paid to the fine structure of the superledge indicated by arrow b in Figure 3(a). The superledge is composed of four small superledges, as shown by arrows a, b, c, and d in Figure 3(b), whose heights are approximately 2.8, 1.2, 6.3, and 12.1 nm, respectively; *i.e.*, a series of superledges have piled up owing to the presence of the barrier.

If a platelike precipitate plate thickens through a ledgewise mechanism, there should exist superledges on the broad faces of the plate, as shown in Figure 7(a). The fine structure of the terrace was investigated.[13] The terrace of the superledge is semicoherent in order to minimize the interfacial energy. Thus, the terrace of a superledge could not migrate by diffusion of atoms across the semicoherent boundary; *i.e.*, it would be sessile owing to its complex structure.[13] The riser, which was supposed to be incoherent in some earlier work,[12] appears to be partially coherent, but with some small ledges or kinks whose risers are expected to be the actual sites of atomic attachment and detachment.[8] In other words, the migration of the riser of a superledge might actually be carried out through the discontinuous lateral migration of risers of these smaller ledges or kinks.

The pileup of superledges is illustrated schematically in Figure 7. In the case of a plate thickening through a ledgewise mechanism (shown in Figure 7(a)), in which there is a barrier that might be another plate crystal of either the same or other phases, such as another bainite plate, carbide, *etc.*, in front of the riser of a superledge (Figure 7(b)), the movement of the superledge riser will be hindered; *i.e.*, the riser of the superledge is "anchored." Thus, pileup of an array of superledge occurs, as shown in Figure 7(c). Pileup of superledges indicates that the risers of these smaller superledges are reasonably mobile.

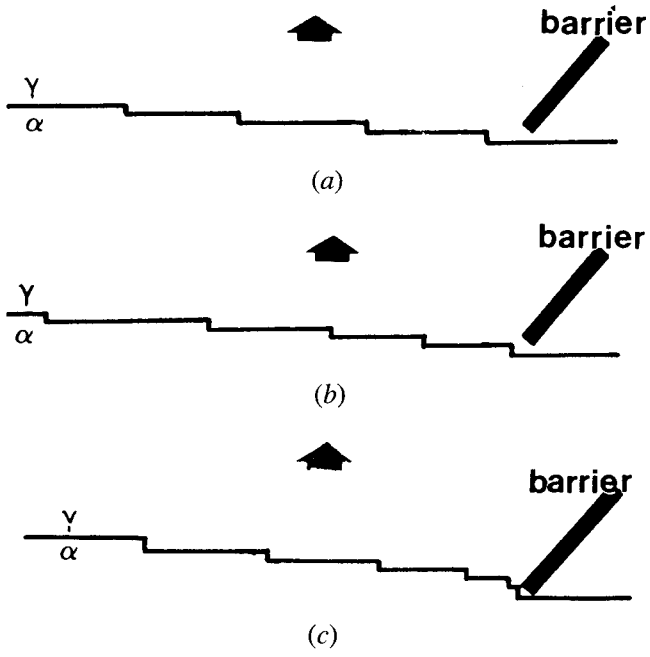


Fig. 7—Schematic model for the pileup of superledges in front of a barrier: (*a*) pileup at the earlier stage; (*b*) pileup after the forward superledge hindered; and (*c*) pileup of superledges.

IV. CONCLUSIONS

1. Carbides nucleate in the γ matrix at the α/γ boundary in the austenite side and also in the austenite between two subunits or two subplates, and grow toward the austenite matrix.
2. Carbides precipitate at the risers of superledges, probably at the junctions between the risers and the "lower" terraces.
3. Superledges are observed on the broad faces of bainitic ferrite plates.
4. Superledges were observed to pile up in front of a barrier phase; this is indicative of the mobility of piled up superledges.

ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China, Beijing Laboratory of Electron Microscope, Chinese Academy of Science and Zhongguancun Measurement and Analysis Fund of the People's Republic of China. The authors are grateful for the careful review and valuable suggestions of Professor Hubert I. Aaronson of the United States Naval Research Laboratory. Appreciation is also expressed to Dr. George Spanos of the United States Naval Research Laboratory for his valuable suggestions.

REFERENCES

1. C.A. Dubé, H.I. Aaronson, and R.F. Mehl: *Kev. Met.*, 1958, vol. 201, pp. 55-57.
2. H.I. Aaronson: *Symp. on Nucleation of Phase Transformations in Metals*, Institute of Metals, London, 1955, pp. 47-50.
3. E.W. Heckel and H.W. Paxton: *Trans. ASM*, 1960, vol. 53, p. 539.
4. H.I. Aaronson: *Decomposition of Austenite by Diffusion Processes*, Interscience Publishers, New York, 1962, p. 387.
5. E.S. Davenport and E.C. Bain: *Trans. AIME*, 1930, vol. 90, p. 117.
6. R.F. Hehemann, K.R. Kinsmann, and H.I. Aaronson: *Metall. Trans.*, 1972, vol. 3, pp. 1077-94.
7. H.K.D.H. Bhadeshia and D.V. Edmonds: *Metall. Trans. A*, 1979, vol. 10A, pp. 895-907.
8. H.I. Aaronson, T. Furuhara, J.M. Rigsbee, W.T. Reynolds, Jr., and M.J. Howe: *Metall. Trans. A*, 1990, vol. 21A, pp. 2369-2409.
9. J.W. Christian: *Decomposition of Austenite by Diffusion Processes*, Interscience Publishers, New York, NY, 1962, pp. 371-86.
10. J.W. Christian: *Mechanisms of Phase Transformations in Crystallography Solids*, Institute of Metals, London, 1969, p. 129.
11. J.W. Christian: *Phase Transformations*, Institution of Metallurgists, London, 1979, vol. 1, p. 1.
12. H.I. Aaronson, C. Laird, and K.R. Kinsman: *Phase Transformations*, ASM, Metals Park, OH, 1970, pp. 313-96.
13. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 351-63.
14. T. Furuhara, A.M. Dalley, and H.I. Aaronson: *Scripta Metall.*, 1988, vol. 22, pp. 1509-14.
15. H.I. Aaronson, E. Eichen, and K.R. Kinsman: *Metall. Trans. A*, 1975, vol. 6, pp. 303-17.

16. H.I. Aaronson: private communication, 1993.
17. J.W. Christian and D.V. Edmonds: *Phase Transformations in Ferrous Alloy*, TMS, Warrendale, PA, 1984, p. 293.
18. H.I. Aaronson, D. Eylon, and F.H. Froes: *Scripta Metall.*, 1987, vol. 21, pp. 1241-45.
19. H.K.D.H. Bhadeshia: *Acta Metall.*, 1980, vol. 28, pp. 1103-14.
20. H.K.D.H. Bhadeshia and J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 767-97.
21. G.J. Jones and R. Trivedi: *J. Appl. Phys.*, 1971, vol. 42, pp. 4299-4304.
22. C. Atkinson: *Proc. R. Soc. London*, 1981, vol. A378, p. 351.
23. G.J. Jones and R. Trivedi: *J. Crystallogr. Growth*, 1975, vol. 29, p. 155.
24. C. Atkinson: *Proc. R. Soc. London*, 1982, vol. A384, p. 107.
25. H.I. Aaronson: private communication, 1992.
26. R.W.K. Honeycombe: *Metall. Trans. A*, 1976, vol. 7A, pp. 915-36.
27. G. Spanos and H.I. Aaronson: *Scripta Metall.*, 1988, vol. 22, pp. 1537-42.
28. Q. Jin and H.S. Fang: *Metall. Trans. A*, 1990, vol. 21A, pp. 2637-41.
29. G. Spanos, H.S. Fang, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1381-90.

Formation Mechanism of Bainitic Ferrite in an Fe-2 Pct Si-0.6 Pct C Alloy

KANEAKI TSUZAKI, AKI KODAI, and TADASHI MAKI

The bainite transformation at 723 K in an Fe-2 pct Si-0.6 pct C alloy (mass pct) was investigated with transmission electron microscopy (TEM) and quantitative metallography to clarify the growth mechanism of the ferritic component of bainite. In early stages of transformation, the bainitic ferrite was carbide free. The laths of bainitic ferrite within a packet were parallel to one another and separated by carbon-enriched retained austenite. The average carbon concentration of the bainitic ferrite was estimated to be 0.19 mass pct at the lowest, indicating that the ferrite was highly supersaturated with respect to carbon. The laths did not thicken during the subsequent isothermal holding, although they were in contact with austenite of which the average carbon concentration was lower than the paraequilibrium value. In the later stage of transformation, large carbide plates formed in the austenite between the laths, resulting in the decrease in the carbon concentration of the austenite. Subsequently, the ferrite with a variant different from the initially formed ferrite in the packet was decomposed for the completion of transformation. The present results indicate that the bainitic ferrite develops by a displacive mechanism rather than a diffusional mechanism.

I. INTRODUCTION

THE mechanism of bainite transformation in steels has not yet been established despite much work both experimentally and theoretically; two completely different models, *i.e.*, a diffusional and a displacive mechanism, have been proposed.[1–5] The diffusional school proposes that the ferritic component of bainite develops by a diffusional ledge mechanism. On the other hand, the displacive school claims that the bainite transformation occurs diffusionlessly, as far as the solvent and substitutional solute atoms are concerned.

For clarifying the transformation mechanism, it is of great importance to examine the thickening process and the carbon concentration of bainitic ferrite. If bainitic ferrite forms by the diffusional mechanism with a local equilibrium condition, it has the carbon concentration in equilibrium with respect to austenite. Moreover, it can grow until the average carbon concentration of the untransformed austenite reaches the equilibrium value with respect to ferrite, if the so-called "solute drag-like effect"[1,2,5] is absent. On the other hand, if bainitic ferrite develops by a displacive mechanism, the carbon concentration of it can be higher than the equilibrium value with respect to austenite. Moreover, it may not thicken after growing to a limited size due to a loss of coherency of the interphase boundaries.[6]

The measurements of the thickening process and carbon concentration of bainitic ferrite are usually difficult for the bainite transformation of conventional carbon steels because of the rapid precipitation of carbides and the absence of untransformed austenite within a bainitic packet or sheave. However, unlike the cases of conventional carbon steels, it is now well known that in the carbon steels containing about 2 mass pct silicon, carbide-free, lath-shaped, bainitic ferrite forms and is separated one from another by carbon-enriched retained austenite in early stages of transformation around 700 K.[7–11] In this case, we can measure the carbon concentration of retained austenite by, *e.g.*, X-ray diffraction analysis. Then, because it is carbide free, we can estimate the carbon concentration of bainitic ferrite from the mass balance of carbon, when the volume fractions of bainitic ferrite and austenite are obtained. Moreover, because bainitic ferrite laths in a packet are in contact with austenite,[7–11] we can examine the thickening of bainitic ferrite by measuring the change in lath width distribution with the reaction time. Fe-Si-C alloys have another benefit for the examination of the ferrite growth; the solute drag-like effect is not necessary to be taken into consideration, because silicon increases carbon activity in austenite.

In the present study, the bainite transformation during the isothermal holding at 723 K in an Fe-2 pct Si-0.6 pct C alloy (mass pct) was investigated with particular emphasis on the thickening and carbon concentration of bainitic ferrite in order to clarify the growth mechanism of the ferritic component of bainite.

II. EXPERIMENTAL PROCEDURE

A high-purity Fe-2 pct Si-0.6 pct C alloy (mass pct) prepared by vacuum induction melting was used. The chemical composition and M_s temperature of the alloy is shown in Table I. The ingot was homogenized at 1473 K for 18 ks followed by hot swaging and rolling at 1473 K to a plate with a thickness of 15 mm. After removing a surface layer of 2.5 mm from both sides, the plates were

KANEAKI TSUZAKI, Associate Professor, and TADASHI MAKI, Professor, are with the Department of Materials Science and Engineering, Kyoto University, Sakyo-ku, Kyoto 606-01, Japan. AKI KODAI, formerly Graduate Student, Kyoto University, Kyoto 606-01, Japan, is Researcher, with the Material Research Department, Akashi Technical Institute, Kawasaki Heavy Industries, Ltd., Kawasaki-cho, Akashi 673, Japan.

This article is based on a presentation made at the Pacific Rim Conference on the "Roles of Shear and Diffusion in the Formation of Plate-Shaped Transformation Products," held December 18–22, 1992, in Kona, Hawaii, under the auspices of ASM INTERNATIONAL's Phase Transformations Committee.

cut into specimens 15 × 10 × 2 mm in size for microstructural observations and X-ray measurements. Specimens were austenitized at 1423 K for 900 s in a vacuum furnace and then quenched into a salt bath set at 723 K. After isothermal holding for various periods of time, the specimens were finally quenched into water.

After removing a surface layer of 0.5 mm or more, the specimens were submitted to X-ray measurements and light microscopy. Three percent nital and an 8-g $Na_2S_2O_5$ + 100-mL H_2O solution were used for etchant to reveal bainitic ferrite and martensite, respectively.

Thin foils for transmission electron microscopy (TEM) were thinned in a solution of 75-g CrO_3 + 400-mL CH_3COOH + 21-mL H_2O using a twin-jet polisher after mechanical and chemical polishing. The foils were examined with a JEM 200CX operated at 200 kV.

The volume fraction of retained austenite was determined from integrated intensities of the 200 ferrite and 220 austenite X-ray diffraction peaks using Co K_α radiation. The volume fractions of bainitic ferrite and martensite were obtained by means of a point counting method using light and TEM micrographs. However, different procedures were used for specimens quenched from different transformation stages. In the specimens quenched from the initial stages of transformation (723 K, 10 and 25 seconds), first, the volume fraction of bainitic packets was obtained using light micrographs. Second, the volume fraction of bainitic ferrite within a bainitic packet was obtained using TEM micrographs, where the interphase boundaries between ferrite and martensite or retained austenite were close to parallel to the incident beam (edge-on). Then, the volume fraction of bainitic ferrite per unit volume was determined from these two values. The volume fraction of martensite was finally obtained by subtracting the volume fractions of the retained austenite and bainitic ferrite. On the other hand, in the specimens quenched from the later stages of transformation (723 K, 50 seconds and longer), the volume fraction of martensite was first obtained using light microstructures etched with an 8-g $Na_2S_2O_5$ + 100-mL H_2O solution, because it had been confirmed that the martensite forms only from relatively large austenite regions between bainitic packets and can be clearly identified with light micrographs. The volume fraction of bainitic ferrite was then determined by subtracting the volume fractions of the retained austenite and martensite.

The width of bainitic ferrite laths were obtained with TEM micrographs, where broad face boundaries of the laths were close to edge-on.

The carbon concentration of retained austenite was determined from the lattice constant using an equation of $a\gamma = 0.3573 + 0.0033C\gamma$,[12,13,14] where $a\gamma$ is the lattice constant of austenite in nanometers and $C\gamma$ is the carbon content of austenite in mass percent. The lattice constant of retained austenite was obtained from the 111, 200, 220, and 311 austenite peaks using Co K_α radiation.

Table I. Chemical Composition (Mass Pct) and M_s Temperature (K) of the Alloy

C	Si	Mn	P	S	Fe	M_s
0.59	2.01	0.001	0.0006	0.0009	balance	510 K

III. RESULTS

A. *Transformation Curve*

Figure 1 shows the fractions of constituents at room temperature in the specimens reacted at 723 K for various periods of time followed by water quenching. When a specimen was directly quenched into water from 1423 K, the structure became almost fully martensite, and the amount of retained austenite was less than 2 vol pct. The isothermal transformation to bainite started within several seconds and finished by 1 ks. Incomplete transformation or stasis in the bainite reaction was not observed, which is in accordance with the result in an Fe-1.73 pct Si-0.38 pct C alloy (mass pct) reported by Reynolds *et al.*[15] The mount of retained austenite became more than 10 vol pct at the middle stage of transformation (50-second holding). Carbide plates formed in the later stage of transformation after the 50-second holding, as will be shown in Figures 6 and 7. Thus, the product phases during the isothermal holding are labeled as "bainitic ferrite + (carbides)" in Figure 1.

B. *Carbon Concentration of Retained Austenite*

The lattice constant and carbon concentration of retained austenite in the specimens reacted at 723 K for 50 seconds are listed in Table II. Although the values are somewhat different among the specimens, it is evident that the carbon concentration of the retained austenite reaches values higher than twice the average carbon concentration of the alloy. Except for the specimens held

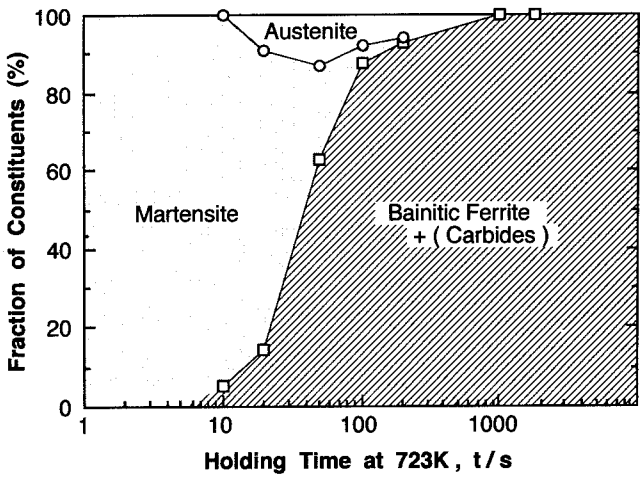


Fig. 1—Variation of fractions of constituents at room temperature with the isothermal holding time at 723 K.

Table II. The Lattice Constant and Carbon Concentration of Retained Austenite in the Specimens Reacted at 723 K for 50 s

Specimen Number	Lattice Constant of γ	Carbon Concentration of γ
1	0.3620 nm	1.42 mass pct
2	0.3625 nm	1.58 mass pct
3	0.3626 nm	1.61 mass pct

for 50 seconds, the lattice constant of retained austenite could not be firmly determined, because clear X-ray diffraction peaks were not obtained due to small amounts of retained austenite.

C. *Light Microscopy*

Figure 2 shows light micrographs of microstructures at (a) middle stage and (b) completion of transformation. In Figure 2(a), it is seen that the morphology of bainitic ferrite is lath rather than plate. Bainitic ferrite laths form separately one from another. Several packets or sheaves consisting of approximately parallel laths are seen in Figure 2(a). Large white regions between bainitic packets correspond to the austenite, which transformed to martensite during quenching after the isothermal holding, as will be shown in Figure 5. Such martensite is not, however, revealed in Figure 2(a) because of etching with nital. At the termination of transformation (Figure 2(b)), individual ferrite laths are hardly identified, although bainitic packets are observed. At 723 K, pearlite did not form.

D. *TEM Microscopy*

Figure 3 shows a TEM microstructure at the early stage of transformation (723 K, 25 seconds). A relatively small bainitic packet surrounded by a martensite matrix is seen in Figure 3(a). Figure 3(b) is a higher magnification micrograph of the region indicated by the arrow in Figure 3(a). The tip part of a ferrite lath is divided by boundaries or gaps as indicated by arrows, whereas such boundaries are not seen in the part away from the tip. Moreover, several small ferrite crystals form near the tip of the ferrite lath. These results suggest that the lengthening of a macroscopic ferrite lath does not occur by a continuous growth of an individual crystal but occurs by a repeated nucleation of small ferrite crystals and the coalescence of them.[7]

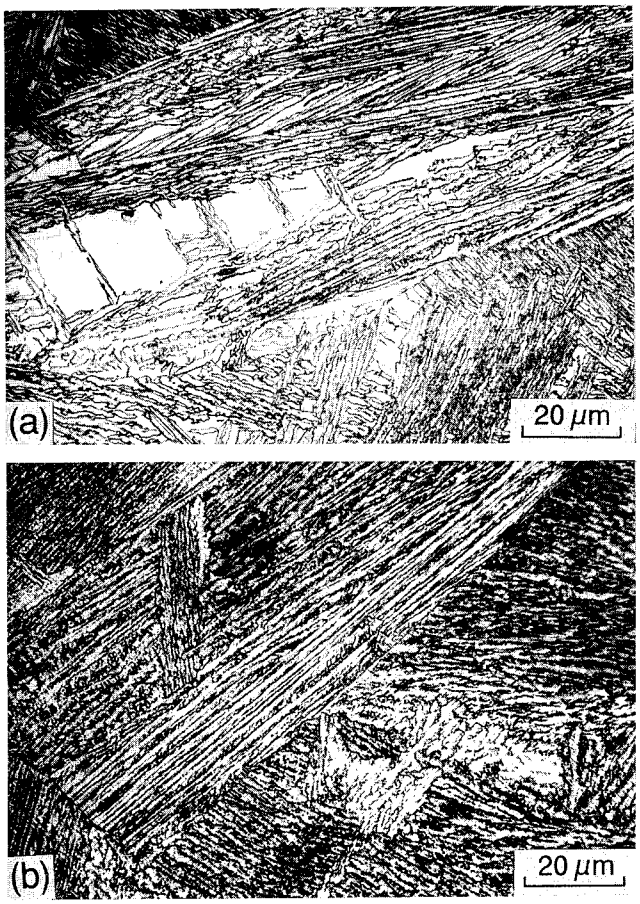


Fig. 2—Microstructures at (*a*) middle stage (723 K, 50 s) and (*b*) completion of transformation (723 K, 1800 s). Light micrographs.

Figure 4 shows TEM microstructures within relatively large bainitic packets at the early stage of transformation (723 K, 25 seconds). In Figure 4(a), a bainitic packet is composed of bainitic ferrite, martensite (M), and retained austenite (A). On the other hand, in Figure 4(b), a packet consists of bainitic ferrite and retained austenite, but martensite is not included. No carbide particles are observed at this stage. It is thus concluded that a bainitic packet consists of bainitic ferrite, martensite and/or retained austenite after quenching from the initial stage of transformation (723 K, 25 seconds).

Figure 5 shows the representative TEM microstructures at the middle stage of transformation (723 K, 50 seconds). Two bainitic packets are seen in Figure 5(a). A large black region marked M between the bainitic packets is the martensite formed during the quenching after the isothermal holding. The laths of bainitic ferrite (BF) are separated by carbon-enriched retained austenite (A), as is clearly seen in Figure 5(b). At this stage, martensite does not exist within a bainitic packet and forms only from relatively large austenite regions between bainitic packets. All of the well-developed bainitic packets are composed of bainitic ferrite and retained austenite. Both the bainitic ferrite and retained austenite involve a large number of dislocations. The parallel ferrite laths in a packet are in the same crystallographic orientation (variant), as will be shown in Figure 7. Carbides are still not observed at this stage.

Figure 6 shows a TEM microstructure in the later stage of transformation (723 K, 100 seconds). Large carbide plates (CB) are seen in the austenite (A) between the initially formed ferrite laths (BF) in a bainitic packet. In some other packets, the austenite between bainitic ferrite laths has decomposed into carbides and ferrite. It is thus clear that in the later stage of transformation, the carbon concentration of untransformed austenite in a packet is decreased by the carbide formation, and subsequently, ferrite forms. This transformation process is the same as that in an Fe-2.4 pct Si-0.79 pct Cr-0.76 pct C alloy (mass pct) reported by Sandvik.[10] According to the nomenclature by Sandvik, the ferrite formed before and after carbide precipitation are hereafter referred to as the primary and secondary ferrite, respectively.

Figure 7 shows TEM microstructures at the termination of transformation (723 K, 1800 seconds). All of the austenite sandwiched between the primary ferrite laths has completely decomposed into carbides and ferrite (Figure 7(a)). The selected-area diffraction pattern (Figure 7(b)) taken from the center region of Figure 7(a) indicates that there are two crystallographic variants of ferrite in the bainitic packet. Figures 7(c) and (d) are the dark-field images of the center part of Figure 7(a) taken

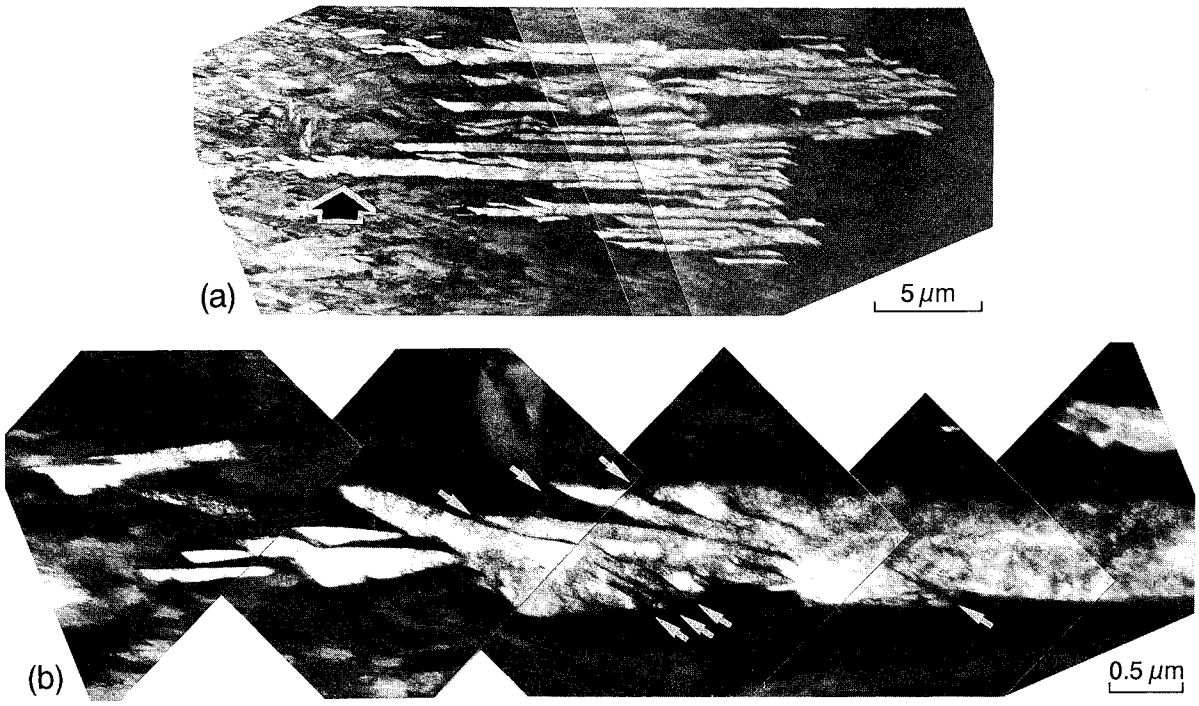


Fig. 3—TEM microstructure of the early stage of transformation (723 K, 25 s): (*a*) shows a relatively small bainitic packet and (*b*) is a high magnification micrograph of the region indicated by the arrow in (a).

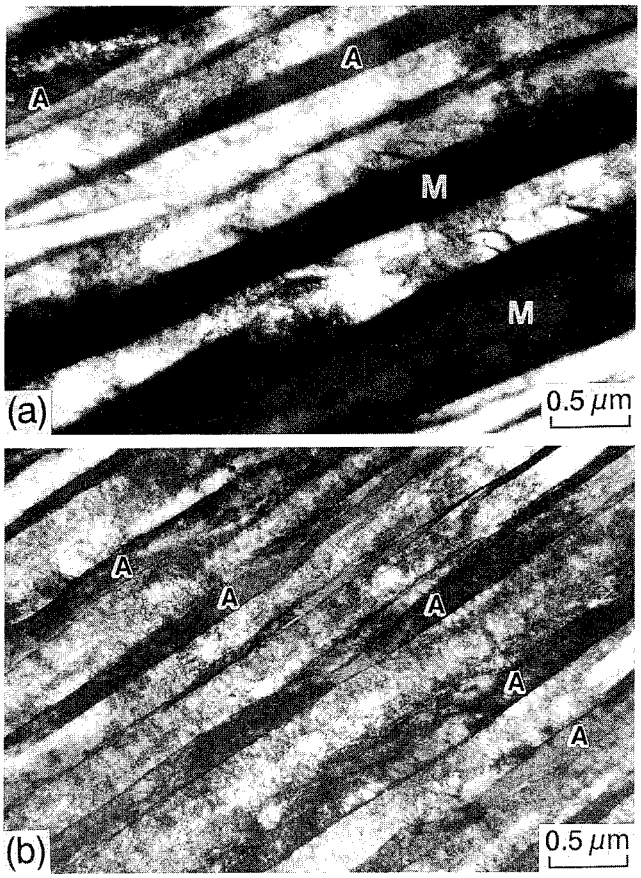


Fig. 4—TEM microstructures of the early stage of transformation (723 K, 25 s). A bainitic packet consists of bainitic ferrite, martensite (M), and retained austenite (A) in (*a*), whereas it is composed of bainitic ferrite and retained austenite in (*b*).

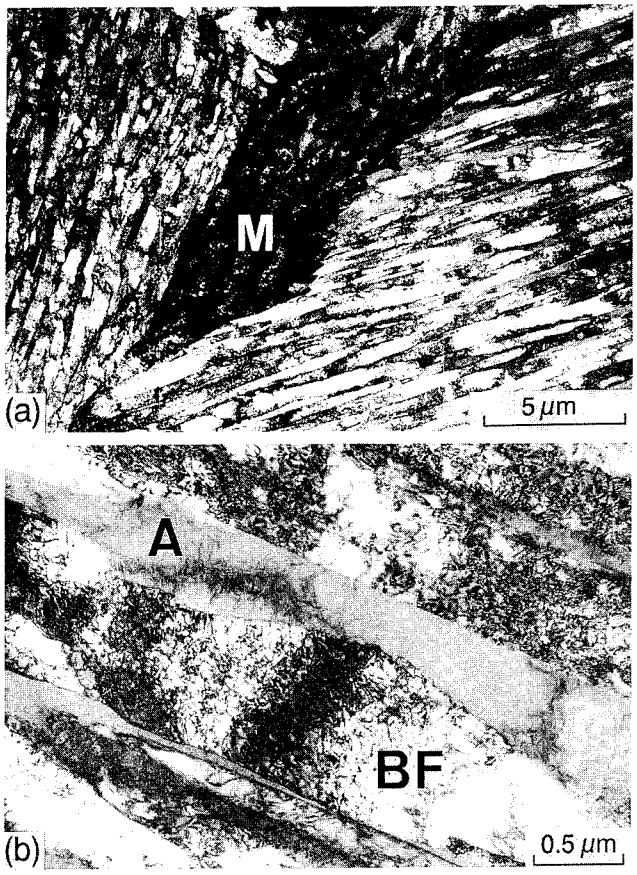


Fig. 5—(*a*) TEM microstructures at the middle stage of transformation (723 K, 50 s); (*b*) is a high magnification micrograph of (a).

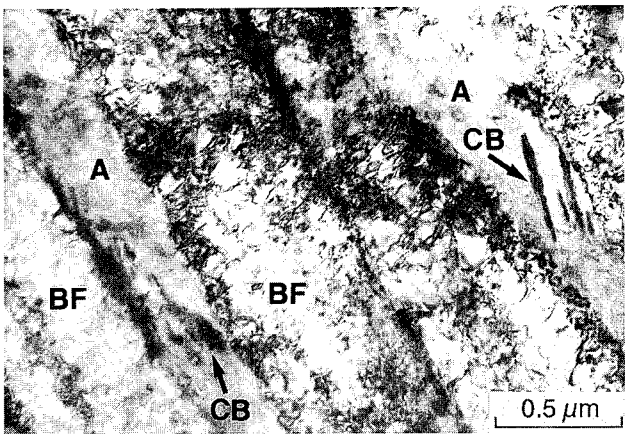


Fig. 6—TEM microstructure at the later stage of transformation (723 K, 100 s) showing large carbide plates (CB).

from a reflection of each variant. It is seen that one variant (BF1) corresponds to the primary ferrite without large carbide plates and the other (BF2) to the secondary ferrite. This indicates that the primary ferrite laths do not thicken during the subsequent isothermal holding, even after the carbide formation in the surrounding austenite and the secondary ferrite with a variant different from the primary ferrite forms for the completion of transformation.

Figure 8 shows high-magnification TEM micrographs of Figure 7(a). Large carbide plates in the secondary ferrite (BF2) are clearly seen in Figure 8(a). Figure 8(b) demonstrates that many fine carbide particles exist in the primary ferrite (BF1). Because no carbides were observed in the primary ferrite at the middle stage of transformation (Figure 5; 723 K 50 seconds), it is clear that the primary ferrite at 50-second isothermal holding was still highly supersaturated with respect to carbon, although the carbon concentration in the surrounding austenite was increased by the carbon rejection from the ferrite (Table II).

E. *Distribution of Ferrite Lath Width*

Figure 9 shows the distribution of the lath widths of the primary ferrite for different isothermal holding times. The primary ferrite could be easily identified even in the final stage of transformation, because it does not contain large carbide plates and has a crystallographic variant different from the secondary ferrite. The distribution hardly changes with the reaction time, and the average value of lath width is a constant of 0.3 μm. This result confirms that the primary ferrite laths grow to a limiting size and do not thicken during the subsequent isothermal holding.

IV. DISCUSSION

Calculated paraequilibrium phase boundaries in Fe-2 pct Si-C alloys are shown in Figure 10,* where α,

*For the calculation, the regular solution model developed by Hillert and Staffanson[16] was used with the thermodynamic parameters reported by Uhrenius[17] and Enomoto.[18] α / (α + θ) and γ / (γ + θ) boundaries were obtained under the condition that cementite does not include silicon. Zener ordering was not included for the T_0 calculations.

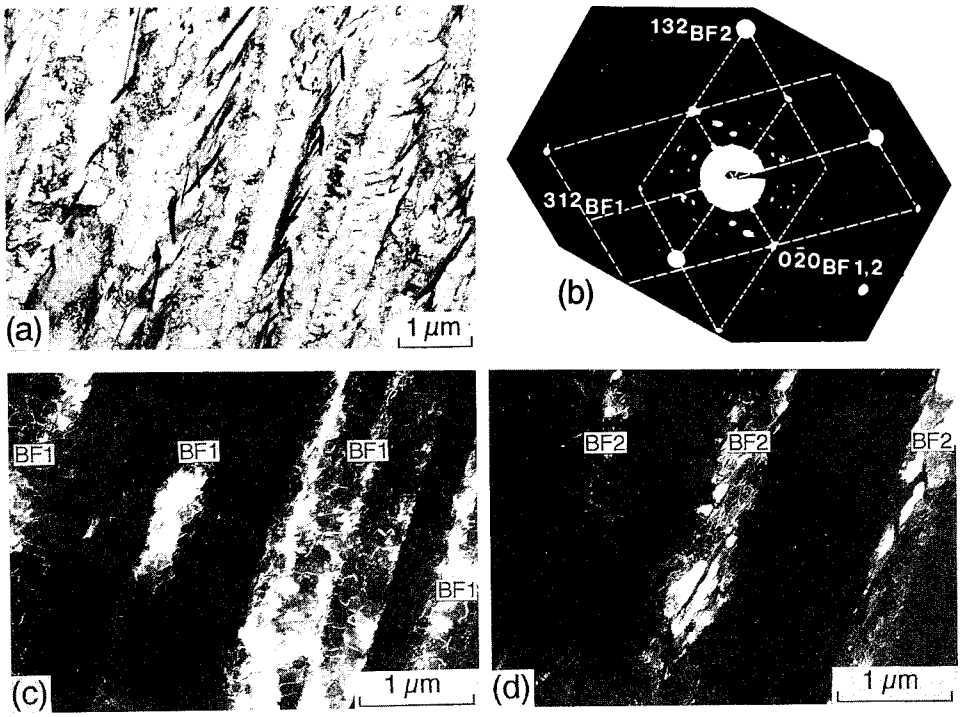


Fig. 7—TEM microstructures at the termination of transformation (723 K, 1800 s): (*a*) bright-field image, (*b*) selected-area diffraction pattern taken from the center part of (*a*), and (*c*) and (*d*) dark-field images taken by reflection from (312)BF1 and (132)BF2, respectively. In (b), the incident beam direction is [$\bar{2}$03] BF1 and [$\bar{2}$01] BF2.

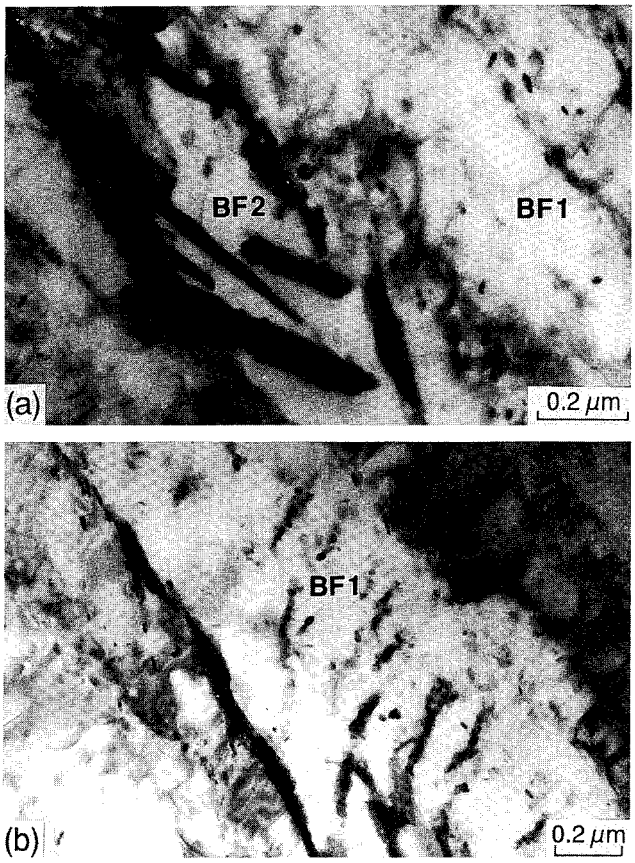


Fig. 8—Higher magnification TEM images of Fig. 7(a) showing (*a*) large carbide plates in the secondary ferrite and (*b*) fine carbide particles in the primary ferrite.

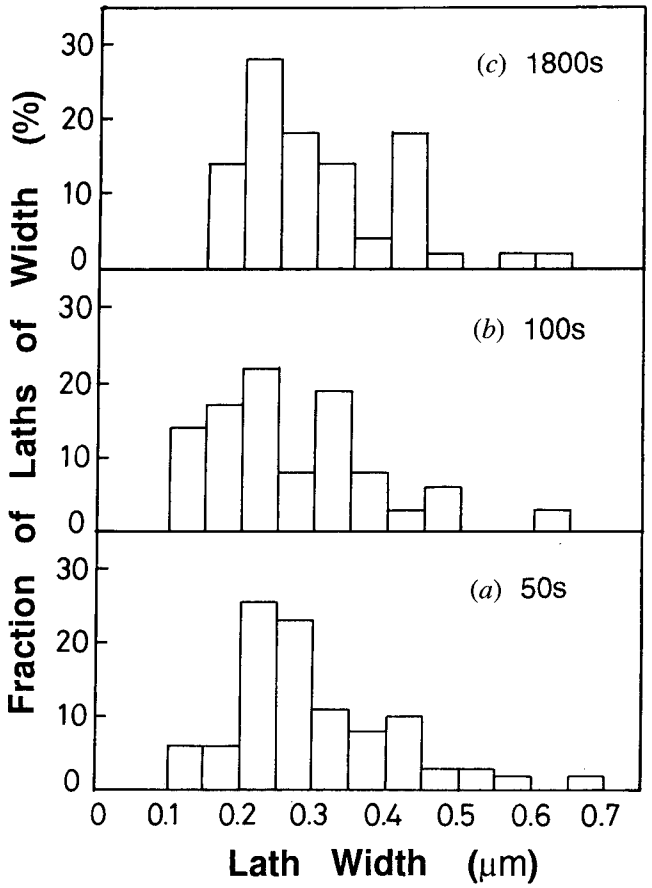


Fig. 9—Distribution of the lath widths of the primary ferrite for different isothermal holding times.

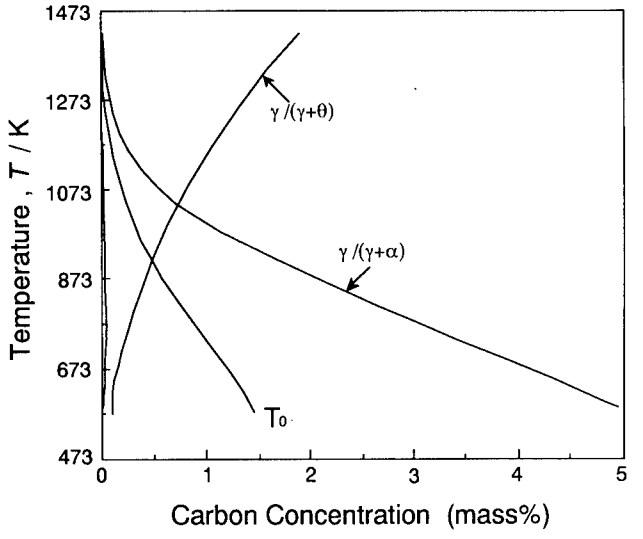
γ, and θ represent ferrite, austenite, and cementite, respectively. At 723 K, the carbon concentration of austenite in equilibrium with ferrite is 3.59 mass pct. In the specimens quenched from the middle stage of transformation (723 K, 50 seconds), the retained austenite existed between the primary ferrite laths within bainitic packets (Figure 5), and the carbon concentration of it was 1.61 mass pct at the highest (Table II). This means that the "average" carbon concentration of the untransformed austenite within bainitic packets at the 50-second holding does not reach the paraequilibrium value. If the bainitic ferrite develops by the diffusional mechanism with a local equilibrium condition, it could grow until the average carbon concentration of the untransformed austenite reaches the paraequilibrium value, because the solute drag-like effect is not expected in the present alloy. However, the primary ferrite did not thicken during the subsequent isothermal holding (Figure 9), and the secondary ferrite with a variant different from the primary ferrite formed after the precipitation of large carbide in the austenite (Figure 7). This result suggests that the bainitic ferrite forms by a displacive mechanism with the glissile movement of coherent interphase boundary, and thus, the growth of it is limited by the loss of coherency of the interface boundary.[6]

Bhadeshia and Waugh[19] directly measured the carbon concentration of a bainitic ferrite in Fe-0.43 pct C-2 pct Si-3 pct Mn alloy (mass pct) by means of field ion microscopy–atom probe (FIM-AP). They reported that the carbon concentration in ferrite was higher than the paraequilibrium value. Similar FIM-AP measurements for supersaturated ferrite have also been reported in other steels.[20,21,22] However, Reynolds *et al.*[22] claimed that the high carbon concentration in ferrite measured by FIM-AP does not simply mean a metastable ferrite. They pointed out that the high carbon concentration was attributed to some finite regions with a high dislocation density (1×10^{12} cm^{-2}) where carbon segregation to dislocations was expected. A similar argument was reported earlier by Bhadeshia and Waugh. Thus, the average carbon concentration in ferrite should be examined to clarify the occurrence of supersaturated ferrite.

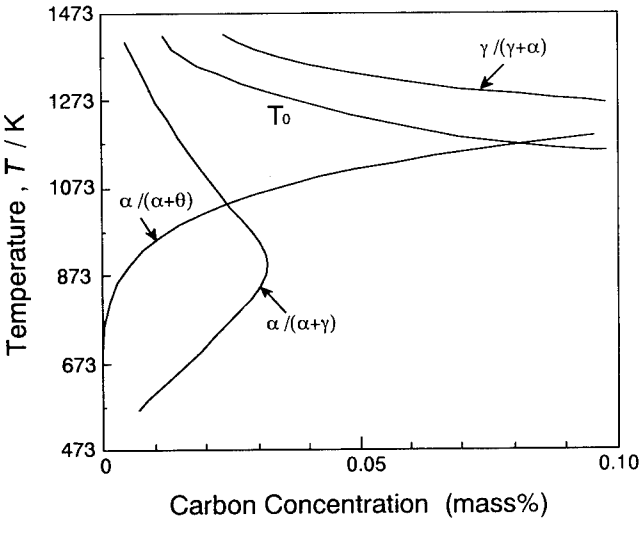
In the specimens quenched from the middle stage of the transformation (Figure 5; 723 K, 50 seconds), carbides were not observed, and the microstructure consisted of primary ferrite, martensite, and retained austenite. In this case, we can estimate the average carbon concentration of the primary ferrite from the mass balance of carbon. Taking the approximation that the bainitic ferrite, martensite, and retained austenite have an identical molar volume, the mass balance is written as

$$(X_F)\,(f_F) + (X_M)\,(f_M) + (X_A)\,(f_A) = \bar{X} \qquad [1]$$

where f_F, f_M, and f_A = volume fractions of the primary ferrite, the martensite, and the retained austenite, respectively, and X_F, X_M, X_A, and $\bar{X}$ = average atom fractions of carbon in the primary ferrite, the martensite, the retained austenite, and the alloy, respectively.



(*a*)



(*b*)

Fig. 10—Calculated paraequilibrium phase boundaries in Fe-2 pct Si-C alloys up to (*a*) 5 mass pct and (*b*) 0.1 mass pct.

As was shown in Figure 1, f_F = 63 vol pct, f_M = 24 vol pct, and f_A = 13 vol pct in the specimen quenched from the middle stage of the transformation.* The carbon concentration of the alloy ($\bar{X}$) is 0.59 mass pct (2.63 at. pct). The carbon concentration of the retained austenite may not be uniform, but the average value of it can be given from the X-ray measurements and adopted to Eq. [1]. The average carbon concentration of the retained austenite (X_A) was 1.61 mass pct (6.94 at. pct) at the highest (Table II). The average carbon concentrations of the primary ferrite and martensite could not be determined from the X-ray measurements because of the overlapping of the primary ferrite and martensite peaks. In the specimens reacted at 723 K for 50 seconds, the martensite formed only from relatively large austenite regions between bainitic packets (Figure 5). As far as the martensite forms, the M_s temperature of such regions must be higher than room temperature (290 K). We thus take the assumption that the average carbon concentration of the martensite corresponds to a value at which M_s temperature becomes room temperature (290 K). Since the carbon concentration of the martensite regions may also be inhomogenious, this assumption apparently overestimates the average carbon concentration of the martensite and thus underestimates the average carbon concentration of the primary ferrite when Eq. [1] is applied. When the specimens were directly quenched from 1423 K, the M_s temperature was obtained to be 510 K (Table I) by means of a dilatometer. According to Andrews' formula of M_s temperature in terms of composition,[23] the addition of 1 mass pct carbon lowers M_s temperature by 423 K. Thus, the average carbon concentration of the martensite (X_M) can be deduced to be 1.11 mass pct (4.87 at. pct) by the present assumption.

Substituting the above values of f_F, f_M, f_A, X_M, X_A, and $\bar{X}$ into Eq. [1], the average carbon concentration of the primary ferrite (X_F) becomes 0.19 mass pct (0.88 at. pct). As is shown in Figure 10(b), the carbon concentration of ferrite in equilibrium with austenite at 723 K is 0.02 mass pct (0.09 at. pct). It is thus apparent that the primary ferrite is highly supersaturated with respect to carbon. This result is in good agreement with the fact that many small carbides exist in the primary ferrite at the termination of transformation (Figure 8; 723 K, 1.8 ks).

The present results for the carbon content and the thickening process of bainitic ferrite strongly indicate that the bainitic ferrite develops by a displacive mode; the bainitic ferrite forms as a metastable ferrite, and the redistribution of carbon atoms mainly occurs after the formation of ferrite. As was shown in Figures 3 and 4, martensite existed within bainitic packets in the specimens quenched from the initial stage of transformation (723 K, 25 seconds), although bainitic packets were composed of ferrite and retained austenite at the middle stage of transformation (Figure 5; 723 K, 50 seconds). This means that the carbon enrichment in the austenite within bainitic packets is not enough at the initial stage of the transformation, because the time after the formation of ferrite is short. The increase in carbon concentration of retained austenite with the bainite reaction time has been already reported in silicon-containing steels by Matas and Hehemann[24] and Le Houillier *et al.*[8]

If the bainitic ferrite develops by a displacive mode, the interfacial structure must be glissile, as that of martensite. Rigsbee and Aaronson[25] examined the interfacial structure of ferrite plates formed by the isothermal holding at 723 to 748 K for 20 to 60 seconds in an Fe-2 pct Si-0.62 pct C alloy (mass pct). The alloy and the heat treatment are similar to those in the present study. They reported that the misfit dislocation structure on the interface was sessile with respect to glide out of the

*As was described in the experimental procedure, the volume fraction of the martensite was obtained by means of a point-counting method using light microstructure etched with an 8-g $Na_2S_2O_5$ + 100-mL H_2O solution. In this case, the volume fraction of the martensite is suspected to be overestimated, because retained austenite might exist between martensite laths within the martensite regions but was counted as martensite. However, it should be remarked that such overestimation results in the underestimation of the volume fraction of bainitic ferrite and thus lowers the average carbon concentration of the primary ferrite obtained from Eq. [1].

interface. This result is contrary to a displacive formation of the ferrite. However, we should consider the transformation stage at which the interfacial structure was observed. The present study clearly indicates that the thickening of the primary ferrite did not occur after the 50 second isothermal holding at 723 K, even though the average carbon concentration of the untransformed austenite within packets was lower than the paraequilbrium value. This suggests that the interface had already lost its mobility as a glissile interface at 723 K with a 50-second holding. Because of high reaction temperatures, the interfacial structure may change by a diffusional mode after stopping the glissile movement in order to release the strain energy caused by a shear, as was suggested by Bhadesia.[26] Such a rearrangement of the interfacial structure will be clarified by the experiment with tempered martensite.

V. CONCLUSION

The bainite transformation during the isothermal holding at 723 K in an Fe-2 pct Si-0.6 pct C alloy (mass pct) has been studied in order to clarify the growth mechanism of the ferritic component of bainite. The main results obtained are as follows.

1. In early stages of transformation (in the reaction time up to 50 seconds), the bainitic ferrite is carbide free. The laths of bainitic ferrite within a packet or sheave are parallel to one another and separated by carbon-enriched retained austenite. Both the bainitic ferrite and retained austenite involve a large number of dislocations.
2. The average carbon concentration of retained austenite in the specimen reacted at 723 K for 50 seconds is 1.61 mass pct at the highest, which is lower than the paraequilibrium value at 723 K (3.59 mass pct). Nevertheless, the bainitic ferrite does not thicken during the subsequent isothermal holding; the distribution of the lath widths of bainitic ferrite hardly changes with the reaction time, and the average value of lath width is a constant of 0.3 μm.
3. The average carbon concentration of bainitic ferrite in the specimens reacted at 723 K for 50 seconds is 0.19 mass pct at the lowest, indicating that the ferrite is highly supersaturated with respect to carbon.
4. In the later stage of transformation (after the 50-second holding), large carbide plates form in the austenite between the laths, resulting in the decrease in the carbon concentration of the austenite. Subsequently, the ferrite with a variant different from the initially formed ferrite in a packet is decomposed for the completion of transformation. Moreover, many fine carbide particles exist in the initially formed ferrite.
5. The results for the carbon concentration and thickening process of bainitic ferrite are contrary to the diffusional ledge mechanism with a local equilibrium condition and strongly indicate that the bainitic ferrite develops by a displacive mode.

ACKNOWLEDGMENTS

The authors would like to express their sincere thanks to Dr. T. Furuhara (Kyoto University), Professor Y. Ohmori (Ehime University), and Professor G. Krauss (Colorado School of Mines) for their stimulating discussion.

REFERENCES

1. H.I. Aaronson: *Inst. Met., Monogr.*, 1969, vol. 33, pp. 270-81.
2. H.K. Hehemann, K.R. Kinsman, and H.I. Aaronson: *Metall. Trans.*, 1972, vol. 3, pp. 1077-94.
3. J.W. Christian and D.V. Edmonds: in *Int. Conf. on Phase Transformation in Ferrous Alloys*, A.R. Marder and J.I. Goldstein, eds., TMS-AIME, Warrendale, PA, 1984, pp. 293-325.
4. H.K.D.H. Bhadeshia and J.W. Christian: *Metall. Trans. A*, 1990, vol. 21A, pp. 767-97.
5. H.I. Aaronson, W.T. Reynolds, Jr., G.J. Shiflet, and G. Spanos: *Metall. Trans. A*, 1990, vol. 21A, pp. 1343-80.
6. R.F. Hehemann and A.R. Troiano: *Trans. AIME*, 1954, vol. 200, pp. 1272-80.
7. J.M. Oblak and R.F. Hehemann: *Transformation and Hardenability in Steels*, Climax Molybdenum Co., Ann Arbor, MI, 1967, pp. 15-30.
8. R. Le Houillier, G. Begin, and A. Dube: *Metall. Trans.*, 1971, vol. 2, pp. 2645-53.
9. H.K.D.H. Bhadeshia and D.V. Edmonds: *Metall. Trans. A*, 1979, vol. 10A, pp. 895-907.
10. B.P.J. Sandvik: *Metall. Trans. A*, 1982, vol. 13A, pp. 777-87; 789-800.
11. K. Tsuzaki, C. Nakao, and T. Maki: *Mater. Trans. JIM*, 1991, vol. 32, pp. 658-66.
12. L. Zwell: quoted by R.C. Ruhl and Morris Cohen: *Trans. TMS-AIME*, 1969, vol. 245, pp. 241-51.
13. L. Cheng, A. Bottger, Th.H. de Keijser, and E.J. Mittemeijer: *Scripta Metall. Mater.*, 1990, vol. 24, pp. 509-14.
14. H.K.D.H. Bhadeshia, S.A. David, J.M. Vitek, and R.W. Reed: *Mater. Sci. Technol.*, 1991, vol. 7, pp. 686-98.
15. W.T. Reynolds, Jr., S.K. Liu, F.Z. Li, S. Hartfield, and H.I. Aaronson: *Metall. Trans. A*, 1990, vol. 21A, pp. 1479-91.
16. M. Hillert and L.I. Staffansson: *Acta Chem. Scand.*, 1970, vol. 24, pp. 3618-26.
17. B. Uhrenius: in *Hardenability Concepts with Applications to Steel*, V.D. Doane and J.S. Kirkaldy, eds., TMS-AIME, Warrendale, PA, 1978, pp. 28-81.
18. M. Enomoto: *Metall. Trans. A*, 1989, vol. 20A, pp. 332-33.
19. H.K.D.H. Bhadeshia and A.R. Waugh: *Acta Metall.*, 1982, vol. 30, pp. 775-84.
20. I. Stark, G.D.W. Smith, and H.K.D.H. Bhadeshia: *Int. Conf. on Solid-Solid Phase Transformations*, Institute of Metals, London, 1988, pp. 211-15.
21. B. Josefsson and H.-O. Adren: *Mater. Sci. Technol.*, 1991, vol. 7, pp. 849-51.
22. W.T. Reynolds, Jr., S.S. Brenner, and H.I. Aaronson: *Scripta Metall.*, 1988, vol. 22, pp. 1343-48.
23. K.W. Andrews: *J. Iron Steel Inst.*, 1965, vol. 203, pp. 721-27.
24. S.J. Matas and R.F. Hehemann: *Trans. TMS-AIME*, 1961, vol. 221, pp. 179-85.
25. J.M. Rigsbee and H.I. Aaronson: *Acta Metall.*, 1979, vol. 27, pp. 365-76.
26. H.K.D.H. Bhadeshia: *Scripta Metall.*, 1983, vol. 17, pp. 1475-79.

High-Temperature Deformation of *B2* NiAl-Base Alloys

IN-GYU LEE, AMIT K. GHOSH, RANJAN RAY, and SUNIL JHA

The high-temperature deformation behavior of three rapidly solidified and processed NiAl-base alloys—NiAl, NiAl containing 2 pct TiB_2, and NiAl containing 4 pct HfC—have been studied and their microstructural and textural changes during deformation characterized. Compression tests were conducted at 1300 and 1447 K at strain rates ranging from 10^{-6} to 10^{-2} s^{-1}. HfC-containing material showed dispersion strengthening as well as some degree of grain refinement over NiAl, while TiB_2 dispersoid-containing material showed grain refinement as well as secondary recrystallization and did not improve high-temperature strength. Hot-pack rolling was also performed to develop thin sheet materials (1.27-mm thick) from these alloys. Without dispersoids, NiAl rolled easily at 1223 K and showed low flow stress and good ductility during the hot-rolling operation. Rolling of dispersoid-containing alloys was difficult due to strain localization and edge-cracking effects, resulting partly from the high flow stress at the higher strain rate during the rolling operation. Sheet rolling initially produced a {111}⟨112⟩ texture, which eventually broke into multiple-texture components with severe deformation.

I. INTRODUCTION

POLYCRYSTALLINE nickel aluminides based on *B2* crystal structure have received much attention due to their low density, microstructural stability, potential for good creep resistance, and oxidation resistance at elevated temperatures. Although this material is generally brittle under ambient conditions, it has been suggested that brittleness may be reduced by grain refinement.[1,2] For this purpose, techniques for producing materials of fine grain size, such as rapid solidification processing and thermomechanical processing, are of considerable interest. One approach toward this end is the utilization of grain-boundary pinning by fine particles.

The room- and elevated-temperature mechanical properties of NiAl have been studied extensively.[2–9] The results indicate that high-temperature creep strength of NiAl is rather low. There have been several attempts in the past to improve high-temperature strength by the development of a second phase through the addition of a third element to the NiAl matrix.[10,11] However, high-temperature stability of such ternary precipitate compounds is always of concern; thus, dispersoid particulates, *e.g.*, TiB_2 and HfC, were considered reasonable candidates for enhancing the high-temperature strength of NiAl. These strengthening refractory phases were preselected because they are generally stable at high temperatures.[12,13] Furthermore, since these particles promote a grain-boundary pinning effect that produces fine grain structure, improved low-temperature toughness was expected.

The poor tensile ductility in polycrystalline *B2* NiAl has been attributed to the lack of a sufficient number of independent slip systems to satisfy von Mises' criterion for compatibility in polycrystalline deformation. At room temperature, only the {110}⟨001⟩ slip system is known to operate, so the number of slip systems is insufficient for flow. Above 673 K, this material shows considerable ductility, because ⟨111⟩, ⟨110⟩, and ⟨100⟩ slip are possible.[14–20] It is possible that sufficient grain refinement may lead to noncube slip in the vicinity of grain boundaries to improve ductility even at room temperature.

The present study was undertaken to characterize the deformation of NiAl-base alloys, to examine the microstructural changes that take place during deformation and the effects of TiB_2 and HfC dispersoids on these changes, and to understand the effects of these changes on mechanical properties at low and elevated temperatures. Hot compression tests were performed to evaluate flow behavior. Based on these results, the production of thin sheet by hot rolling was attempted. Pole figures were determined in order to monitor changes in crystallographic texture during rolling to better understand the operating deformation mechanisms.

II. EXPERIMENTAL PROCEDURES

A. *Materials Fabrication*

Compositions of the NiAl-base alloys studied in the present investigation are listed in Table I. Additions of TiB_2 and HfC were designed for dissolution in the melt and subsequent precipitation as fine dispersion during rapid solidification to impart creep resistance to the alloys. A small amount of Zr was added to improve room-temperature ductility by trapping oxygen during hot-working processes and to enhance high-temperature oxidation resistance.

The alloys were prepared as ingots by arc melting high-purity (>99.9 pct) constituent elements. The prealloyed ingots were subsequently converted into rapidly solidified filaments by an advanced melt-spinning technique developed for processing reactive and refractory alloys. The alloy was arc-melted in vacuum on a water-cooled copper hearth, and the melt was delivered with

IN-GYU LEE, formerly Graduate Student, The University of Michigan, is with the Ceramics Division, Korea Institute of Science & Technology, Cheongryang, Seoul, Korea. AMIT K. GHOSH, Professor, is with the Department of Materials Science and Engineering, The University of Michigan, Ann Arbor, MI 48106. RANJAN RAY, President is with Marko Materials, Inc., North Billerica, MA 01862. SUNIL JHA, formerly with Marko Materials, Inc., is with Texas Instruments, Attleboro, Ma.

Manuscript submitted October 31, 1990.

Table I. Compositions of NiAl-Base Alloys

Alloy Number	Composition (Wt Pct)
1 (NiAl)	Ni-31.5Al-0.1Zr
2 (NiAl + 2 pct TiB_2)	Ni-31.5Al-0.1Zr-2.0TiB_2
3 (NiAl + 4 pct HfC)	Ni-31.5Al-0.1Zr-4.0HfC

a regulated flow to the rim of a rotary-chilled molybdenum casting wheel. As the melt contacted the solid substrate, the cooling rate was very high (usually $\geq 10^6$ K/s), and a rapidly solidified filament formed. These filaments were pulverized by a rotating hammer mill into powders with an average size below 40 mesh; particles finer than 200 mesh were discarded.

The powders were cold-compacted in mild steel cans, degassed at 673 K, and vacuum sealed. The billets were soaked at temperatures ranging between 1422 and 1505 K and extruded as round and rectangular bars. Table II summarizes the extrusion processing conditions for the various alloys studied. NiAl + 4 pct HfC was densified in two stages: closed-die compaction followed by extrusion.

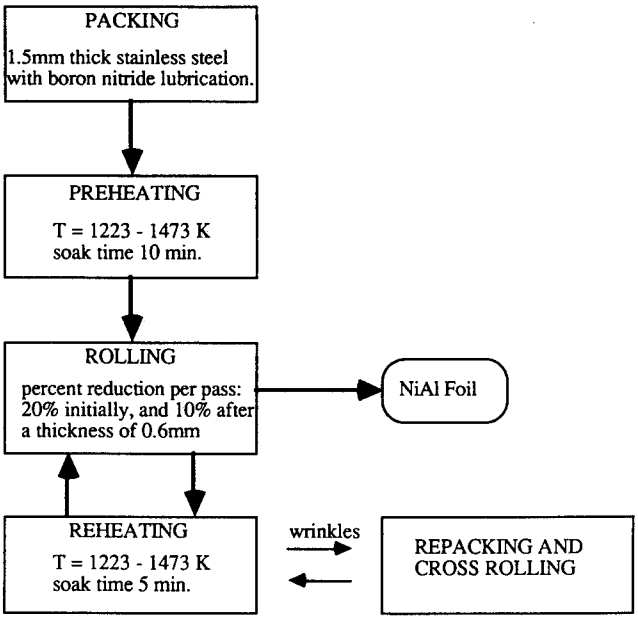


Fig. 1—Flow diagram of the hot-rolling process for NiAl alloys.

B. *High-Temperature Compression Testing*

Compression tests were performed to determine the flow stress and strain-rate sensitivity of NiAl materials. Tests were conducted on NiAl at 1300 and 1477 K. Both NiAl containing 2 pct TiB_2 and NiAl containing 4 pct HfC were tested at 1477 K. Test specimens were 5 mm in diameter and 10 mm in height. A universal testing machine was used, with specimen temperature variation within ±1 K. Through the use of boron nitride lubrication on alumina compression platens, the extent of barreling was found to be relatively small even after a true strain level of 0.3. Strain rates were maintained constant during testing by continuously decreasing crosshead velocity in an exponential manner as specimen height decreased. True stress–true strain data were prepared from these tests, and microstructural characterization of deformed specimens was carried out.

C. *Hot Rolling of NiAl Materials*

Hot rolling of the NiAl alloys at a rapid deformation rate was performed to initiate dynamic recrystallization. To achieve a fine, recrystallized grain size in this material, the lowest possible temperature that would not cause cracking was used. Rolling was attempted at 1223 to 1473 K with these alloys. Initial rolling passes were approximately 1.7 mm in reduction. The pieces were packed in 1.5-mm-thick stainless steel with a copious amount of boron nitride lubrication. The packs were preheated to rolling temperature and rolled with reductions of 0.6 to 0.8 mm per pass at the slowest possible rates (1270 surface mm/min). When the sheet thickness dropped below 0.6 mm, the reduction per pass was reduced to 0.05 to 0.08 mm to avoid cracking. Wrinkling was observed in the sheets as they were pack-rolled in the uniaxial direction. Continued uniaxial rolling caused cracking in the foils. Therefore, foils were repacked in stainless steel after a few passes and cross-rolled just enough to remove the wrinkles. Rolling passes were then continued at the slow deformation rate. After each pass, intermediate reheating of the foils was performed at the rolling temperature for 5 minutes. This hot-rolling process is schematically illustrated in Figure 1.

D. *Pole Figure Measurement*

To understand the mechanisms of deformation and changes in grain structure and orientation during the hot rolling of NiAl, pole figures were determined at several stages of rolling to indicate the crystallographic texture changes. Three different samples of NiAl were prepared for pole figure measurement: as-extruded material, sheet after hot rolling to 1.27 mm (true strain = 1.9), and

Table II. Summary of Extrusion Processing

Alloy Number	Shape	Billet Diameter (mm)	Die Dimensions (mm)	Extrusion Temperature (K)	Extrusion Ratio	Result
1	bar	51.8	12.7	1422	16:1	—
1	plate	64.6	28.7 × 9.5	1494	12:1	—
2	bar	51.8	12.7	1422	16:1	—
3	—	64.6	blind die	1505	—	hot upset
(Hot upset)	bar	51.8	15.9	1505	10.7:1	—

sheet after hot rolling to 0.25 mm (true strain = 3.6). A Rigaku diffractometer was used to measure the preferred orientation in samples by rotating the sample around two axes. Only the inner portion of the pole figure was measured, using the Schultz reflection method. A normalization was carried out on the entire pole figure after the data were collected.

III. RESULTS AND DISCUSSION

A. *Initial Microstructures*

The microstructure for extruded NiAl, shown in Figure 2(a), exhibits a mixture of fine and coarse grains. Grains are somewhat elongated along the extrusion direction; however, an average linear intercept from various directions gives a grain-size value of 8 μm. Metallographic examination of the microstructure generally revealed grain refinement with additions for second-phase particles. TiB_2-containing material showed very small grains intermixed with a few coarse, elongated grains (Figure 2(b)), a likely result of secondary recrystallization. The fine grain size was ~2 to 3 μm, while the coarse grains exceeded 100 μm. The coarse grains are seen to contain subgrains. It is probable that a nonuniform particle distribution resulted in the duplex microstructure. The microstructure of NiAl + 4 pct HfC (Figure 2(c)) was also fine (although somewhat coarser than that of NiAl + 2 pct TiB_2); the average grain size was 4 μm and grain shape was similar to NiAl.

B. *Flow Stress Behavior and Corresponding Microstructural Changes*

The hot compression test results are plotted in Figures 3 and 4 for NiAl at 1300 and 1477 K, respectively, and in Figures 5 and 6 for NiAl + 2 pct TiB_2 and NiAl + 4 pct HfC, respectively, at 1477 K.

1. *NiAl*

This alloy reaches a saturated steady-state flow stress after approximately 10 pct strain and does not exhibit any further work hardening (Figures 3 and 4). This behavior is seen at both 1300 and 1477 K. Thus, at least in the strain-rate range of 10^{-2} to 10^{-4} s^{-1}, little change in grain structure other than in grain aspect ratio is expected; changes in the subgrain structure also occur during this stage. Assuming no other microstructural change occurs during deformation, the change in grain aspect ratio due to the imposed compressive strain can be described by a simple equation:

$$A_f = A_i \exp(1.5\ \varepsilon_c) \qquad [1]$$

where A_f is the grain aspect ratio after deformation, A_i is the initial grain aspect ratio of the specimen, and ε_c is the compressive strain.

At 1477 K and at strain rates below 10^{-5} s^{-1}, some strain hardening was observed, which suggests that concurrent dynamic grain growth may have occurred.[21] Grain size was measured after 1477 K compression tests and is plotted in Figure 7. Figure 7 shows that a significant amount of dynamic grain growth occurred (an increase from 10 μm to as large as 26 μm) and that this growth was much greater initially in the transverse direction. This indicates an instability of the as-extruded microstructure due to the imposed strain.

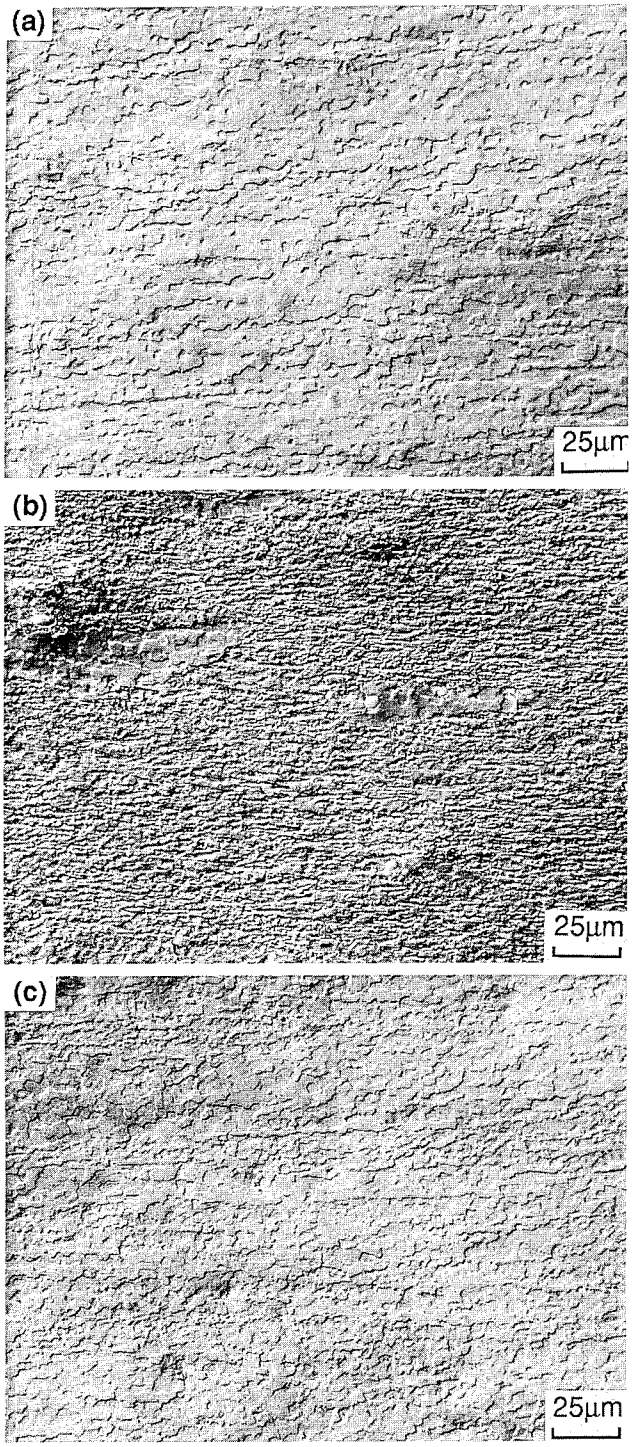


Fig. 2—Optical micrographs of extruded materials (horizontal extrusion direction): (*a*) NiAl; (*b*) NiAl + 2 pct TiB_2; and (*c*) NiAl + 4 pct HfC.

2. *NiAl + 2 pct TiB_2*

For this alloy, strain hardening was observed at all of the strain rates studied (Figure 5). At higher strain rates, the strain-hardening behavior is believed to arise from interactions between second-phase particles and dislocations. The fact that interactions do occur is evidenced

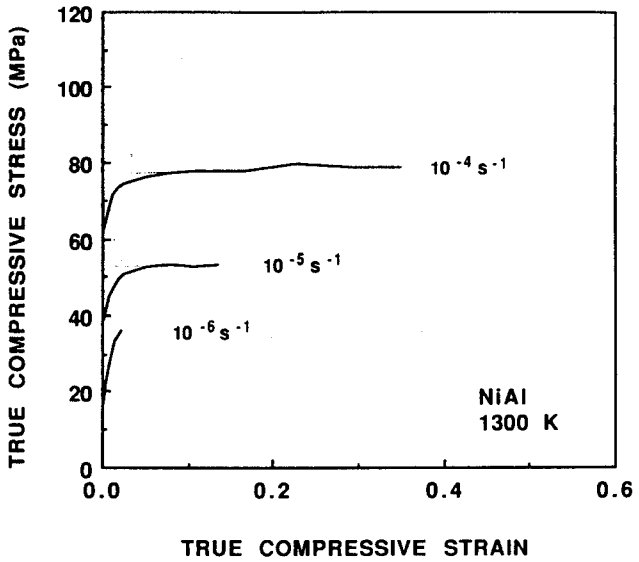


Fig. 3—Stress-strain curves for NiAl in compression for three different strain rates at 1300 K.

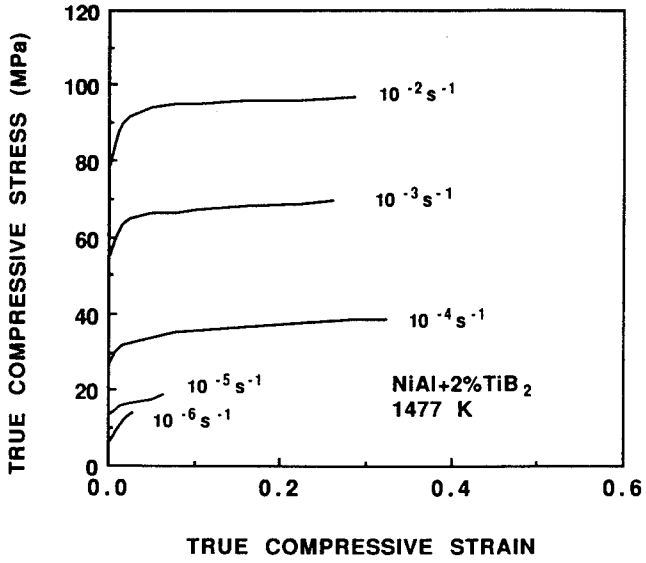


Fig. 5—Stress-strain curves for NiAl + 2 pct TiB_2 in compression for five different strain rates at 1477 K.

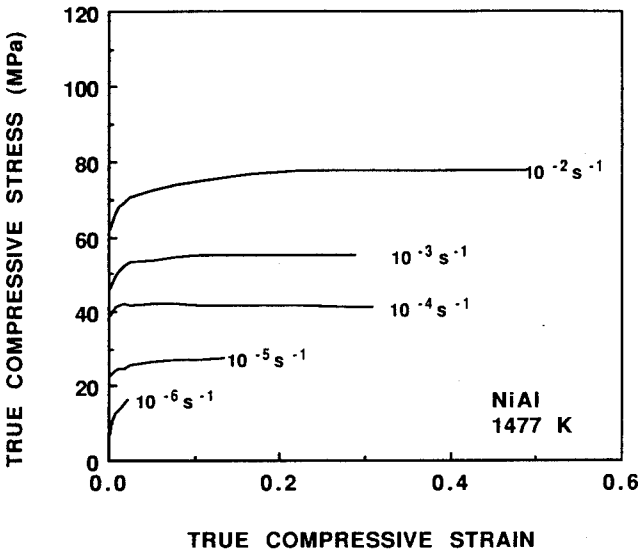


Fig. 4—Stress-strain curves for NiAl in compression for five different strain rates ar 1477 K.

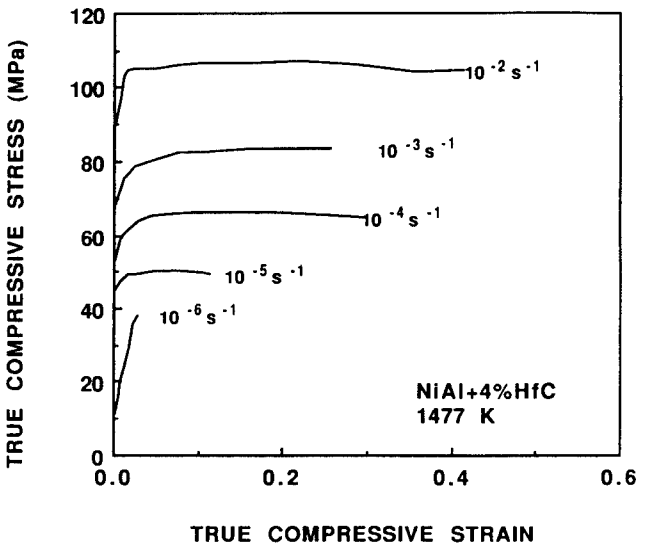


Fig. 6—Stress-strain curves for NiAl + 4 pct HfC in compression for five different strain rates at 1477 K.

by the development of a particle-pinned fine subgrain structure in this alloy. In low-strain-rate tests, in addition to this phenomenon, dynamic grain growth was observed. Since some degree of diffusional creep accommodation is believed to be occurring at this temperature, strain hardening can result from grain coarsening as well as coarsening of the substructure with continued deformation.[21,22]

Figure 8(a) shows a photomicrograph of a specimen static-annealed for 1 hour. A mixture of coarse, elongated grains and finer grains can be seen. The fine grain area also had a duplex grain size, *i.e.*, a mixture of grains with sizes ranging from ~2 to 3 μm to ~10 to 15 μm. (Recent studies using transmission electron microscopy have indicated that the ~2- to 3-μm grain-like features are subgrains.) The ~2- to 3-μm subgrains are believed to grow into coarse, elongated grains with continued annealing. This leads to an increase in the volume fraction of coarse grains.

Figures 8(b) and (c) show microstructures of NiAl + TiB_2 samples after compression tests were performed at 1477 K at two different strain rates. After compression testing, the microstructure of this alloy exhibits a duplex grain structure. Apparently, secondary grain growth occurred in the course of testing. With decreasing strain rate, the volume fraction and size of large grains were increased. The average grain size of the fine grains became 4 μm and the larger grains grew up to 2000 μm. Although grain refinement was achieved by adding TiB_2 particles in the fabrication process, these particles were not able to prevent grain-boundary migration during high-temperature tests. The high-temperature strength of

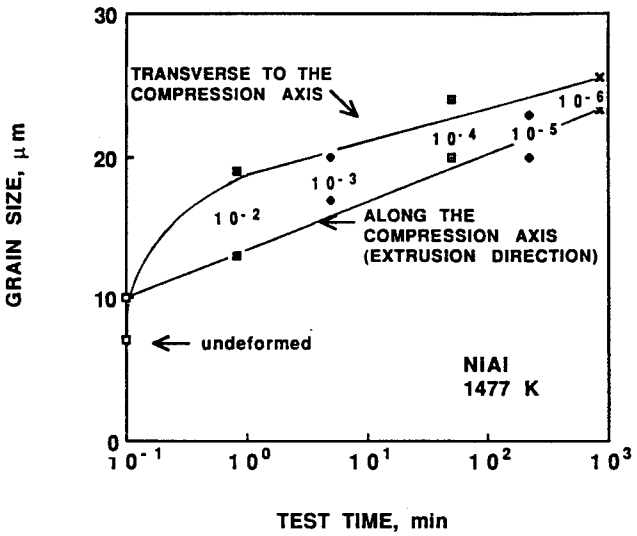


Fig. 7—Grain size as a function of compression test time for NiAl at 1477 K. Test strain rates (s^{-1}) are indicated.

this material was not improved, possibly because of the finer grain size.

3. *NiAl + 4 pct HfC*

Figure 6 shows that NiAl + 4 pct HfC reached a steady value of flow stress at 1477 K at all strain rates except the slowest (10^{-6} s^{-1}). Little hardening was observed in this alloy, which suggests that little change in the microstructure occurred during testing. In the test conducted at 10^{-6} s^{-1}, continued strain hardening was observed, which suggests that dynamic grain growth may have occurred at this slow rate.

The grain structure of NiAl + 4 pct HfC after hot compression testing is shown in Figures 9(a) and (b). The microstructure of the sample from the higher strain rate (as in Figure 9(a)) shows that grains were fairly elongated along the test direction. Coarse HfC particles are also visible in the micrograph, and these particles are not located on high-angle grain boundaries. At slow strain rate, grains are more equiaxed and fewer fine particles are visible on the grain boundary as compared with the specimen deformed at higher strain rates.

For NiAl + 4 pct HfC, the grain size and corresponding aspect ratio were measured after compression testing; results are plotted in Figure 10. As test time increased, the grain structure became more equiaxed; grain size along the compression axis decreased, and grain size transverse to the compression axis increased. This kind of spheroidization effect is clearly a result of stress-assisted grain-boundary migration (which is possibly controlled by grain-boundary diffusion) and counteracts flattering of grains due to strain.

Grain-growth behavior and grain-shape change during high-temperature compression testing for these alloys are schematically described in Figure 11. For NiAl + 2 pct TiB_2, the dispersoid particles helped to produce a fine-grain material by retarding grain growth during the extrusion process. After compression testing, grains in NiAl + 4 pct HfC became more elongated and coarser than in the other two alloys.

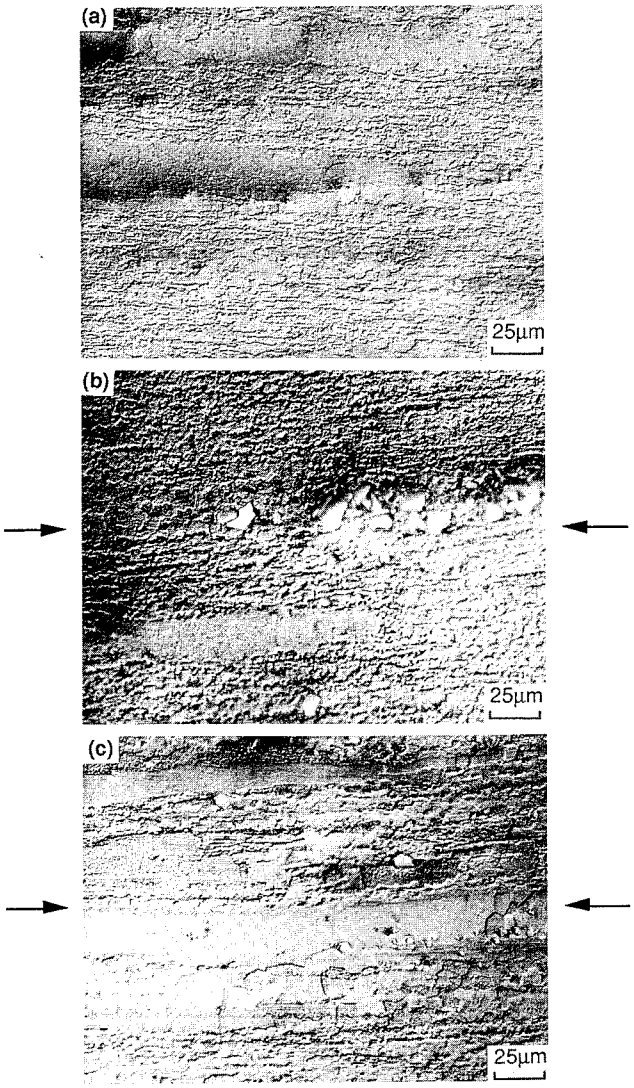


Fig. 8—Optical micrographs of NiAl + 2 pct TiB_2 (arrows indicate compression directions): (*a*) annealed for 1 hr at 1477 K; (*b*) compression tested at 1477 K, strain rate = 10^{-2} s^{-1}, $\varepsilon = 0.5$; (*c*) compression tested at 1477 K, strain rate = 10^{-5} s^{-1}, $\varepsilon = 0.15$.

The stress–strain rate data in Figures 3 to 6 were derived from the standard power law

$$\dot{\varepsilon} = A\sigma^n \qquad [2]$$

where A is a constant and n is the stress exponent. NiAl data also resulted from using the temperature-compensated power law rate expression

$$\dot{\varepsilon} = B\sigma^n \exp\left(-\frac{Q}{RT}\right) \qquad [3]$$

where B is a constant, Q is the activation for creep, R is the gas constant, and T is absolute temperature. Linear-regression techniques were used for the fitting, and the results of these calculations for A, B, n, and Q, as well as the coefficients of determination R_d^2, are given in Table III. For NiAl the stress exponent, n, was 6.23 and activation energy, Q, was 356 KJ mol^{-1}. This is in agreement with Whittenberger's data.[8] The stress exponent was 5.75 and the activation energy was 314 KJ

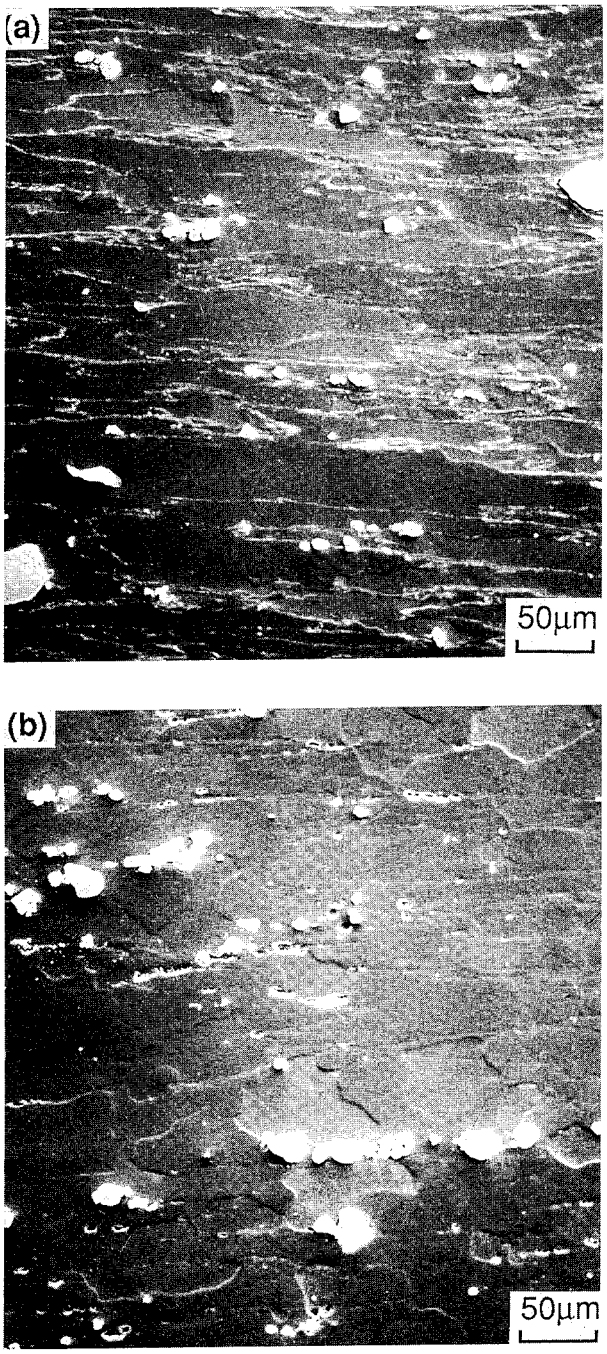


Fig. 9—Scanning electron micrographs of NiAl + 4 pct HfC after compression testing at 1477 K (horizontal compression axis): (*a*) strain rate = 10^{-2} s^{-1}, ε = 0.5; (*b*) strain rate = 10^{-6} s^{-1}, ε = 0.05.

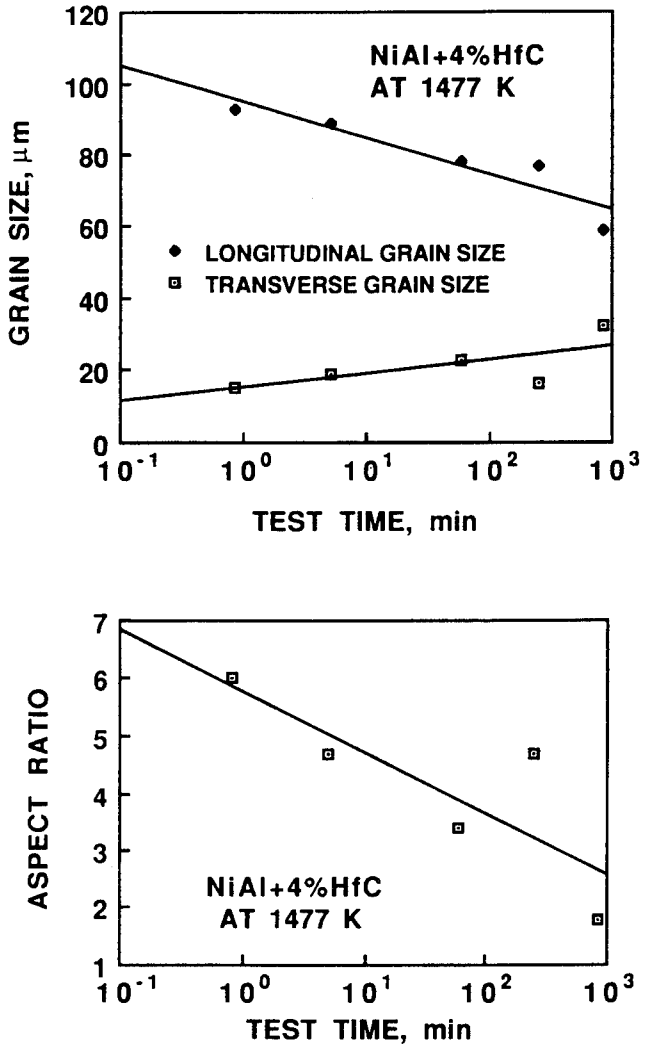


Fig. 10—Grain size and corresponding aspect ratio as a function of compression test time for NiAl + 4 pct HfC at 1477 K.

mol^{-1} for powder-processed NiAl with a grain size of 15 μm. The stress exponents for NiAl + 2 pct TiB_2 and NiAl + 4 pct HfC tested at 1477 K were 3.9 and 8.6, respectively. Whittenberger *et al.*[23] tested dispersoid-containing NiAl alloys produced by same processing technique at 1300 K and reported n values of 5.3 for NiAl + 2 pct TiB_2 and 12.0 for NiAl + 4 pct HfC.

Figure 12 shows stress *vs* strain-rate plots on a log-log scale for NiAl at the two temperatures examined and for NiAl + 2 pct TiB_2 and NiAl + 4 pct HfC at 1477 K. The data points in this plot are fitted with the best least-square polynomial curves, and the slopes of these curves represent strain-rate sensitivity (m) values, which are plotted in Figure 13. If a steady-state situation prevails, then m is equivalent to the reciprocal of n in Eq. [2]. At 1477 K, the NiAl + 4 pct HfC alloy exhibited the highest flow stress and NiAl + 2 pct TiB_2 showed the highest strain-rate dependency of steady flow stress among the alloys examined. The low m value of 0.2 indicated that deformation occurs primarily by dislocation creep.

The TiB_2-containing alloy has somewhat higher m values (0.3) compared to the other alloys. The significantly higher m value for NiAl + 2 pct TiB_2 suggests possibly a finer structure and a greater degree of plasticity, which is helpful from a processing standpoint. Fine TiB_2 dispersion, leading to finer grain size in this material, seems to be responsible for this behavior. For this alloy, grain-boundary diffusional accommodation as well as dislocation climb are believed to be occurring even at higher strain rates. Since diffusional accommodation by grain-boundary Coble creep is strongly dependent on grain size (inverse d^3 relationship), strain hardening is

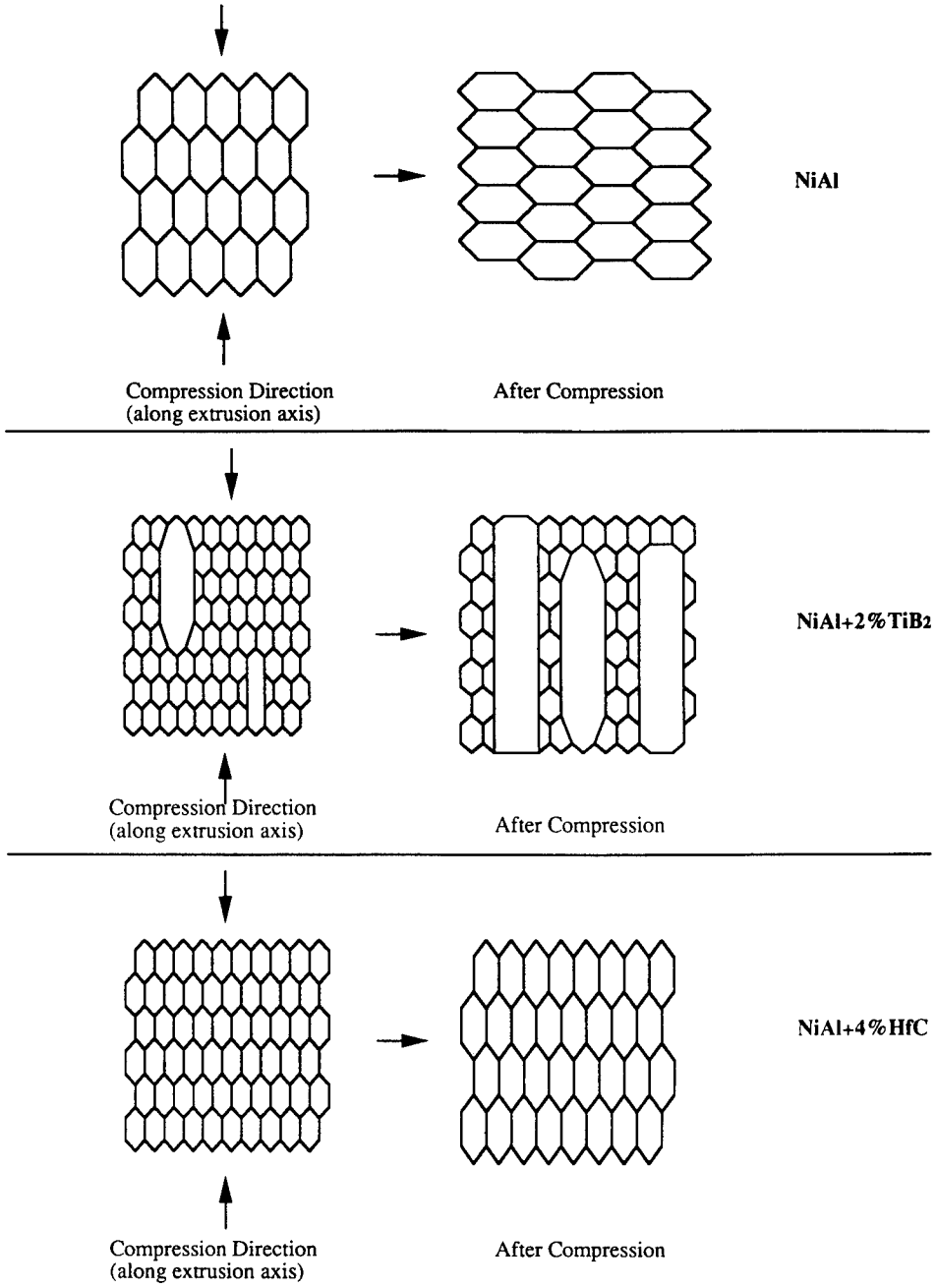


Fig. 11—Schematic diagram of changes in grain shape as a result of compression testing for three different NiAl alloys.

Table III. Summary of Regression Fits for NiAl-Base Alloys

Material	Temperature (K)	A (s^{-1})	B (s^{-1})	n	Q (KJ mol^{-1})	R_d^2
Power Law						
NiAl	1300	2.53×10^{-16}	—	6.16	—	1.000
	1477	1.31×10^{-14}	—	6.30	—	0.985
NiAl + 2 pct TiB_2	1477	1.28×10^{-10}	—	3.90	—	0.987
NiAl + 4 pct HfC	1477	3.67×10^{-20}	—	8.64	—	0.997
Temperature-Compensated Power Law						
NiAl	1300 to 1477		5.18×10^{-2}	6.23	356	1.000

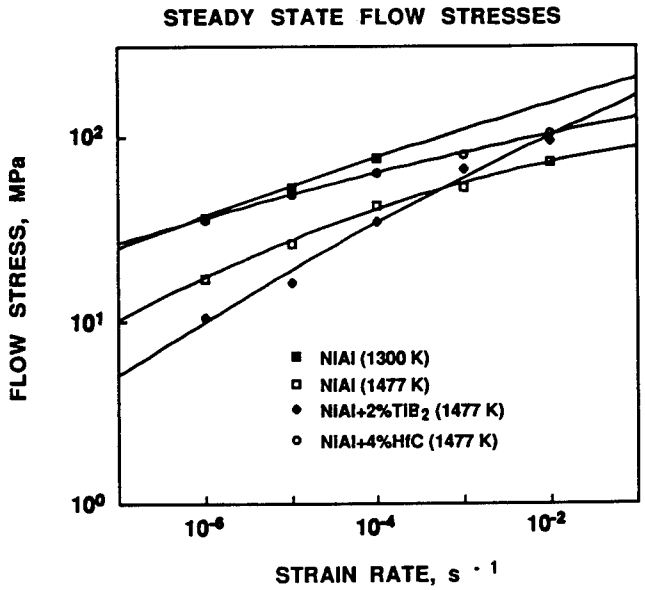


Fig. 12—Flow stress–strain rate data for three different NiAl materials. NiAl specimens were tested at 1300 and 1477 K.

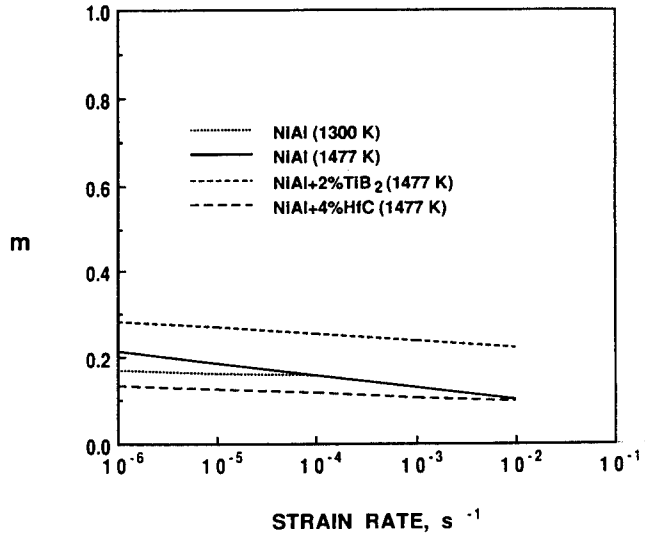


Fig. 13—The corresponding *m* values derived from the slope of curves in Figure 12.

expected with dynamic grain growth; however, the extent of this hardening is somewhat reduced by concurrent dislocation creep.[21,22] This type of flow-stress behavior is shown in Figure 5.

C. *Hot Rolling*

1. *Hot rolling of NiAl*

At 1323 K, this alloy was rolled to produce thin sheets down to 0.2 mm in thickness. Slight cracking occurred during the operation; therefore, conditions for suppressing edge cracks were investigated. At 1373 K, no cracks were visible and crack-free foils down to 0.13 mm were produced.

High-speed rolling of this alloy was also attempted to initiate dynamic recrystallization and possible grain refinement, but this produced considerable cracking of the pieces. Grain-boundary cracking was observed in this material, and it is believed that dynamic recrystallization is not a suitable means of grain refinement for these alloys since the grain boundaries cannot withstand the high deformation stress, particularly at the highest deformation rates. An optical micrograph of rolled NiAl sheets (Figure 14) shows grain growth and pancake grain formation after rolling, indicating that grain refinement of this material by the hot-rolling process cannot be achieved without the presence of grain refiners.

2. *Hot rolling of NiAl + 2 pct TiB_2 and NiAl + 4 pct HfC*

These particle-containing alloys were rolled to a sheet thickness of 0.8 mm at 1373 K. Severe cracking occurred when production of a thinner sheet was attempted. Further reduction in thickness did not appear promising due to the high flow stress and cracking of these alloys. During rolling of NiAl + 2 pct TiB_2 at 1473 K, cracks formed along the extrusion axis in a location where the steel casing had separated.

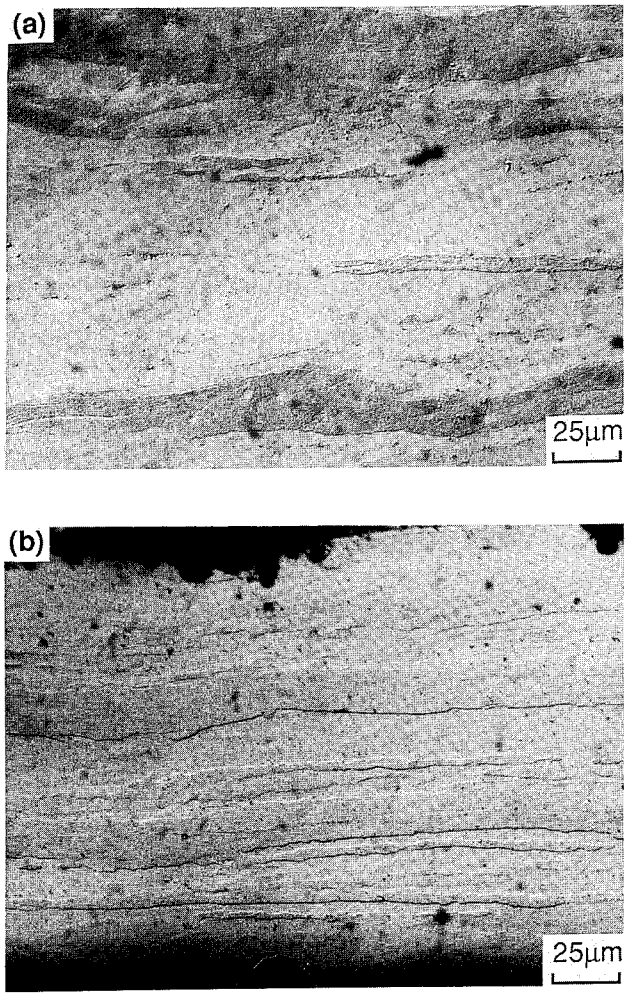


Fig. 14—Optical micrographs of hot-rolled NiAl sheets (horizontal rolling direction) Sheet thickness of (*a*) 1.25 mm and (*b*) 0.25 mm.

D. *Texture Change*

The pole figures of the as-extruded NiAl specimen given in Figure 15 show a {110}⟨110⟩ texture. The shape of the as-extruded NiAl was rectangular, so it was not a true fiber texture. Similar ⟨110⟩ fiber texture was reported by Vedula *et al.*[24] for an extruded NiAl alloy. The {110}⟨110⟩ texture of the as-extruded alloy changed to {111}⟨112⟩ texture after the sample was rolled down to 1.27 mm (Figure 16). Khadkikar *et al.*[25] also reported ⟨111⟩ fiber texture for cast and extruded and powder-extruded NiAl materials and powder-extruded FeAl. For the cast and extruded FeAl, ⟨110⟩ fiber texture was observed; however, texture was not examined in the cast state. A {111}⟨112⟩ texture was found from the pole figures of the 0.25-mm sample in Figure 17, but this texture was weaker than that of the 1.27-mm sample. Weak {100}⟨110⟩ texture was also found from the pole figure of the same sample. This result may be due to the significant degree of cross rolling that was required to produce this foil. A TiB_2-containing NiAl alloy was forged at 1573 K, and the same ⟨111⟩ fiber texture was observed. In contrast, HfC-containing material forged at 1623 K showed a ⟨110⟩ and weak ⟨100⟩ fiber texture. At high temperature it was reported that {100}⟨100⟩ and {110}⟨100⟩ slip systems operated in NiAl single crystals.[15,19,20] In compression, the axis rotates toward the slip plane normal, ending between three ⟨100⟩ poles, *i.e.*, ⟨111⟩. Sometimes ⟨110⟩ is possible, ending between two ⟨100⟩ poles. This may be the case for the heavily rolled sheet and HfC-containing NiAl. In tension, the alternate operation of primary and conjugate slip systems leads to a final orientation of ⟨112⟩. Rolling corresponds roughly to a superposition of the two processes: tension in the rolling direction and compression at right angles to the sheet. Correspondingly, a {111}⟨112⟩ rolling texture is expected for NiAl, *i.e.*, a [111] orientation of the normal to the sheet plane and a [112] orientation in the rolling direction.

IV. CONCLUSIONS

1. The high-temperature strength of NiAl can be improved by adding fine HfC and TiB_2 particles, with HfC providing the largest increase. This strength increase is more than 50 pct the strength of NiAl at all strain rates.

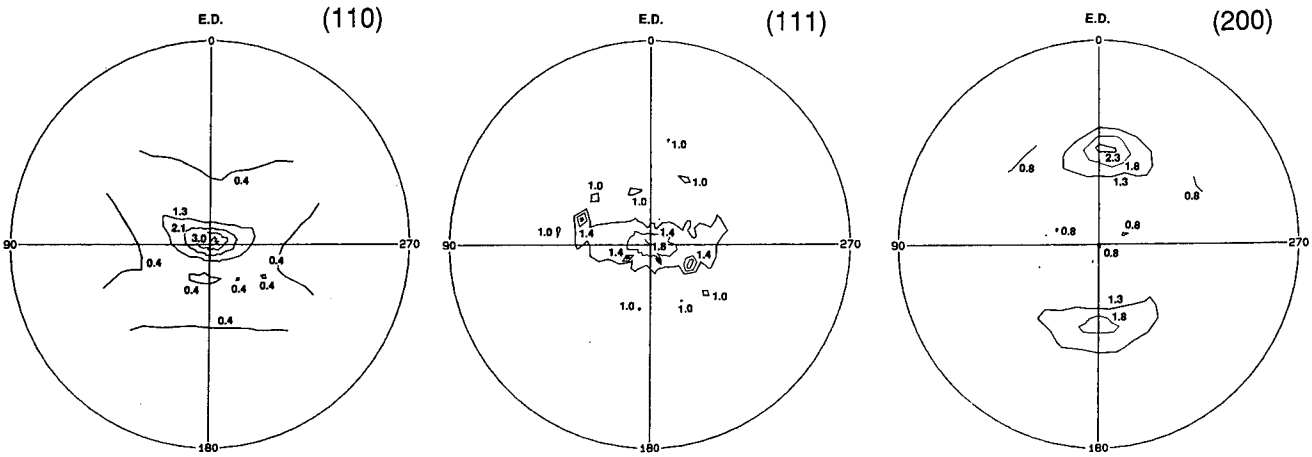


Fig. 15—Normalized pole figures of NiAl in the extruded condition for (110), (111), and (200) pole reflections.

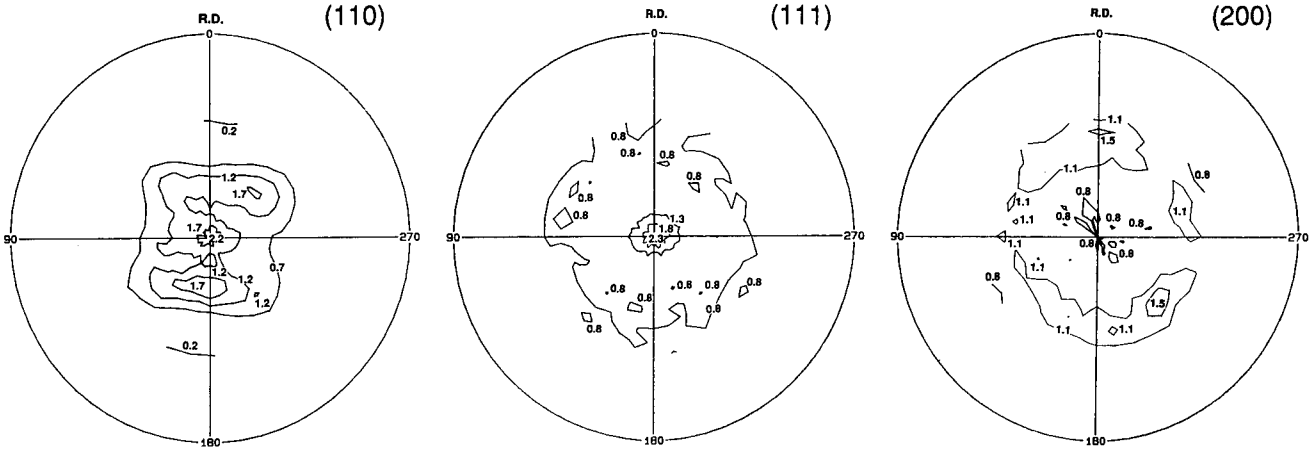


Fig. 16—Normalized pole figures of NiAl in the hot-rolled condition for (110), (111), and (200) pole reflections. Hot-rolled to a thickness of 1.27 mm at 1323 K.

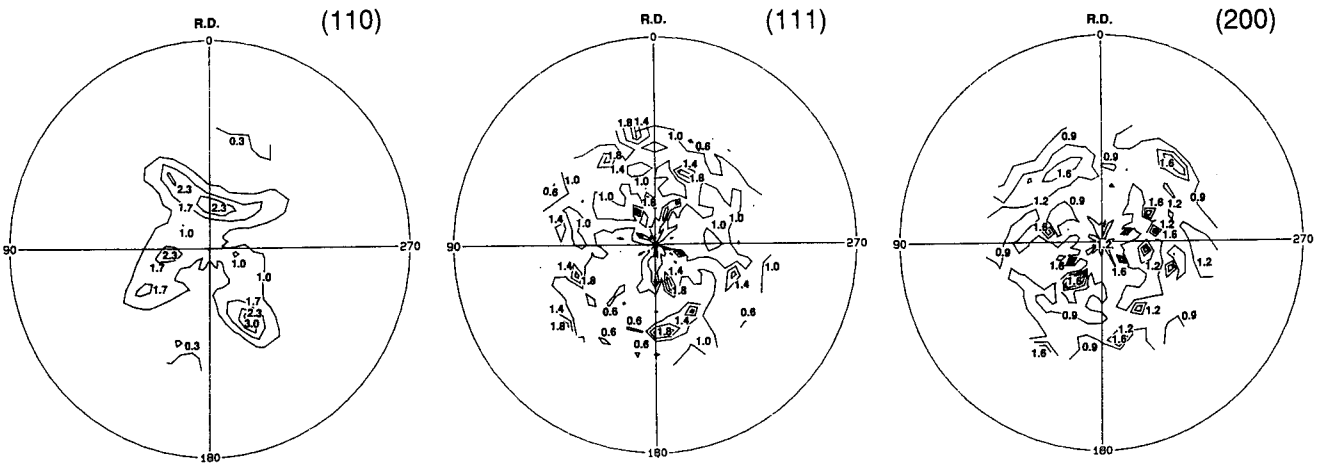


Fig. 17—Normalized pole figures of NiAl in the hot-rolled condition for (110), (111), and (200) pole reflections. Hot-rolled to a thickness of 0.25 mm at 1323 K.

2. TiB_2 dispersoids are not as promising for improving high-temperature strength at low strain rates due to fine grain size and poor particle-pinning effects.
3. Microstructural change occurred during compression testing, typically resulting in grain growth and changes in grain shape, which depends on second-phase additions.
4. Conventional hot-working processes, such as extrusion, forging, and rolling, can be applied to NiAl-base alloys, but rolling of the particle-containing materials is extremely difficult.
5. The strain-rate sensitivity calculated from compression tests is about 0.2, indicating that deformation of this alloy is primarily by dislocation creep. TiB_2-containing material shows that diffusional creep accommodation is also operating, with significant grain-boundary migration.
6. Rolled specimens developed a {111}⟨112⟩ texture, which severe deformation broke into multiple finer-texture components.

ACKNOWLEDGMENTS

I.G. Lee acknowledges Marko Materials, Inc. for partial support of this work. Helpful comments by Dr. J.D. Whittenberger of NASA Lewis Research Center are also acknowledged.

REFERENCES

1. E.M. Schulson: *Res. Mech. Lett.*, 1981, vol. 1, p. 111.
2. E.M. Schulson and D.R. Barker: *Scripta Metall.*, 1983, vol. 17, p. 519.
3. E.M. Grala: in *Mechanical Properties of Intermetallic Compounds*, J.H. Westbrook, ed., Wiley, New York, NY, 1960, p. 358.
4. A.G. Rozner and R.J. Wasilewski: *J. Inst. Met.*, 1966, vol. 94, p. 169.
5. R.R. Vanervoort, A.K. Mukherjee, and J.E. Dorn: *Trans. ASM*, 1966, vol. 59, p. 930.
6. W.J. Yang and R.A. Dodd: *Met. Sci. J.*, 1973, vol. 7, p. 41.
7. H.L. Fraser, R.E. Smallman, and M.H. Loretto: *Phil. Mag.*, 1973, vol. 28, p. 651.
8. J.D. Whittenberger: *J. Mater. Sci.*, 1987, vol. 22, p. 394.
9. K.H. Hahn and K. Vedula: *Scripta Metall.*, 1989, vol. 23, p. 7.
10. D.M. Vedula, V.M. Pathare, I. Asanidis, and R.H. Titran: in *High Temperature Ordered Intermetallic Alloys*, C.C. Koch, C.T. Liu, and N.S. Stoloff, eds., MRS, Pittsburgh, PA, 1985, vol. 39, p. 411.
11. M. Sherman and K. Vedula: *J. Mater. Sci.*, 1986, vol. 21, p. 1974.
12. V. Mandorf, J. Hartwig, and E.J. Seldin: in *High Temperature Materials*, G.M. Ault, ed., Wiley-Interscience, New York, NY, 1963, vol. II, p. 455.
13. E.K. Storms: in *Refractory Materials*, Academic Press, New York, NY, 1967, vol. 2, p. 35.
14. A. Ball and R.E. Smallman: *Acta Metall.*, 1966, vol. 14, p. 1517.
15. R.J. Wasilewski, S.R. Butler, and J.E. Hanlon: *Trans. TMS-AIME*, 1967, vol. 239, p. 1357.
16. R.T. Pasco and C.W.A. Newey: *Phys. Status Solidi*, 1971, vol. 29, p. 357.
17. M.H. Loretto and R.J. Wasilewski: *Phil. Mag.*, 1971, vol. 23, p. 1311.
18. J. Bevk, R.A. Dodd, and P.R. Strutt: *Metall. Trans.*, 1973, vol. 4, p. 159.
19. H.L. Fraser, R.E. Smallman, and M.H. Loretto: *Phil. Mag.*, 1973, vol. 28, p. 651.
20. H.L. Fraser, M.H. Loretto, and R.E. Smallman: *Phil. Mag.*, 1973, vol. 28, p. 657.
21. A.K. Ghosh and C.H. Hamilton: *Metall. Trans. A.*, 1979, vol. 10A, p. 6-9.
22. A.K. Ghosh: in *Deformation in Polycrystals: Mechanisms and Microstructures*, Proc. 2nd Riso Int. Symp., N. Hansen *et al.*, eds., Roskilde, Denmark, 1981, vol. 1, p. 277.
23. J.D. Whittenberger, R.K. Viswanadham, S.K. Mannan, and B. Sprissler: *J. Mater. Sci.*, 1990, vol. 25, p. 35.
24. K. Vedula, K.H. Hahn, and B. Boulogne: in *High Temperature Ordered Intermetallic Alloys III*, C.T. Liu, A.I. Taub, N.S. Stoloff, and C.C. Koch, eds., MRS, Pittsburgh, PA, 1989, vol. 133, p. 299.
25. P.S. Khadkikar, G.M. Michal, and K. Vedula: *Metall. Trans. A*, 1990, vol. 21A, p. 279.

Forging Limits for an Aluminum Matrix Composite: Part I. Experimental Results

D.-G.C. SYU and A.K. GHOSH

Forging limits in a discontinuously reinforced aluminum (DRA) matrix composite, 2014 Al/15 vol pct Al_2O_3, were determined by compressing samples of various cylindrical geometries under different conditions of temperature, strain rate, and lubrication and measuring the limit strains attained prior to incipient crack formation. In some cases, circumferential grids were machined on the sample surface to obtain the local fracture strain states. Crack formation was caused by the secondary tensile stresses; however, crack propagation was relatively slow and somewhat more severe at 300 °C than at 400 °C. The forging limit of the composite was found to be higher at 400 °C than at 300 °C and also higher at slower strain rates. The plane-strain forging limit of the composite at 300 °C and a strain rate of 0.5 s^{-1} was less than 0.05, while that of the matrix was higher than 0.5. It was found that the forging limits can be influenced by the depth of the circumferential grids and can be lower than those for the smooth surface samples.

I. INTRODUCTION

ALL forging processes basically consist of the compressive deformation of the workpiece within a pair of dies. Depending on the geometry of the dies, varying amounts of lateral constraint may be imposed on the workpiece so that forging operations are classified into two categories: open-die forging and closed-die forging. In open-die forging, the lateral constraint is minimal, and the amount and distribution of lateral metal flow are controlled by factors, such as the geometry of the workpiece, friction and heat transfer between the dies and the workpiece, and the total reduction in the workpiece dimension parallel to the compression direction. In open-die forging, formability is primarily determined by the material's intrinsic properties and the forging conditions. A material's intrinsic properties include the grain structure, the presence or absence of second phase (precipitates or reinforcements), homogeneity (distribution and shape of second phases or texture), and crystal structure. The heat transfer induced by thermal conductivity and specific-heat capacity is also important in hot forging. Other forging conditions that are important include heating rates, forging temperatures (which influence grain growth, recrystallization, phase transformation, the occurrence of a liquid phase, and dissolution or growth of a second phase, *etc.*), as well as strain rates and loading histories (including variation of strain rates and strain states).

During the early stages of closed-die forging, the lateral constraint is minimal, and the process is similar to the open-die forging operation. Cracks and defects of concern can often occur during this stage. As the constraint increases, material begins to fill the shape of the die, and the ease of the metal to flow becomes the key factor for attaining the final shape, *i.e.*, to fill in the corner of the die. Hence, the two important parameters that determine a successful closed-die-forging operation are the strain states that avoid cracking and the ability of plastic flow to fill the shape of the die.

The focus of this investigation is on the experimental determination of the limit strains or *forging-limit diagram* that can be used by the die designer in actual forging operations. The forging limit is defined by a combination of compressive strain (ε_1) and tensile strain (ε_2) on the surface of a forging at the onset of an incipient crack. Forging limit refers to strain limit reached in a localized region of a forging but does not provide a sense of the material's capability for die filling during forging process. To test the capacity for plastic flow, a separate forgeability test was designed, which is discussed in the Appendix and not directly dealt with in the text.

The forging-limit diagram is divided into three regions—the safe, the marginal, and the failure regions. Several investigators have determined the forging-limit diagrams of aluminum alloys[1,2] and steels.[3,4,5] The most common method for determining the forging-limit diagram is to use the upset test in which the strain path (*i.e.*, $\varepsilon_2/\varepsilon_1$) can be varied by using different sample geometries and lubrication conditions. The upset test is an axial-compression test of a right-circular cylinder between two flat, parallel dies. The advantage of the upset test is that the strain state on the cylindrical surface can be easily measured and controlled by varying the frictional boundary conditions, the ratio of height to diameter, and the geometry of the cylinder. Figure 1 shows the possible strain paths in different upset tests. The strain-path variations are caused mainly by the different degree of barreling on the surface of upset samples. At a given circumferential strain, ε_2, which is tensile, the well-lubricated upset sample produces a more negative compressive axial strain, ε_1, than the poorly lubricated one.[2–8] Samples with higher ratios of height to diameter also produce more negative ε_1's than those with lower ratios. Samples with a raised collar on the surface (described in detail in the Section II) provide strain states that are close to positive plane strain ($\varepsilon_1 = 0$).

Experimental data from several previous works[1–7] show

D.-G.C. SYU, formerly Graduate Student, the University of Michigan, is Senior Engineer, Taipei Municipal Government, Hsintien, Taipei, Taiwan, Republic of China. A.K. GHOSH, Professor of Materials Science and Engineering, is with the University of Michigan, Ann Arbor, MI 48109-2136.

Manuscript submitted August 11, 1992.

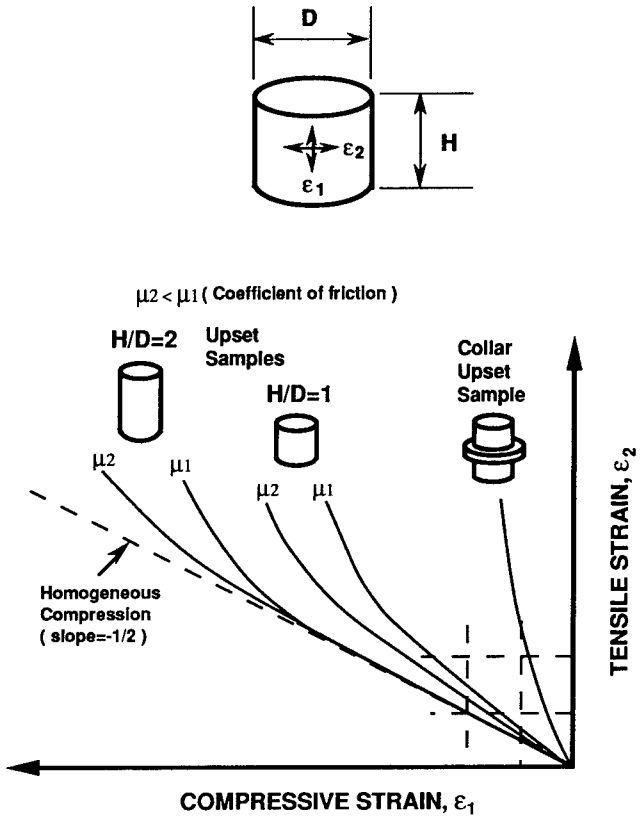


Fig. 1—Change of strain path due to varying sample geometry and frictional condition.

that the forging-limit line is a straight line having a slope of $-1/2$ (parallel to the homogeneous compression line) with an intercept on the positive plane-strain state ($\varepsilon_1 = 0$). Kivivuori and Sulonen[8] have proposed that the slope of the forging-limit line for Al-killed steel (spheroidized Ck-35 with 0.36 pct C) was less than $-1/2$, and that this straight line reached a maximum at around $\varepsilon_2 = 1$ and then became almost parallel to the ε_1 axis. The same observation was also reported by Darvas[3] for Al-killed steel (spheroidized Ck-15 with 0.14 pct C). They explained that a possible reason for this horizontal forging-limit line was the formation of shear bands, since the ductility of the material decreased rapidly during severe deformation. They also observed three fracture modes, namely longitudinal cracking, oblique cracking, and double oblique cracking in their Al-killed steel.

In the present investigation, the material of interest is a discontinuously reinforced aluminum (DRA) matrix composite, *e.g.*, 2014 Al/15 vol pct Al_2O_3. The addition of 15 vol pct Al_2O_3 particles can increase the Young's modulus of 2014 Al alloy from 73 to 92 GPa and the yield strength from 413 to 475 MPa.[9] These composites are produced by ingot metallurgy methods, and the cost of fabricating these materials has declined considerably in recent times. Thus, the DRA matrix composites have potential applications in the automotive industry for manufacturing engine components and brake systems[10–14] most conveniently by hot forging methods.

Although forging-limit diagrams have been determined for many aluminum alloys and steels, forging-limit data for DRA matrix composites are not established at this time. Our objective is to determine the forging limits of this DRA composite under various temperatures and strain rates.

II. EXPERIMENTAL PROCEDURES AND RESULTS

The material used in this investigation was 2014 Al containing 15 vol pct Al_2O_3 particulates.* The composition, microstructures, and mechanical properties of this composite are discussed in Reference 15. The matrix alloy chemistry is approximately Al-4 pct Cu-1 pct Mn-0.5 pct Mg-0.8 pct Si-1.2 pct Ti. The average Al_2O_3 particle size was about 5.6 μm, and the subgrain size was about 2 μm in the composite in the extruded condition.[15] A 2014 Al matrix alloy that was used for comparison was also produced *via* ingot metallurgy and extrusion. It had a grain size of 500 to 2000 μm.

*This material was manufactured by Duralcan Aluminum Composites Corporation, San Diego, CA, and was produced by dispersing aluminum oxide particles into the molten aluminum alloy that was then cast and extruded into a rod with a diameter of 50.8 mm.

Prior to determination of the forging-limit data, it was of interest to select a meaningful temperature and strain-rate range where some ductility can be present for conducting the investigation. Several investigators have reported testing methods and sample geometries to assist in determining an optimum forging condition, such as the diametral-compressive-test method for brittle materials,[16] the heading process to obtain the headability that is defined as a material's ability to be upset without cracking,[17,18] and the wedge-forging test that is used to determine the breakdown temperature ranges for ingot structures,[19] *etc.* In the present investigation, a sample geometry with a notch was designed so that it could be easily fractured, and an approximate fracture strain could be determined under various forging conditions. This test, called the sensitivity test,[20] is discussed in the Appendix of Reference 15. While this test showed some degree of forgeability between 300 °C and 400 °C for our test material, a clear peak in forgeability was not seen.

From a survey of the literature, work done by Pickens *et al.*[21] provides help in the selection of these parameters. They had suggested that for a strain rate of 0.1 s^{-1}, the optimum temperatures for maximizing hot ductility and minimizing flow stress were 427 °C and 507 °C for 7090 Al/20 vol pct SiC and 6061 Al/20 vol pct SiC, respectively. Because the reinforcement content in our material is somewhat lower, a slightly lower temperature was thought to be adequate; thus, 300 °C and 400 °C were chosen as the forging temperatures. The strain rates selected, 0.015 and 0.5 s^{-1}, were guided by the available equipment.

A. *Specimen Geometry and Test Procedure*

To construct the forging-limit diagram, six upset-sample geometries were designed and machined. These samples are identified as SU (shorter upset, no lubricant), SUL (shorter upset, lubricated), TU (taller upset, no lubricant), TUL (taller upset, lubricated), CU (collar upset, no lubricant), and CUL (collar upset, lubricated), as shown

in Figure 2. All samples had the extrusion direction as the compression direction. Graphite was used as a high-temperature lubricant. To provide adequate space for the lubricant during the tests, a shallow well was machined in both ends of the samples and filled with graphite paste before testing.

For SU, SUL, TU, and TUL samples, a portion of the sample surface was machined with circumferential grids, as shown in Figure 2. During the test, strain concentrated within the grooved grid area and resulted in reducing the effective height-to-diameter ratio for these samples, as shown in Figure 3. Regions adjacent to machined grooves act as raised collars (less deformable regions) and cause the strain state to move toward plane strain. Thus such surface grooves lower the forging limits slightly. This effect can be minimized by machining grids over the entire sample surface. To experimentally verify whether the strain paths determined by these six kinds of samples were sufficient to cover the forging-limit diagram, as well as to roughly obtain the strain state beyond which failure occurs, several samples were initially tested in an interrupted manner. Subsequently, the forging-limit diagrams were determined by a single-step process, as discussed in Section II–A–3.

A hydraulic press with programmable movement was used for the interrupted and single-step tests at low strain rates. The press was fitted with a furnace for obtaining isothermal test conditions. An Instron universal testing machine, model 4505, was used for the single-step test with high strain rates. This was also fitted with a furnace with a heating zone of 15.2-cm height and 10.2-cm diameter. Testing was conducted after the desired test temperature was reached and had stabilized for at least 10 minutes. Parallel platens, 10.2 cm in diameter, were used for compressing the samples; these were made of 17-4 PH steel.

1. *Determination of forging limit strains*

The forging limit was defined by the occurrence of an incipient crack at the surface of the deformed sample. By measuring strain states (combination of ε_1 and ε_2) before and after such incipient surface-crack formation on the deformed samples, the forging-limit diagram was constructed as a function of temperature and strain rate for use in forging-process design.

For the SU, SUL, TU, and TUL samples that had machined grids, h_{def} and h_o were the distances between two grooves on the deformed and undeformed samples, respectively, and were measured by using an optical microscope, while d_{def} and d_o were the maximum diameters of the deformed and undeformed samples, respectively, and were measured by a micrometer. For determining strains from the deformed specimens, the following equations were used:

$$\varepsilon_1 = \ln (h_{def}/h_o) \qquad [1a]$$

$$\varepsilon_2 = \ln (d_{def}/d_o) \qquad [1b]$$

For the CU and CUL samples, h_{def} and h_o were the collar heights of the deformed and undeformed samples, respectively, and d_{def} and d_o were the collar diameters of the deformed and undeformed samples, respectively, and were measured by a micrometer.

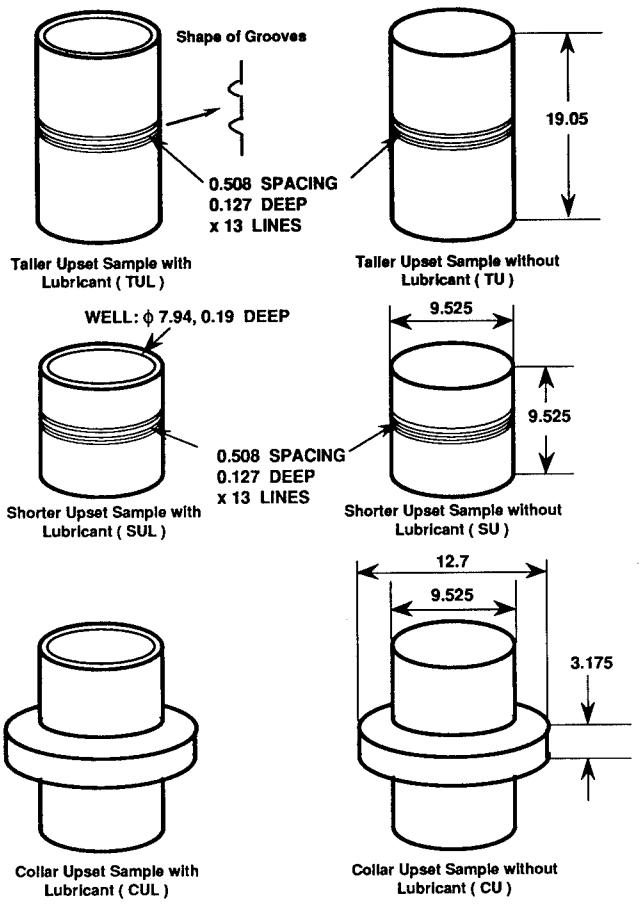


Fig. 2—Six sample geometries for constructing forging-limit diagrams, in millimeters.

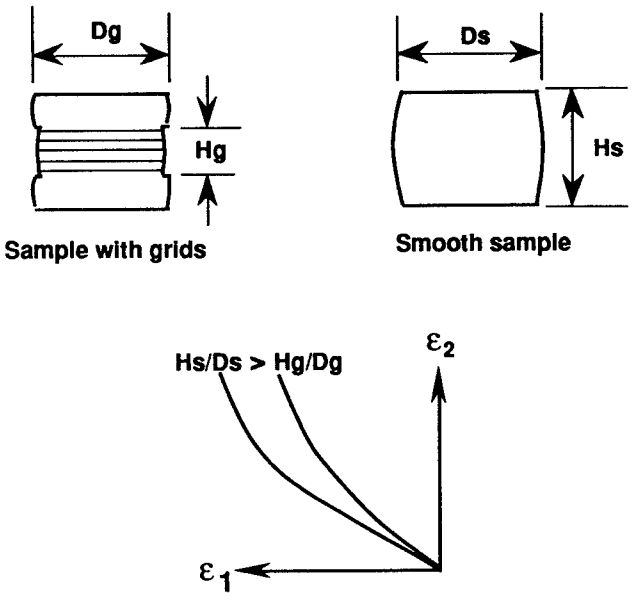


Fig. 3—Ratio of height to diameter for the sample with machined grids at the barreled area, H_g/D_g, smaller than that for the smooth-surface sample, H_s/D_s. The strain state for the sample with machined grids has a less-negative compressive strain and a lower tensile strain.

2. *Determination of strain paths and approximate forging limits by interrupted testing*

Initially, several samples were tested in an interrupted manner at 400 °C to determine the strain paths (ε_2 *vs* ε_1) and obtain an approximate forging limit that could be a guide for the single-step test. For this study, samples were deformed to given levels of strain. After deformation, the sample was removed from the furnace, and ε_1 and ε_2 were measured. After measuring the strains, the deformed sample was placed back in the furnace, heated to 400 °C, held for 20 to 30 minutes to stabilize the temperature, and compressed to a higher strain level. At several stages of interruption, ε_1 and ε_2 were determined, and an effective strain was calculated from the relation[22]

$$\varepsilon_{\text{eff}} = [4/3\ (\varepsilon_1^2 + \varepsilon_1\varepsilon_2 + \varepsilon_2^2)]^{0.5} \qquad [2]$$

The average strain rate was defined as the incremental effective strain divided by the incremental time. Based on the combinations of ε_1 and ε_2 in each step, the strain paths of these six kinds of samples were determined, as shown in Figure 4(a). Examination for the appearance of incipient cracks was made by unaided eyes during the tests. If there was no crack on the surface, this safe strain combination (ε_1 and ε_2) was then plotted in the diagram and labeled as ○ (open circle). On the other hand, if any incipient crack appeared on the deformed-sample surface, the strain state was measured and labeled as x (cross) in the same diagram. Figure 4(b) shows data points for 400 °C and a strain rate of 0.1 s^{-1}. A forging-limit diagram was eventually drawn to separate the circles from the crosses.

3. *Determination of forging-limit diagrams by single-step test*

Because forging is generally a single-step process, the forging-limit diagrams under various conditions were largely determined by single-step deformation. Samples with machined grids (SU, SUL, TU, and TUL samples) were tested to a selected strain level (based on the limit strains obtained from the interrupted test) and then visually examined. The safe strain state was labeled as ○, and the strain state at the local cracked area was measured and labeled as x in a diagram, as shown in Figures 5 and 6 at 300 °C and 400 °C, respectively, within a strain-rate range of 0.015 s^{-1} to 0.5 s^{-1}. Then a forging-limit diagram was constructed by drawing two lines (marginal region) to separate the failed and safe strain states. It was found that the forging limit decreased toward the plane-strain ($\varepsilon_1 = 0$) state. The absolute magnitude of the slope was around 0.2 to ~0.4, which is

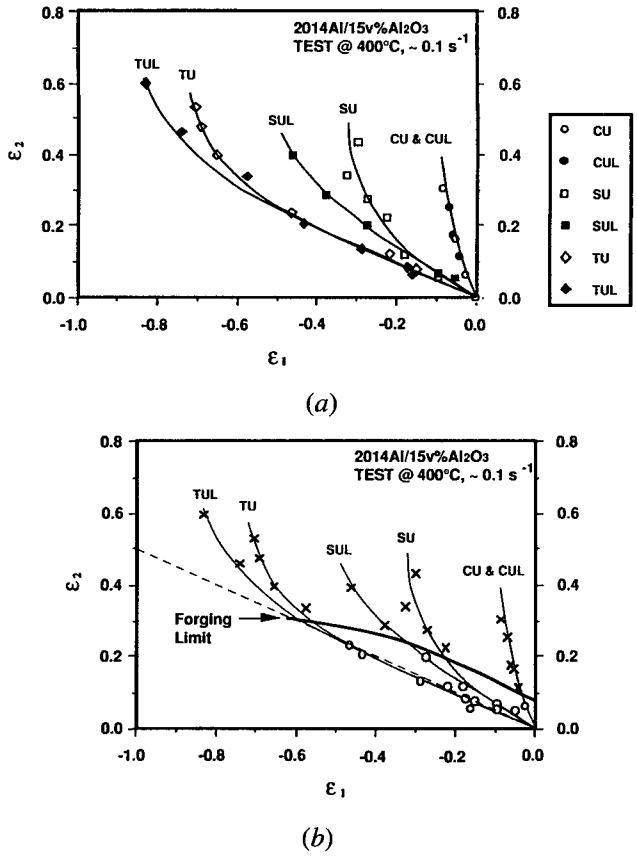


Fig. 4—(*a*) Strain-path variation and (*b*) forging-limit diagram by the interrupted test method for 2014 Al/15 vol pct Al_2O_3. Samples were tested at 400 °C and an average strain rate of 0.1 s^{-1}.

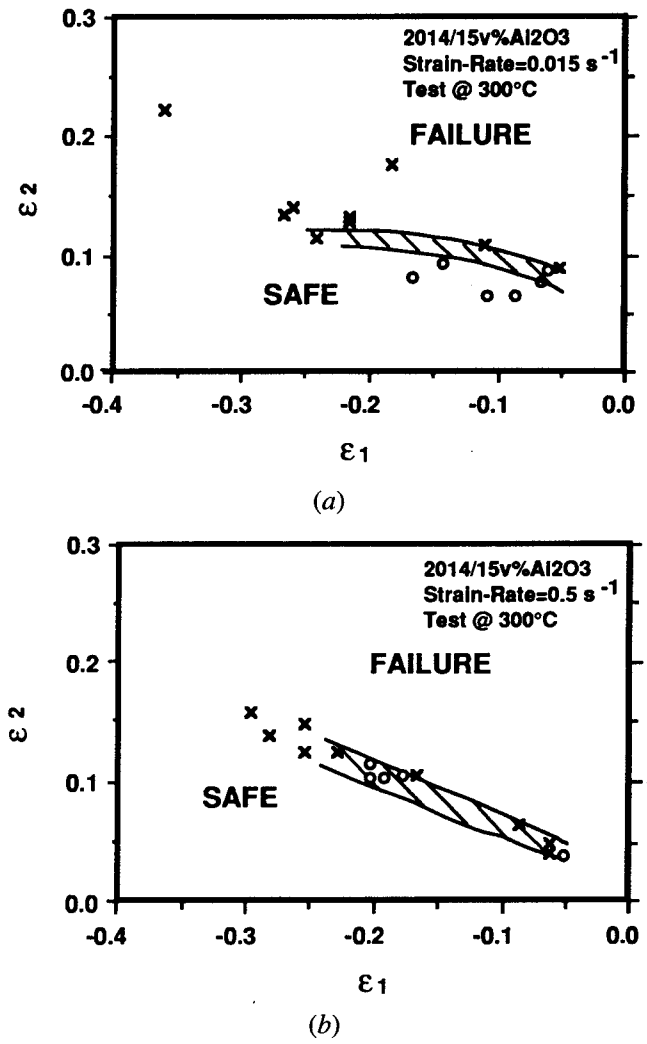


Fig. 5—Forging-limit diagrams of 2014 Al/15 vol pct Al_2O_3 determined from SU, SUL, TU, and TUL tests at 300 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In these tests, samples had machined grids on the surface.

less than that from previous work on Al alloys and steels (slope = $-1/2$).[2–8] In addition, the forging-limit curves in Figures 5 and 6 were not a straight line, possibly because they contain effects from the machined grids.

As mentioned before, the machined grids on the SU, SUL, TU, and TUL samples acted as raised collars, as seen in Figure 7. At a given ε_2, the circumferential grooves compressed more than the adjoining regions (raised), *e.g.*, the strain state at failure was closer to plane strain ($\varepsilon_1 = 0$). Consequently, the samples with machined grids would have a lower forging limit in comparison to the smooth-surface sample. To avoid this effect, the SUL samples without machined grids were also tested. The strain states from the smooth-surface SUL samples, then, were combined with the strain states of the CU and CUL samples (there were no machined grids on the CU and CUL samples) using the method described in Section II–A–4 to construct a forging-limit diagram without the machined-grid effect.

4. *Approximate method for estimating the forging limit for smooth-surface specimens*

Since smooth-surface specimens do not offer a direct method to measure the axial strain, ε_1, before and after incipient cracking, an approximate method was adopted. This method first utilizes an upper bound and a lower bound curve from tests performed on samples with machined grids and subsequently utilizes a refined approach based on the Kobayashi analysis,[23] as shown in Figures 8 and 9.

The upper bound is dictated by the interrupted test in which samples had machined grids, as discussed in Section II–A–2 (Figure 4(a)). Since samples with machined grids provide strain states closer to plane strain than a smooth-surface sample, as shown in Figure 7, the actual strain path for the smooth-surface SUL sample in Figures 8 and 9 should be on the left-hand side of the strain path obtained from this method.

The lower bound is based on the assumption that the upset sample deforms in a homogeneous manner, *i.e.*, that $\varepsilon_1 = -2\varepsilon_2$. The ratio of $\varepsilon_2/\varepsilon_1$ is, however, less than $-1/2$, because the smooth-surface SUL samples barrel slightly, as seen in Figure 10. Since this introduces an additional tensile component, the actual strain path for the smooth-surface SUL sample in Figures 8 and 9 should be on the right-hand side of the strain path obtained from the homogeneous compression method. In Figure 10, it was also found that the barreling effect was less at 300 °C than at 400 °C for the smooth-surface SUL samples probably because of a decrease in the strain-hardening rate of the workpiece with increasing temperature. It is known that a higher strain-hardening rate promotes more uniform deformation in the workpiece.

Having defined the upper and lower bounds on the strain path, an approach to the more exact strain path was made by using a detailed analysis by Kobayashi.[23] Figures 8 and 9 show strain paths from the Kobayashi analysis and the two other bounds, with data points from safe and failed samples. Subsequently these data points from the Kobayashi analysis were used to determine the forging limits.

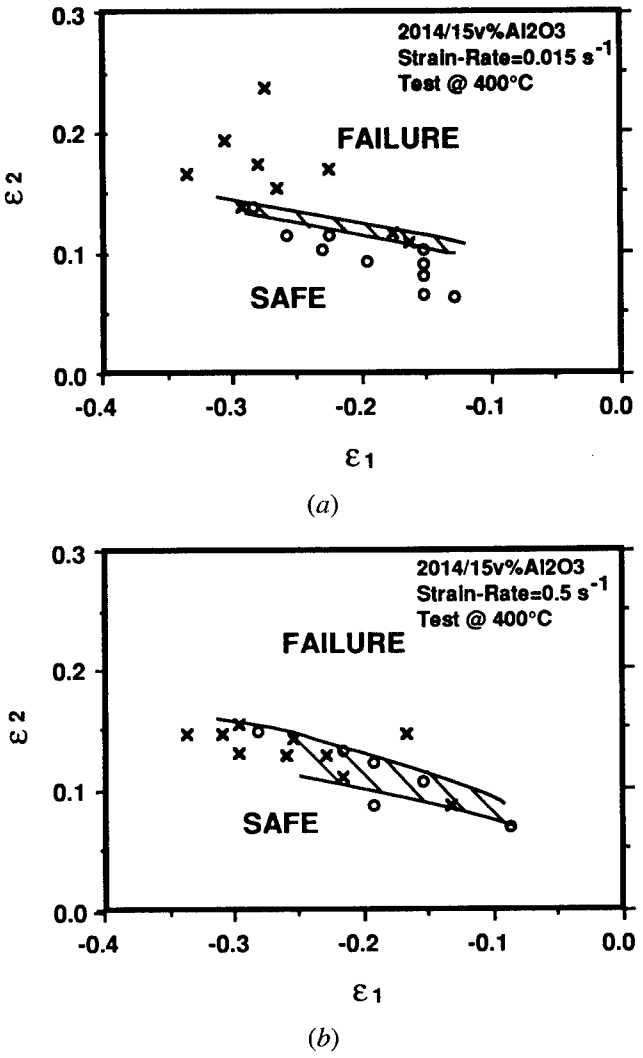


Fig. 6—Forging-limit diagrams of 2014 Al/15 vol pct Al_2O_3 determined from SU, SUL, TU, and TUL tests at 400 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In these tests, samples had machined grids on the surface.

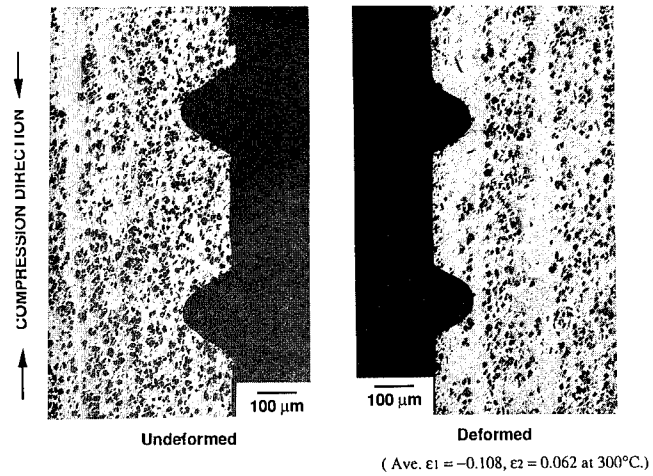


Fig. 7—Optical micrograph showing grooves compressed more than raised collars in an upset sample with machined grids.

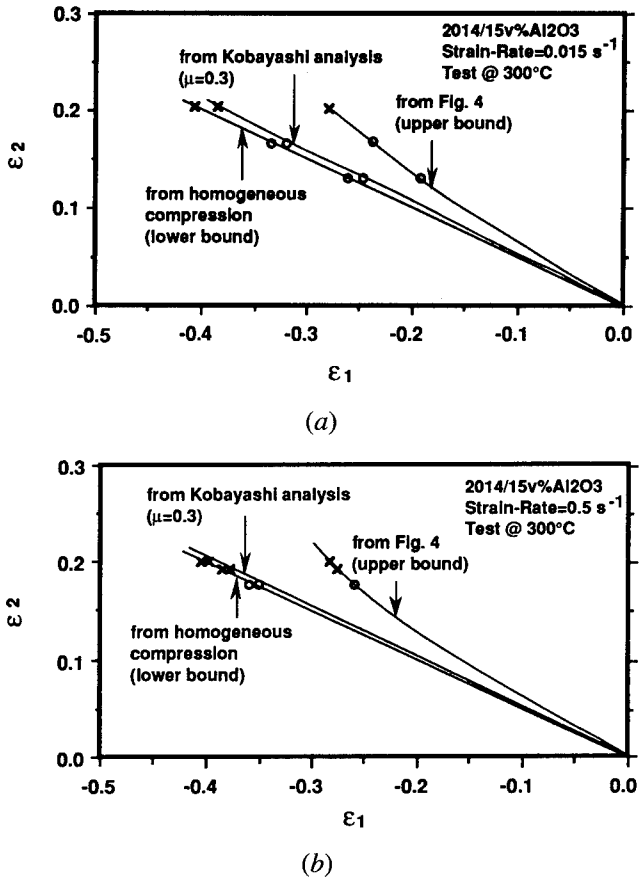


Fig. 8—Assumed strain paths in the SUL test for 2014 Al/15 vol pct Al_2O_3 at 300 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In this test, samples had smooth surfaces (ground by 600-grit sandpaper). The determination of ε_1 was based on the strain path in Figure 4, homogeneous compression, and Kobayashi analysis.

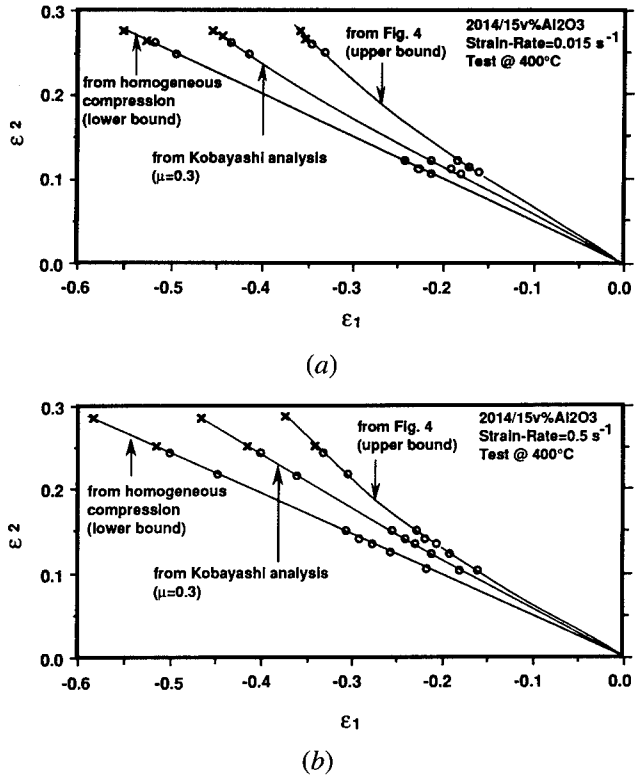


Fig. 9—Assumed strain paths in the SUL test for 2014 Al/15 vol pct Al_2O_3 at 400 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In this test, samples had smooth surfaces (ground by 600-grit sandpaper). The determination of ε_1 was based on the strain path in Figure 4, homogeneous compression, and Kobayashi analysis.

B. *Barreling Correction on Smooth-Surface Specimens*

Several researchers[3,23,24] have also attempted an analytical treatment for the barreling effect that generates a secondary tensile stress parallel to the compression direction and changes the value of ε_1. Banerjee[24] tested aluminum cylinders with various ratios of height to diameter and lubrication conditions and found that the curvature of the barrel fit closely to a circular arc, and its radius followed a power law with the true axial compressive stress. Darvas[3] proposed a relationship between ε_1 and ε_2 that fit the experimental data quite well. He concluded that ε_1 was an arc tangent function of ε_2,

$$\varepsilon_1 = -2C \tan^{-1}(\varepsilon_2/C) \qquad [3]$$

where C is a constant. In his experiment a C value of 0.1369 was selected to fit the experimental results for an upset sample with a height-to-diameter ratio of 1 under adhesive frictional conditions.

Kobayashi[23] also assumed that the curvature of the barrel is a circular arc and developed a closed-form solution for the relationship between ε_1 and ε_2. He assumed that the incremental radial displacement du, in upsetting solid cylinders, is given by

$$du = a_1 r(1 - \mu z^2) + a_2 r^3 \qquad [4a]$$

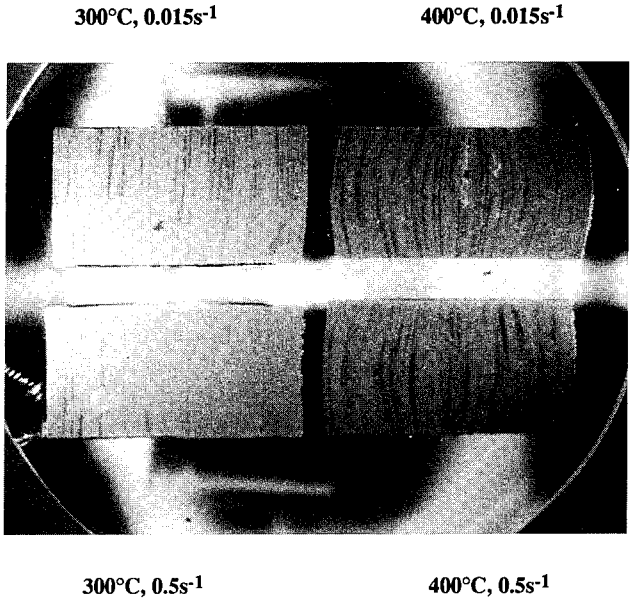


Fig. 10—Photograph of deformed-smooth-surface SUL samples (magnification 4 times) of 2014 Al/15 vol pct Al_2O_3 at 300 °C and 400 °C with effective strain rates of 0.015 s^{-1} and 0.5 s^{-1}. The barreling effect is more severe at 400 °C than at 300 °C.

and the incremental axial displacement dv can be obtained from the incompressibility equation,

$$\frac{\partial(du)}{\partial r} + \frac{du}{r} + \frac{\partial(dv)}{\partial z} = 0 \qquad [4b]$$

where r is the radial coordinate and z is the axial coordinate of material points in the sample and $z = 0$ at the half-height of the sample. The term μ is the coefficient of friction following the Coulomb friction law. By introducing the radius of the equatorial free surface, R, and the radius of curvature of the barrel, ρ, a_1 and a_2 can be solved under the boundary conditions, *i.e.*, when $\rho \to \infty$ or $\mu \to \infty$, then $\varepsilon_1 = -2\varepsilon_2$. From these, he obtained

$$d\varepsilon_1 = \int \frac{8\frac{\rho}{R} + 2 - 4\mu R\rho}{4\frac{\rho}{R} + 4 - \mu R\rho} d\varepsilon_2 \qquad [4c]$$

With measurements of R and ρ from deformed samples using the interrupted-test method, and, selecting the value of μ to fit the experimental strain paths in the various upset tests, a good predictability of barrel shape was given by his analysis.

In the present investigation, the smooth-surface SUL sample was tested to a given strain by the single-step method. Then, R, ρ, and ε_2 were measured from each deformed smooth-surface SUL sample, as shown in Table I. Based on Table I and an incremental form of Eq. [4c], the ε_1 and strain paths for the smooth-surface SUL samples were determined. These are shown in Figures 8 and 9 at 300 °C and 400 °C, respectively, where μ was chosen as 0.3 for the calculation. The data points of the SUL tests calculated by using the Kobayashi analysis fell between ε_1 from Figure 4 and that from homogeneous compression. These data points were then plotted with the data from the CU and CUL tests (no machine grids) to determine approximate forging-limit diagrams by drawing lines to separate the safe and failed strain states, as shown in Figures 11 and 12 at 300 °C and 400 °C, respectively. Again the forging limit was found to drop toward the plane-strain state.

III. DISCUSSION

A. *Forging Limits*

A comparison of the forging-limit diagrams for the various samples is presented in this section. Figure 13(a) shows the mean lines of the marginal region (shaded area) from each of Figures 5 and 6. Figure 13(b) shows the forging limits obtained from the smooth-surface SUL (strain paths determined by using the Kobayashi analysis), CU, and CUL tests under various forging conditions. We find that at 400 °C and a strain rate of 0.015 s^{-1}, the limiting tensile strain (ε_2) was about 25 pct lower for the machined-grid samples. This difference may be reduced by using shallower grids (<0.1-mm deep). Figures 13(a) and (b) also show that the forging limit was lower at the lower temperature (300 °C) and higher strain rate (0.5 s^{-1}). For example, at the slower strain rate (0.015 s^{-1}), the plane-strain ($\varepsilon_1 = 0$) forging limit of the composite at 400 °C was around 0.08; however, it was about 0.04 at 300 °C. Also, at 400 °C, the plane-strain forging limit of the composite at the higher strain rate (0.5 s^{-1}) was around 0.03 as compared with 0.08 at the lower strain rate (0.015 s^{-1}).

Table I. Values of R, ρ, and ε_2 Measured from the Deformed-Smooth-Surface SUL Samples*

Rate (s^{-1})	Sample Number	R (mm)	ρ (mm)	ε_2	$\delta\varepsilon_2$	ε_1
400 °C						
0.015	1	5.28	30.0	0.102	0.102	−0.180
0.015	2	5.33	28.5	0.107	0.005	−0.189
0.015	3	5.38	25.5	0.115	0.008	−0.212
0.015	4	6.10	14.0	0.247	0.132	−0.415
0.015	5	6.20	12.5	0.262	0.015	−0.437
0.015	6	6.22	11.5	0.266	0.004	−0.443
0.015	7	6.30	10.5	0.281	0.015	−0.446
0.5	1	5.33	23.0	0.108	0.108	−0.185
0.5	2	5.41	19.5	0.123	0.015	−0.210
0.5	3	5.46	18.0	0.134	0.011	−0.228
0.5	4	5.51	17.5	0.143	0.009	−0.243
0.5	5	5.54	17.25	0.148	0.005	−0.251
0.5	6	5.94	15.0	0.222	0.074	−0.367
0.5	7	6.10	12.5	0.247	0.025	−0.405
0.5	8	6.15	11.5	0.255	0.008	−0.417
0.5	9	6.38	10.0	0.292	0.037	−0.469
300 °C						
0.015	1	5.44	—	0.132	0.132	−0.263
0.015	2	5.59	~100	0.161	0.029	−0.320
0.015	3	5.79	~75	0.196	0.035	−0.370
0.5	1	5.66	—	0.176	0.176	−0.350
0.5	2	5.77	—	0.190	0.014	−0.380
0.5	3	5.79	—	0.198	0.008	−0.390

*The value of ε_1 can be determined from the values in this table and the Kobayashi analysis.

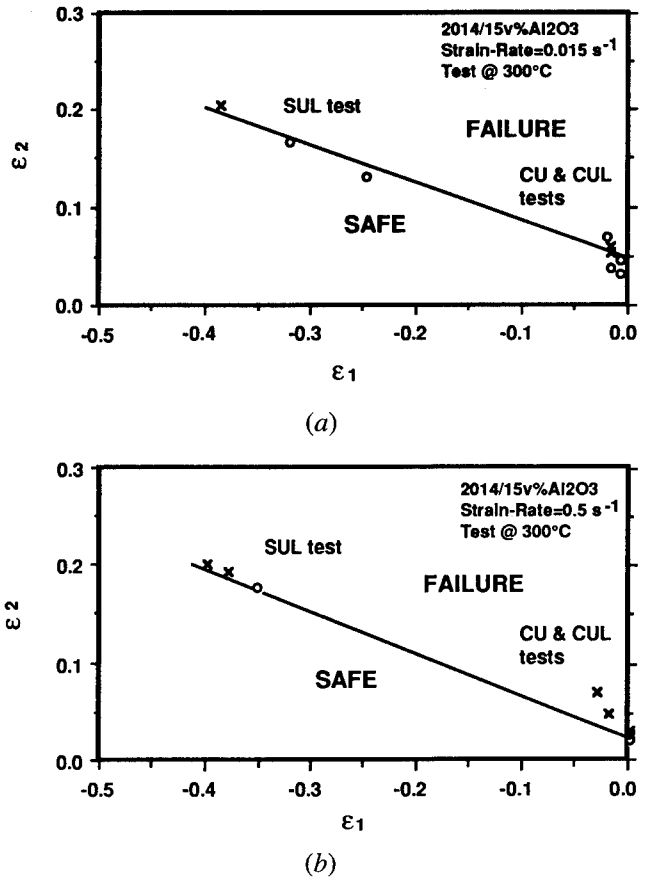


Fig. 11—Forging-limit diagrams of 2014 Al/15 vol pct Al_2O_3 determined from CU, CUL, and SUL tests at 300 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In these tests, samples had smooth surfaces (ground by 600-grit sandpaper). The determination of ε_1 in the SUL test with smooth-surface samples was based on the Kobayashi analysis.

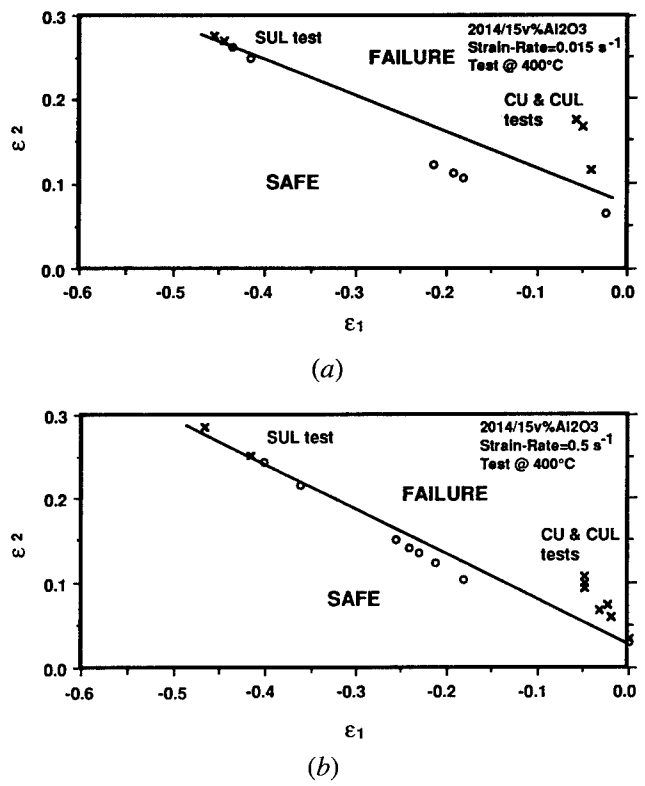


Fig. 12—Forging-limit diagrams of 2014 Al/15 vol pct Al_2O_3 determined from CU, CUL, and SUL tests at 400 °C with effective strain rates of (*a*) 0.015 s^{-1} and (*b*) 0.5 s^{-1}. In these tests, samples had smooth surfaces (ground by 600-grit sandpaper). The determination of ε_1 in the SUL test with smooth-surface samples was based on the Kobayashi analysis.

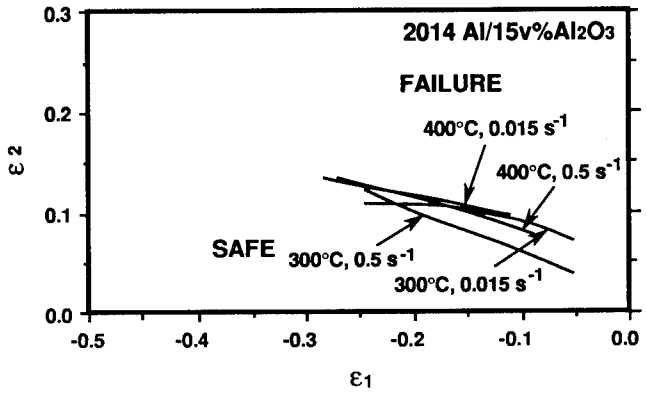
Figure 14 shows a comparison of forging limits between 2014 Al and the composite at 300 °C with a strain rate of 0.5 s^{-1}. Although the forging-limit diagram of 2014 Al was not obtained, its plane-strain forging limit was found to be higher than 0.5, while that of the composite was less than 0.05. This demonstrates the significant change of the forging-limit diagram due to the presence of Al_2O_2 particles.

B. *Fracture Mode*

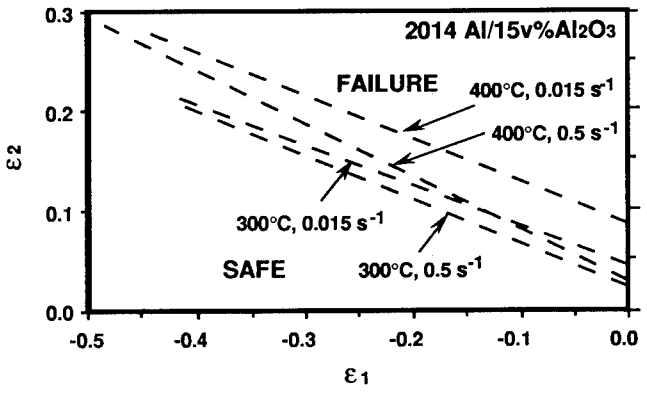
The fracture mode of the incipient crack in this composite material was longitudinal cracking parallel to the compression direction caused by tensile stress, as seen in Figure 15. However, if the samples were continuously deformed, regardless of the occurrence of the incipient crack, to approximately −1.1 true strain in height reduction, double-oblique cracking was observed, as shown in Figure 16. This has also been found in Al-killed steels Ck 35[8] and Ck 15 and in Cu-40 wt pct Zn-2 wt pct Pb brass.[3] Kivivuori and Sulonen[8] have proposed a possible explanation that the occurrence of double-oblique cracking is due to the formation of shear bands[25,26,27] at high deformation levels and is associated with a rapid decrease in ductility. Because of the poor ductility in this composite, the observed double-oblique cracking after high deformation is in agreement with prior work.

In the present work, incipient cracks at the deformed-sample surface were used to define the forging limit. However, these surface incipient cracks did not grow rapidly into the interior of the composite until a high compressive strain was achieved. Figures 17(a) and (b) are cross-sectional micrographs of upset samples showing cracks that have propagated into the interior of the composite at a height strain of 0.9. Most of the fracture occurred in the matrix or at the Al-Al_2O_3 interface during the upset test at 300 °C and 400 °C. Matrix cracks were also observed within particle clusters a small distance ahead of the propagating crack possibly due to the formation of a plastic zone. It is also seen that crack propagation was less severe at 400 °C than at 300 °C when similar high reductions were given.

The resistance to crack propagation was determined quantitatively by measuring the maximum crack depth at various temperatures and stress states. Figure 18 shows the relation between maximum crack depth and true strain in a single-step test for SU and CU samples, *i.e.*, each point indicates one deformed sample. The strain was measured by height reduction in the SU samples and collar-height reduction in the CU samples. Comparing the SU samples at 300 °C and 400 °C, it was found that the maximum crack depth significantly increased above a height strain of about 0.6 at 300 °C but did not change with increasing true strain at 400 °C. When the strain



(*a*)



(*b*)

Fig. 13—Comparisons of forging-limit diagrams of 2014 Al/15 vol pct Al_2O_3 at 300 °C and 400 °C with effective strain rates of 0.015 s^{-1} and 0.5 s^{-1}. (*a*) Determined from SU, SUL, TU, and TUL tests. In these tests, samples had machined grids on the surface. (*b*) Determined from CU, CUL, and SUL tests (Kobayashi analysis, $\mu = 0.3$). In these tests, samples had smooth surfaces (ground by 600-grit sandpaper).

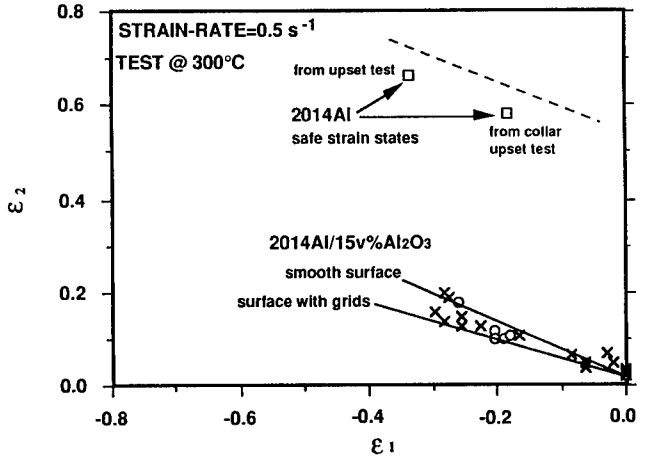


Fig. 14—Comparison of forging-limit diagrams between 2014 Al/15 vol pct Al_2O_3 and 2014 Al at 300 °C and a strain rate of 0.5 s^{-1}. The two strain states for 2014 Al indicate a safe condition, *i.e.*, the forging limit for 2014 Al (dashed line) should be higher than those two strain states.

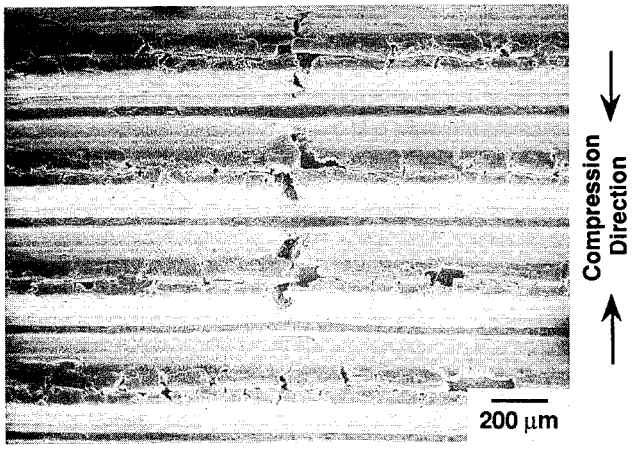


Fig. 15—Scanning electron micrograph illustrating the incipient cracking on the machined grids in a deformed 2014 Al/15 vol pct Al_2O_3 upset sample. Test condition: 300 °C, $e_1 = -0.26$, $e_2 = 0.137$.

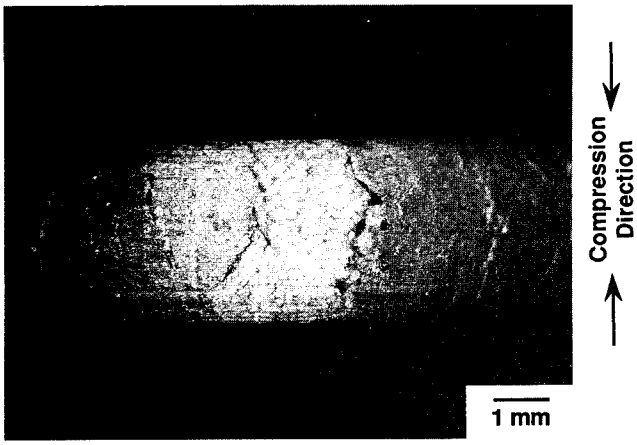


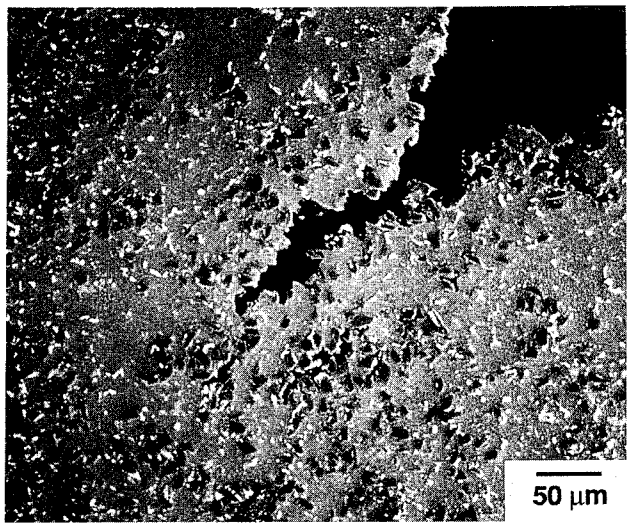
Fig. 16—Photograph showing double-oblique cracking on a deformed 2014 Al/15 vol pct Al_2O_3 upset sample. Test condition: 250 °C, strain rate of 0.1 s^{-1}, true strain of ln $(h_0/h) = 1.1$.

state was closer to positive plane strain, *i.e.*, in the CU test, crack propagation became even more severe at 400 °C.

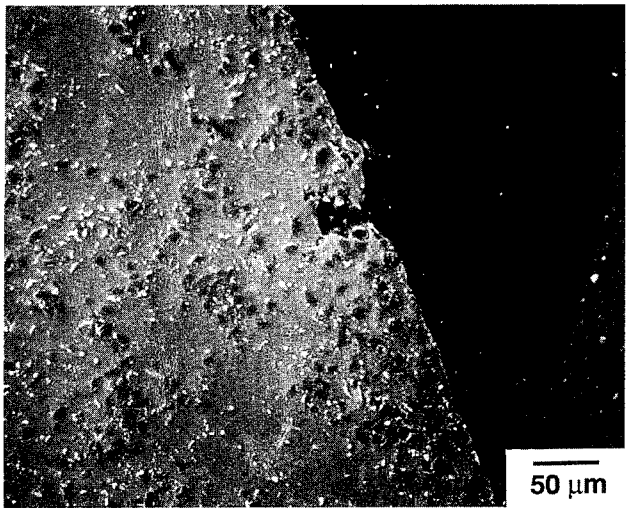
IV. SUMMARY AND CONCLUSIONS

An experimental method for determining forging-limit diagrams was extended to DRA matrix composites by conducting a series of hot upset tests with varying sample geometry and frictional conditions.

1. The experimental results showed that the forging limit of 2014 Al/15 vol pct Al_2O_3 is higher at 400 °C than at 300 °C, and it increases further at lower strain rates in the range 0.015 s^{-1} to 0.5 s^{-1}. For example, at 0.015 s^{-1} the plane-strain ($\varepsilon_1 = 0$) forging limit of the composite at 400 °C is around 0.08; however, it is about 0.04 at 300 °C.
2. The plane-strain ($\varepsilon_1 = 0$) forging limit of the composite at 300 °C and a strain rate of 0.5 s^{-1} is less than 0.05, while that of the matrix is above 0.5. That is the 15 vol pct of Al_2O_3 particles significantly lowers the forging limit of 2014 Al.



(*a*)



(*b*)

Fig. 17—Scanning electron micrographs showing cracks propagating into 2014 A/15 vol pct Al_2O_3 at (*a*) 300 °C and (*b*) 400 °C. Test condition: a strain rate of 0.1 s^{-1} and a height true strain of 0.9 for both temperatures. Matrix cracks were observed within particle clusters a small distance ahead of the propagating crack.

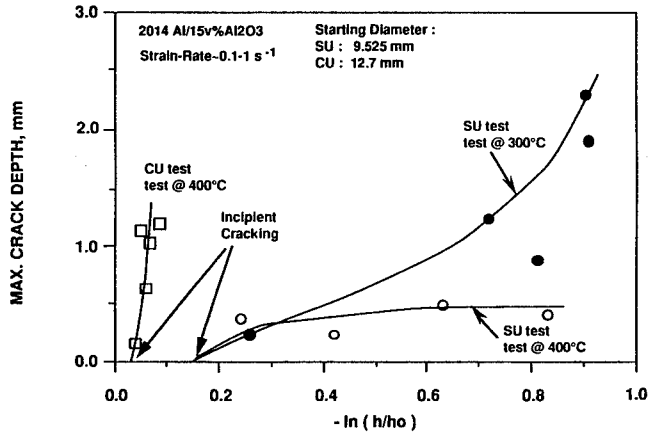


Fig. 18—Crack propagation becoming more severe at lower temperature as well as at the strain state closer to plane-strain state.

3. The machined grids on the sample surfaces concentrate the strain during upsetting. The local strain state moves closer to positive plane strain ($\varepsilon_1 = 0$), which lowers the forging limit. The limiting tensile strain (ε_2) decreases about 25 pct at 400 °C and a strain rate of 0.015 s^{-1}. This difference may be reduced by improving the gridding method. To estimate the forging limit for a sample without machined grids, the Kobayashi analysis can be employed.
4. In the composite, incipient cracks are caused by the secondary tensile stress and are parallel to the compression direction (longitudinal cracking) in the upset test. When the deformation state is highly negative (toward pure shear), double-oblique cracking appears.
5. Incipient cracks do not rapidly grow once initiated in the composite. The resistance to crack propagation is greater at 400 °C than at 300 °C. This resistance to crack propagation at 400 °C decreases when the strain state is closer to positive plane strain.

APPENDIX

Forgeability test

The forgeability test was used to determine the ability of the metal to flow at different temperatures while under a given applied load. The purpose of this test was to assess the optimum temperature and strain rate at which the material could flow with relative ease to fill a complex die shape. A programmable PHI hydraulic hot press

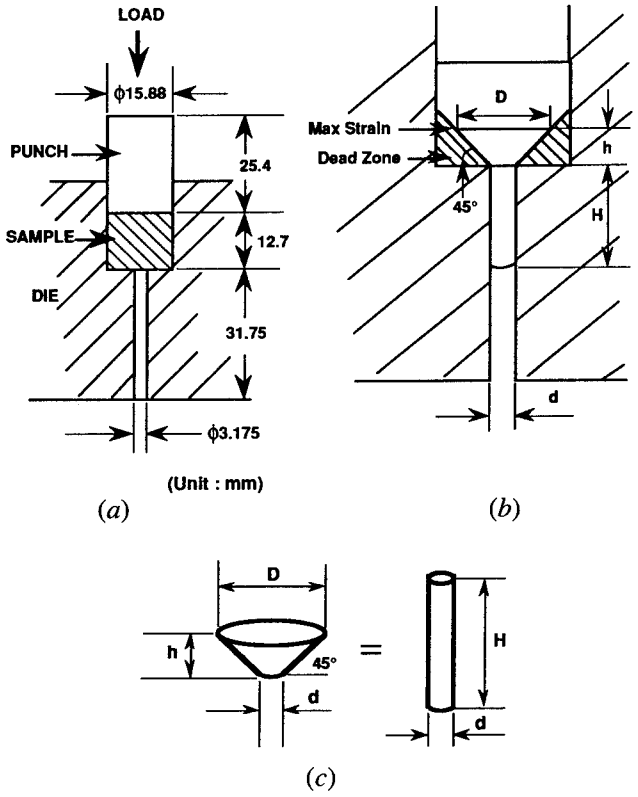


Fig. A1—(*a*) Schematic diagram of the forgeability test (*b*) and (*c*) The volume of the extruded cylinder equal to the volume of the cone.

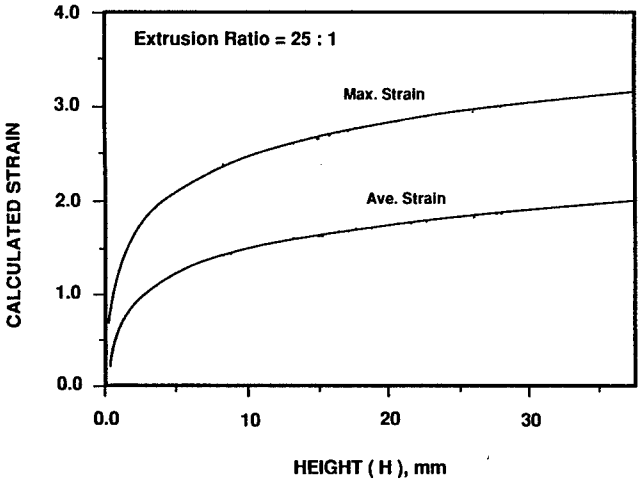


Fig. A2—Calculated maximum and average strains plotted as a function of extrusion height in the forgeability test.

was used in this test. In order to determine the strain with respect to the extrusion height, a dead zone was assumed, as seen in Figure A1. The maximum strain and the average strain were calculated from the following:

$$\frac{1}{3}\frac{\pi D^2}{4}\left(h+\frac{d}{2}\right)-\frac{1}{3}\frac{\pi d^2}{4}$$

$$\frac{d}{2}=\frac{\pi d^2}{4}H \qquad \text{(volume constancy)} \qquad [A1]$$

$$D=2\left(h+\frac{d}{2}\right) \qquad \text{(from geometry)} \qquad [A2]$$

$$\text{Maximum strain} = 2\ln(D/d) \qquad [A3]$$

where d is the diameter of the extruded cylinder and H is the extrusion height. D is the maximum diameter of the cone calculated from Eqs. [A1] and [A2]. Once the value of D was determined, the maximum strain could be obtained by Eq. [A3]. To determine the average strain between d and D of the cone, the cone was divided into 100 discs of equal thickness. The average strain was determined from

$$Dn = d + \frac{(D-d)}{100}\times n \qquad (n = 1, 2, \ldots, 100) \qquad [A4]$$

$$\text{Average strain} = \frac{2\Sigma \ln(Dn/d)}{100} \qquad [A5]$$

Based on these equations, the maximum strains and average strains were plotted as functions of the extrusion height, as shown in Figure A2. It should be noted that these equations are only valid in the conic region ($D <$ 15.88 mm), which was maintained during the present study. Figure A3 shows the result of the forgeability test at strain rates from 0.09 to 0.13 s^{-1}, where it was found that the metal flow became easier above 400 °C. The effect of the loading history, to determine the various strain rates *vs* temperature, was also studied, as seen in Figure A4. For both loading histories in Figure A4, the shapes of the forgeability curves were similar; that is, at the range of 0.06 to 0.13 s^{-1}, the strain rate did not significantly influence the shape of the forgeability curve. Consequently, it is concluded that the ability of metal to flow for 2014 Al/15 vol pct Al_2O_3 is mainly dominated by temperature.

ACKNOWLEDGMENTS

The authors would like to thank the AES, GM Technical Center, for supporting this work. Fruitful discussions with Dr. B. Taylor and Mr. A. McMillan of GM Technical Center are gratefully acknowledged.

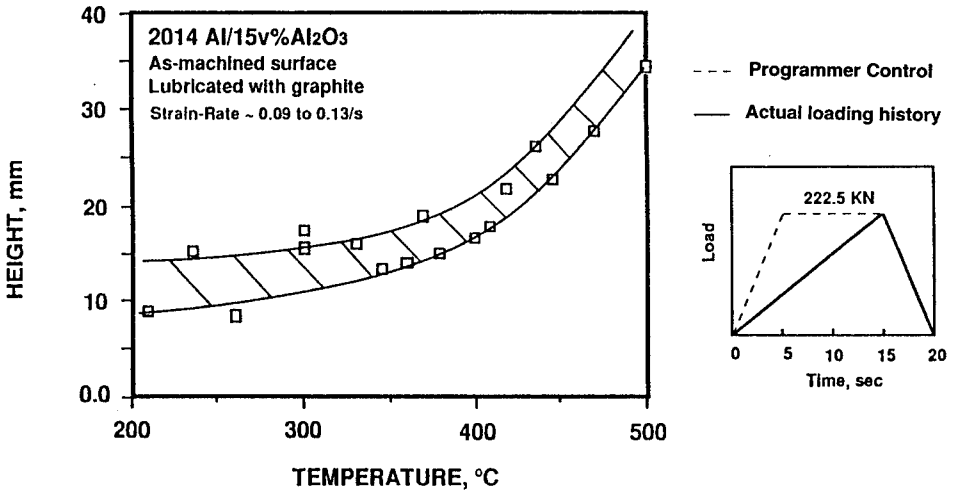


Fig. A3—Temperature effect in the forgeability test. Above 400 °C, the metal plastically flows more easily.

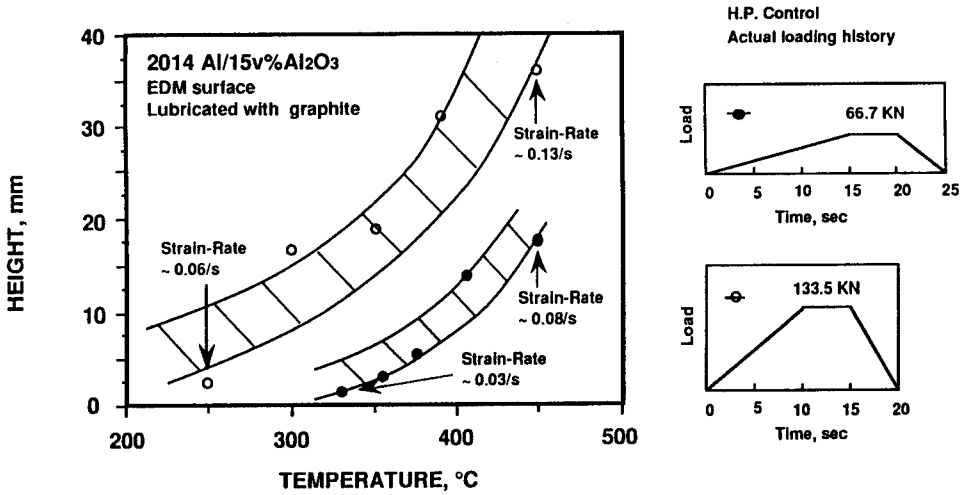


Fig. A4—Effect of loading history in the forgeability test. A change in loading history will vary the extrusion height but not the shape of forgeability curves in the forgeability test.

REFERENCES

1. D.A. Domfeld and E. Diei: *ASME J. Eng. Mater. Technol.*, 1982, vol. 104, pp. 145-52.
2. T. Erturk and E. Kazazoglu: in *Formability of Metallic Materials—2000 A.D.*, ASTM STP 753, J.R. Newby and B.A. Niemeier, eds., ASTM, Philadelphia, PA, 1982, pp. 19-34.
3. Z. Darvas: *Mater. Sci. Eng.*, 1985, vol. 70, pp. 101-10.
4. H.A. Kuhn, P.W. Lee, and T. Erturk: *ASME J. Eng. Mater. Technol.*, 1973, Oct., pp. 213-18.
5. R. Sowerby, I. O'Reilly, N. Chandrasekaran, and N.L. Dung: *ASME J. Eng. Mater. Technol.*, 1984, vol. 106, pp. 101-06.
6. G. Ziaja, L. Cser, and Z. Darvas: in *Advanced Technology of Plasticity 1987*, K. Lange, ed. Springer-Verlag Ltd., Berlin, vol. 2, pp. 999-1003.
7. N.L. Dung and O. Mahrenholtz: in *Advanced Technology of Plasticity 1987*, K. Lange, ed. Springer-Verlag Ltd., Berlin, vol. 2, pp. 1013-20.
8. S. Kivivuori and M. Sulonen: *Ann. CIRP*, 1978, vol. 27, pp. 141-45.
9. Duralcan W2A 15A Product Specification, Dural Aluminum Composites Corporation, San Diego, CA, Apr. 7, 1989.
10. D.B. Rogers: Paper presented at the *17th Annual Automotive Materials Symp.*, Ann Arbor, MI, May 1990.
11. J.E. Allison: Paper presented at the *17th Annual Automotive Materials Symp.*, Ann Arbor, MI, May 1990.
12. A.K. Sachdev: Paper presented at the *17th Annual Automotive Materials Symp.*, Ann Arbor, MI, May 1990.
13. J.M. Corwin: Paper presented at the *17th Annual Automotive Materials Symp.*, Ann Arbor, MI, May, 1990.
14. C.P. You: Paper presented at the *17th Annual Automotive Materials Symp.*, Ann Arbor, MI, May, 1990.
15. D.-G.C. Syu: Ph.D. Dissertation, University of Michigan, Ann Arbor, MI, 1992.
16. H. Awaji and S. Sato: *ASME J. Eng. Mater. Technol.*, 1979, vol. 101, pp. 139-47.
17. J.L. Tevaarwerk, A. Plumtree, and R. Sowerby: *ASME J. Eng. Mater. Technol.*, 1975, Apr., pp. 144-50.
18. G.W. Vickers, A. Plumtree, R. Sowerby, and J.L. Duncan: *ASME J. Eng. Mater. Technol.*, 1975, Apr., pp. 127-35.
19. A.M. Sabroff, F.W. Boulger, and H.J. Henning: *Forging Materials and Practices*, Rheinhold Book Corp., New York, NY, 1968.
20. P.F. Thomason: *Int. J. Mech. Sci.*, 1969, vol. 11, pp. 65-73.
21. J.R. Pickens, T.J. Langan, R.O. England, and M. Liebson: *Metall. Trans. A*, 1987, vol. 18A, pp. 303-12.
22. W.F. Hosford: *Metal Forming: Mechanics and Metallurgy*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
23. S. Kobayashi: *J. Eng. Ind.*, 1970, May, pp. 91-99.
24. J.K. Banerjee: *ASME J. Eng. Mater. Technol.*, 1985, vol. 107, pp. 138-44.
25. S.L. Semiatin and G.D. Lahoti: *Metall. Trans. A*, 1983, vol. 14A, pp. 105-15.
26. S.L. Semiatin and G.D. Lahoti: *Metall. Trans. A*, 1983, vol. 14A, pp. 743-50.
27. S.L. Semiatin and G.D. Lahoti: *Metall. Trans. A*, 1982, vol. 13A, pp. 275-88.

Forging Limits for an Aluminum Matrix Composite: Part II. Analysis

D.-G.C. SYU and A.K. GHOSH

Plasticity analysis has been carried out to calculate the forging limits of a particulate-reinforced aluminum matrix composite under various forging conditions. A geometric defect (*i.e.*, variation in cross-sectional area) that can concentrate stresses and strain and accelerate local deformation was assumed to represent all possible defects in this material. Thus, in effect, the local stress concentrations around nondeformable particles, inhomogeneous distribution of particles and grain sizes, porosity and cracked particles, *etc.*, were assumed to be simulated adequately by such a defect factor. The analysis followed a method suggested by Marciniak and Kuczynski (MK) to determine the strain path within the defect region of the composite during multiaxial deformation. A relationship of stress, strain, and strain rate obtained from the uniaxial tension test was used for the calculation of the strains. To terminate the plasticity analysis, a rate-independent fracture criterion was used that is based on Cockcroft's model of a constant work performed by the tensile component of stress. It was found that the calculated results predicted the experimental forging limits for 2014 Al/15 vol pct Al_2O_3 reasonably well. At 400 °C and a slow strain rate (0.015 s^{-1}), the predicted curve was higher than the experimental result. This was probably because the fracture mechanism, and thus fracture criterion, changed with temperature. The effect of assumed defect factor on predicted forging limits was also studied. It was found that the size of the defect factor did not significantly change the forging limits at 300 °C for strain rates from 0.015 s to 0.5 s^{-1}; however, it did have a large effect on the forging limit at 400 °C for the lower strain rate of 0.015 s^{-1}.

I. INTRODUCTION

THE experimental forging limits of 2014 Al/15 vol pct Al_2O_3 have been determined as a function of strain state in a companion article.[1] To develop a fundamental understanding of the magnitudes of the strains associated with incipient cracking, it is necessary to calculate the forging limits analytically for comparison with the experimental results. A reasonable approach for the analysis is based on the assumption that materials contain defects.[2] These defects can accelerate local deformation in the defect region, while a macroscopically homogeneous state of stress is applied to the overall sample. This can cause the growth of a defect until material separation or failure occurs in the defect region. In the composite material of interest, the defects may arise from the presence of nondeformable Al_2O_3 particles, a nonuniform distribution of the Al_2O_3 particles and matrix grain sizes, cracked Al_2O_3 particles and porosity, *etc.* For simplicity, the sample was assumed to be subjected to a macroscopically linear strain path (*i.e.*, a constant ratio of compressive to tensile strain). However, microscopically, the strain path can change due to the concentrated stresses and/or strains induced by the local defects.

Prediction of local-limit strain[3–8] in composites by finite element modeling of strain concentration at many particles, and the coalescence of voids, is prohibitively time consuming and costly. In these calculations, the results often become less accurate due to many numerical approximations. Thus, a simple method was adopted following the work of Marciniak and Kuczynski (MK)[2] which represents the defects by an equivalent geometric inhomogeneity (*i.e.*, variation in cross-sectional area). In their analysis, a groove running in a direction perpendicular to the larger principal stress (σ_1) was considered in a sheet metal subjected to biaxial tension when the ratio of the principal stresses was $0.5 \leq \sigma_2/\sigma_1 \leq 1$, as shown in Figure 1. In Figure 1, the local effective strain in the groove (region *B*) begins to concentrate gradually under biaxial tension and eventually reaches the fracture or infinite strain ratio (ratio of incremental compressive strain, $d\varepsilon_1$, to incremental tensile strain, $d\varepsilon_2$). The limiting strain state in their analysis is, then, defined as the strain state in region *A* when the local effective strain in region *B* reaches the failure criterion. They reported that the initial inhomogeneity of the material, $f = t_B/t_A$, exerts a great influence on the limiting strain. They also concluded that the inhomogeneity of the sheet metal results from other defects, such as a nonuniform distribution of impurities, varying texture, different size and orientation of grains, and can always be reduced to the equivalent geometric inhomogeneity, f.

In sheet metal-forming calculations,[9–12] primarily tensile-strain components are present. However, if this concept is applied to the upset test in bulk forging, it is necessary to consider not only tension but also compression, since this comprises the major strain component. Figure 2 shows the relationship between strain state and stress state in a plot of tensile strain (ε_2) *vs* compressive strain (ε_1) where a linear strain path is assumed. In region I, there is a biaxial tensile-stress state. The strain

D.-G.C. SYU, formerly Graduate Student, the University of Michigan, is Senior Engineer, Taipei Municipal Government, Hsintien, Taipai, Taiwan, Republic of China. A.K. GHOSH, Professor of Materials Science and Engineering, is with the University of Michigan, Ann Arbor, MI 48109-2136.

Manuscript submitted August 11, 1992.

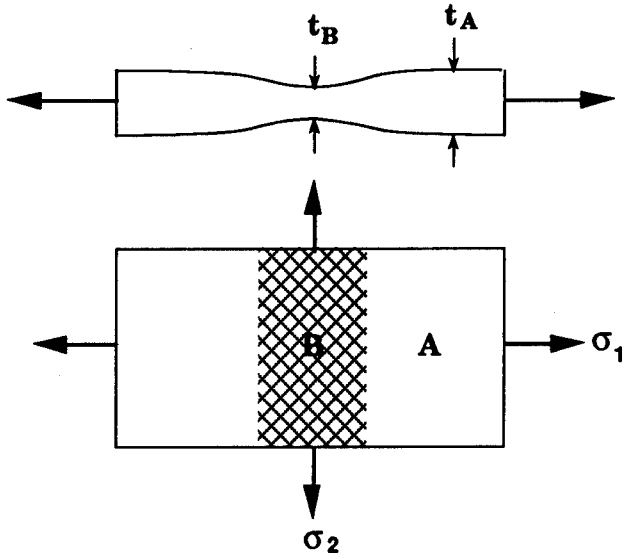


Fig. 1—Schematic diagram of the MK analysis by assuming a groove running in a direction perpendicular to the larger principal stress under biaxial tension.

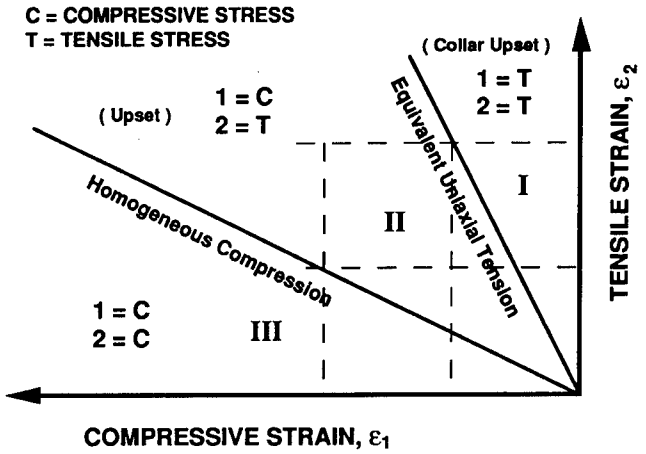


Fig. 2—Relationship between strain state and stress state in an ε_1-ε_2 diagram.

path for the collar upset test is in this region. In region II, both compressive and tensile stresses are developed. The upset test belongs in this region. In region III, the material is under a biaxial compressive-stress state, and surface fracture will never occur. A typical example of region III is the plane-strain compression test, $\varepsilon_2 = 0$.[1]

In the current investigation, a small, thin element containing defects was selected near the surface of an upset sample, as shown in Figure 3(a). These defects may include the nonuniform distribution of Al_2O_3 particles and matrix grain sizes, cracked Al_2O_3 particles and porosity, *etc*. Similar to the MK analysis, this groove in a direction perpendicular to the secondary tensile stress (the 2 direction in our calculations) was assumed to represent the largest of all the defects in the composite. This groove can be expressed by a defect factor, F ($F = 1 - t_B/t_A = 1 - f$), shown in Figure 3(b), and will induce the strain-path change in region B and early fracture, as

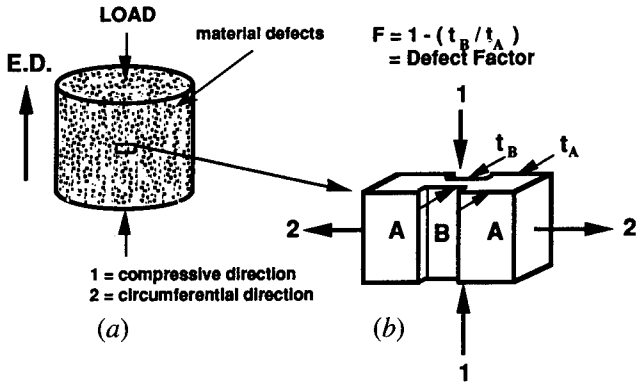


Fig. 3—(*a*) Defects in a small element from an upset sample reduced to an equivalent geometric defect factor, F, as seen in (*b*).

described in the MK analysis. For example, if compression along the 1 direction at a constant strain rate is applied in an upset test, the secondary tensile stress along the 2 direction in region B would be higher than that in region A due to the force balance. Consequently the strain rate along the 2 direction in region B would be higher than that in region A, and the strain in this direction (ε_2) in region B would reach the fracture criterion or infinite strain ratio prior to region A, as shown in Figure 4. A similar method was also reported by Kuhn *et al.*[13] and Lee and Kuhn.[14]

Satisfaction of force equilibrium in two sections of the specimen requires the knowledge of material constitutive relation, *i.e.*, a relationship among stress, strain, and strain rate of the composite. This was determined experimentally. Using this relationship, the instantaneous cross-sectional areas, and instantaneous stresses in regions A and B, the forces in the 2 direction can be balanced[21] by using the following equations based on the von Mises yield criterion for effective stress and strain for plane stress ($\sigma_3 = 0$). The following relations are used for proportional straining.

$$\alpha = \frac{(2\rho + 1)}{(2 + \rho)} \qquad [1]$$

$$\sigma_2 = \sigma (1 - \alpha + \alpha^2)^{-0.5} \qquad [2]$$

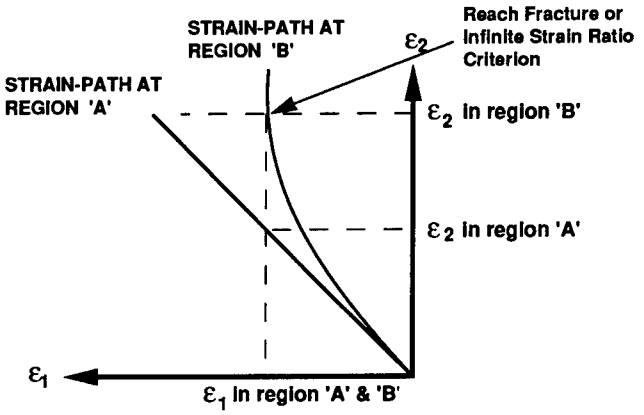


Fig. 4—Strain-path deviation based on the MK analysis and applied to the upset tests. The strain path in region B deviated from that in region A due to the defect factor.

$$\sigma_{2A} \times \exp(-(\varepsilon_{1A} + \varepsilon_{2A})) = \sigma_{2B} \times (1 - F) \times \exp(-(\varepsilon_{2B} + \varepsilon_{2B})) \quad [3]$$

where α ($= \sigma_1/\sigma_2$) and ρ ($= d\varepsilon_1/d\varepsilon_2$) are the stress and strain ratios, respectively. The term σ is the effective stress from the tensile test. The term σ_2 is the stress in the 2 direction, *i.e.*, σ_{2A} and σ_{2B} are the stresses in the 2 direction in regions A and B, respectively. The values of ε_{1A} and ε_{2A} are the true strains in the 1 and 2 directions, respectively, in region A. The values of ε_{1B} and ε_{2B} are the true strains in the 1 and 2 directions, respectively, in region B. The term F is the defect factor.

Failure can occur by localized necking or the growth of internal voids followed by separation. When the fracture strain of the material is extremely high, the limiting strain is achieved when the strain ratio ($d\varepsilon_1/d\varepsilon_2$) reaches an infinite value. However, in upset tests of engineering materials, fracture generally occurs long before necking, and an infinite strain ratio is not attained. Hence, it is necessary to introduce an appropriate fracture criterion to terminate the plasticity calculation. The accumulated strain outside the defect at this point, then, is the predicted forging limit.

Several fracture criteria for ductile fracture have been proposed in the literature, such as fracture due to the deformation work induced by the tensile component,[15] due to the stress state,[16] and due to the growth and coalescence of cylindrical holes,[17] *etc.* The application of these fracture criteria has been reported by many researchers.[13–20] In the present case, three fracture criteria were examined. They are the Cockcroft fracture criterion,[15] a fracture-stress-based criterion,[16] and a constant effective fracture-strain criterion.

Cockcroft and Latham[15] proposed that ductile fracture will take place when the amount of deformation work due to the maximum tensile stress reaches a critical value C^*. In the present case, the fracture criterion was only considered in region B, since this region is closer to fracture due to the presence of the defect. This criterion can be written as

$$\int_0^{\varepsilon_{\text{eff}}} \left(\frac{\sigma_{2B}}{\sigma_{\text{eff}}}\right) \sigma_{\text{eff}}\, d\varepsilon_{\text{eff}} > C^*, \quad \text{where } C^* = \int_0^{\varepsilon_f} \sigma d\varepsilon \quad [4]$$

where σ_{2B} is the maximum tensile stress in region B. σ_{eff} and ε_{eff} are the effective stress and strain, respectively, in region B. σ, ε, and ε_f are the equivalent stress, strain, and fracture strain, respectively, from the uniaxial tension test. C^* is the total deformation work that the material can suffer without fracture.

Other possible fracture criteria include the constant effective fracture-strain- and fracture-stress-based criteria that are discussed in the Appendix. The constant effective fracture-strain criterion states that fracture will occur when the effective strain in region B reaches a constant effective fracture strain.

The fracture-stress-based criterion[16] is based on the assumption that the growth rate of cavitation is proportional to the mean stress, and the shear strain rate is proportional to the maximum shear stress. The criterion proposes that fracture will occur in region B when the product of the mean stress (initiation and growth of voids) and the maximum shear stress (coalescence of voids) is greater than or equal to the square of the effective fracture stress from the uniaxial tension test divided by 6.

II. EXPERIMENTAL PROCEDURES AND RESULTS

Experimentally, the forging limit is determined at the onset of an incipient crack on the upset sample.[1] However, for the calculations, the experimentally determined true strain (from area reduction) at the incipient crack in a *tensile specimen* was used as the termination point for the forging-limit prediction. Figure 5 shows a part of the fractured tensile specimen where an incipient crack was seen at the edge of the sample and far from the fracture surface. The width and thickness at this location were then measured by a micrometer, and their product, A_f, was the cross-sectional area at the onset of incipient fracture. The incipient fracture strain, ε_f, is defined as $-\ln(A_f/A_0)$, where A_0 is the starting cross-sectional area (also measured by a micrometer). If there was more than one incipient crack on the tensile sample, the farthest one from the fracture surface was measured to determine the A_f.

To correlate with the previous forging tests that showed cracks parallel to the compression axis (which is also the extrusion direction), as shown in Figure 6, uniaxial-tension tests were conducted along the transverse direction (T direction) to obtain the constitutive equation of the composite. This constitutive equation was then used as the input data for the calculations.

The composite material, 2014 Al/15 vol pct Al_2O_3, used in this investigation was manufactured* by dispersing aluminum oxide particles into a molten aluminum alloy that was then cast and extruded into a rod with a diameter of 50.8 mm. The composition, microstructures, and mechanical properties of this composite

*This composite material was manufactured by Duralcan Aluminum Composites Corporation, San Diego, CA.

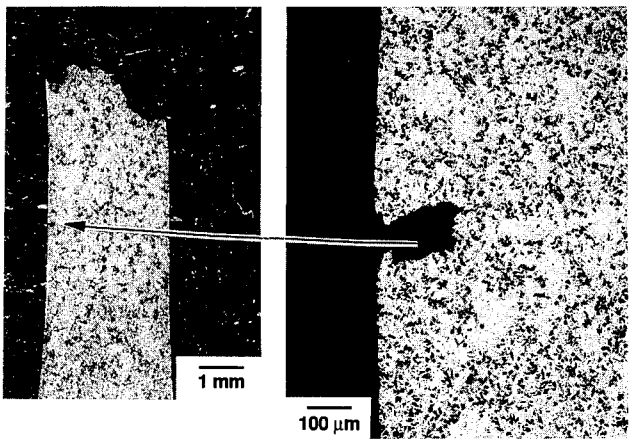


Fig. 5—Incipient cracking occurred far from fracture surface under uniaxial-tension test for 2014 Al/15 vol pct Al_2O_3. Test condition: T direction at 300 °C and a strain rate of 0.5 s^{-1}.

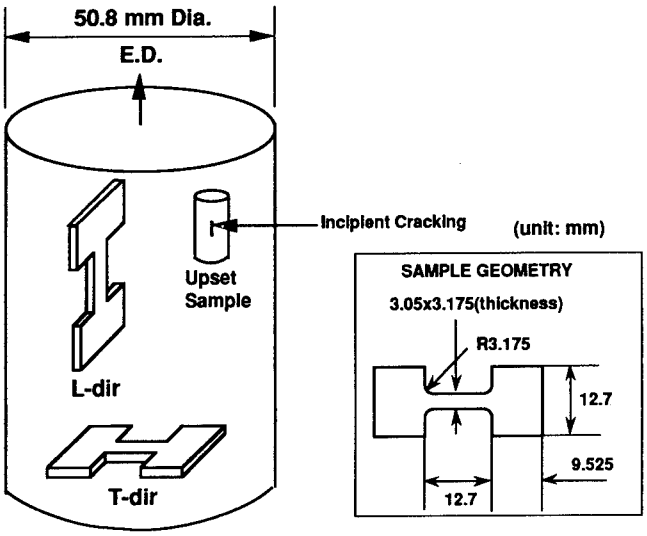


Fig. 6—Sample dimensions and testing directions for uniaxial-tension test.

are as discussed in Reference 1. The average particle size of our composite was about 5.6 μm, and the subgrain size was about 2 μm.[1] A 2014 Al matrix alloy that was used for comparison was produced *via* ingot metallurgy and extrusion and had a grain size of 500 to 2000 μm.

A. *Determination of the Relation for Stress, Strain, and Strain Rate*

Samples were tested under uniaxial tension along the T direction to obtain the σ-ε curves for various strain rates. To compare calculated forging limits with the previous experimental ones, 300 °C and 400 °C were used as the testing temperatures. The sample geometry is shown in Figure 6, which had a gauge length of 11.11 mm. An Instron universal testing machine, model 4505, was used for tensile testing of the test materials to obtain the relationship between load and displacement. The σ-ε curves were calculated from these in terms of plastic strain (by subtracting the elastic displacements). Figures 7(a) and (b) show the σ-ε curves at 300 °C and 400 °C, respectively, for various strain rates. From the tested specimens, the overall fracture strains were determined from $-\ln (A_f/A_0)$, where A_f is the area of the fracture surface and was determined by measuring the width and thickness of the fracture profile under an optical microscope. Since true σ-ε curves cannot be obtained beyond the load maximum point, the σ-ε curves in Figures 8(a) and (b) were extended by the best-fit polynomial curve at 300 °C and 400 °C, respectively, to cover the fracture strain for calculation. The σ-ε curves along the L direction were also shown in Figures 7(a) and (b) to study the high-temperature anisotropic behavior of the composite, which will be discussed in Section III.

From the σ-ε curves at different strain rates, a best-fit constitutive equation was determined for each temperature. First, for all σ-ε curves, the best-fit polynomial equation for (log σ) was expressed up to the second

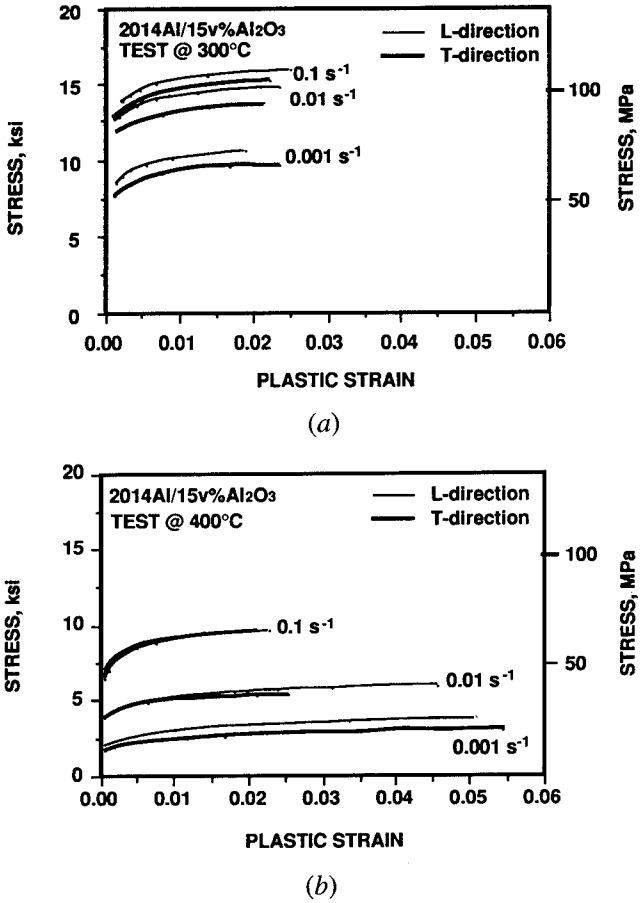


Fig. 7—σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under uniaxial-tension test in longitudinal (L) and transverse (T) directions at (*a*) 300 °C and (*b*) 400 °C. It is seen that the anisotropy becomes less significant at higher temperature.

order of (log ε) at each strain rate: 0.1, 0.01, and 0.001 s^{-1},

$$\log \sigma = c_0 + c_1 (\log \varepsilon) + c_2 (\log \varepsilon)^2 \qquad [5]$$

where c_i (i = 0, 1, and 2) is a function of strain rate and can be expressed by the best-fit polynomial equation up to the second order of (log $\dot{\varepsilon}$), as follows:

$$c_0 = a_0 + a_1 (\log \dot{\varepsilon}) + a_2 (\log \dot{\varepsilon})^2 \qquad [6a]$$

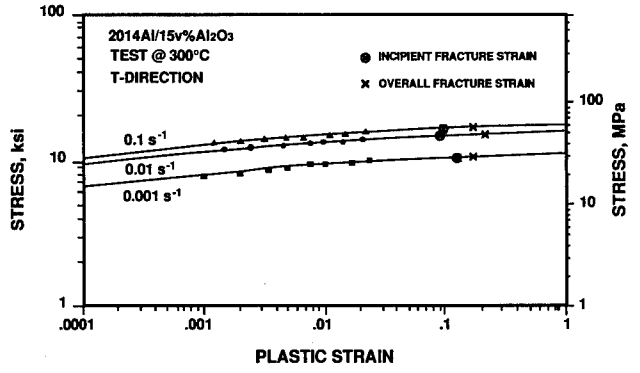
$$c_1 = b_0 + b_1 (\log \dot{\varepsilon}) + b_2 (\log \dot{\varepsilon})^2 \qquad [6b]$$

$$c_2 = d_0 + d_1 (\log \dot{\varepsilon}) + d_2 (\log \dot{\varepsilon})^2 \qquad [6c]$$

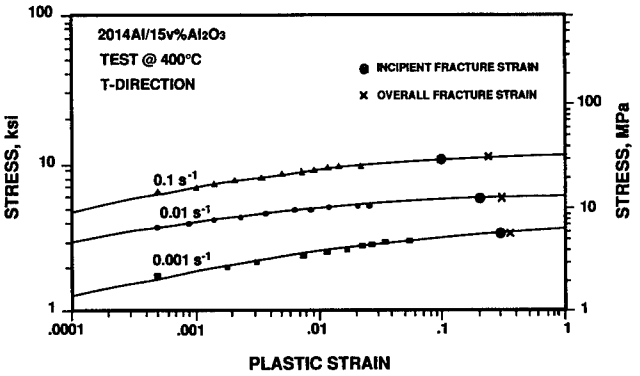
where $\dot{\varepsilon}$ is the strain rate, a_i, b_i, and d_i (i = 0, 1, 2) are constants as shown in Table I. Therefore, the constitutive Eqs. [5] and [6] of the composite at 300 °C and 400 °C were determined and used as the input data to calculate the strain-path change in region B.

B. *Strain-Path Calculation*

To calculate the strain path within region B, an incremental compressive strain ($\Delta\varepsilon_1$), incremental time, Δt, and strain ratio, ρ (= $\Delta\varepsilon_1/\Delta\varepsilon_2$), were also given as the input parameters. A flowchart for the computer calculations is shown in Figure 9. The incremental compressive strains in regions A and B were given the same value



(*a*)



(*b*)

Fig. 8—Extended σ-ε curves of 2014 Al/15 vol pct Al_2O_3 at (*a*) 300 °C and (*b*) 400 °C. The points on the curves were calculated from testing charts. The last data point in each curve indicates the point of load instability. Curves between the instability and the fracture strain were interpolated by the best-fit equation.

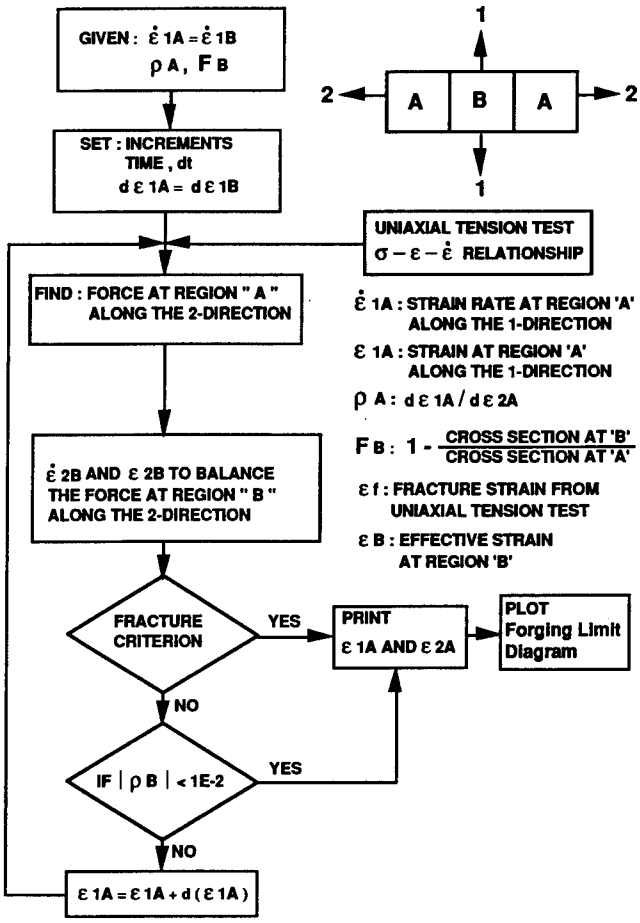
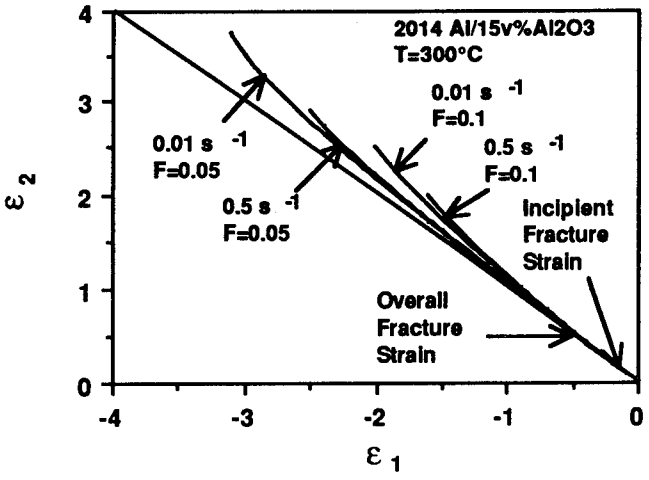


Fig. 9—Computer flowchart for the calculation of forging limits using the defect model and fracture criterion.

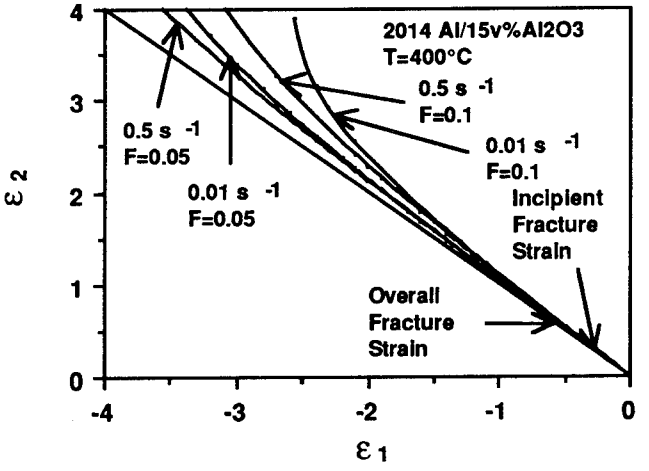
($\varepsilon_{1A} = \varepsilon_{1B}$) throughout the calculations, since the samples were compressed with the same axial strain during upset testing. Using Eqs. [1] and [2] for a given ρ and ε_{1A}, the stress (σ_{2A}) and strain (ε_{2A}) in region A along the 2 direction could be calculated at various time increments. For a given value of the defect factor, F, using Eq. [3], the corresponding ε_{1B} and σ_{2B} were then determined by an iterative numerical procedure. Since region B is weaker than region A, the effective strain in region B will reach the fracture criterion faster than that in region A. Once the effective strain in region B satisfied the fracture criterion (Eq. [1]), the calculations were stopped, and the corresponding strain state in region A (ε_{1A} and ε_{2A}) was plotted in an ε_1-ε_2 diagram and indicated as the forging limit. By performing similar calculations for several values of ρ for a given F, the forging-limit line can be drawn as a locus of the strain combinations in region A at failure. The strain-path deviation due to the defect factor, F, is shown in Figures 10(a) and (b) at 300 °C and 400 °C, respectively, without using the fracture criterion. In these figures, the strain path in region A was a straight line with a strain ratio ($\Delta\varepsilon_1/\Delta\varepsilon_2$) of 1. It is clear that when fracture criterion is applied, the limit strains would become lower.

Table I. Constants for Constructing the Relationship of Stress, Plastic Strain, and Strain Rate in 2014 Al/15 vol pct Al_2O_3

		c_0			c_1			c_2		
		a_0 (MPa)	a_1 (MPa · s)	a_2 (MPa · s^2)	b_0 (MPa)	b_1 (MPa · s)	b_2 (MPa · s^2)	d_0 (MPa)	d_1 (MPa · s)	d_2 (MPa · s^2)
300 °C	L-DIR	1.8837	−0.1902	−0.0621	−0.12219	−0.070913	−0.012145	−0.048046	−0.024557	−0.005296
	T-DIR	2.0484	−0.11365	−0.05375	0.012524	−0.044846	−0.013338	−0.021024	−0.023738	−0.0068877
400 °C	L-DIR	2.2013	0.34035	0.040995	0.07494	0.074197	0.025536	0.000484	0.017852	0.0041065
	T-DIR	2.3640	0.51709	0.069545	0.1622	0.18947	0.05338	−0.008604	0.012079	0.003521



(*a*)

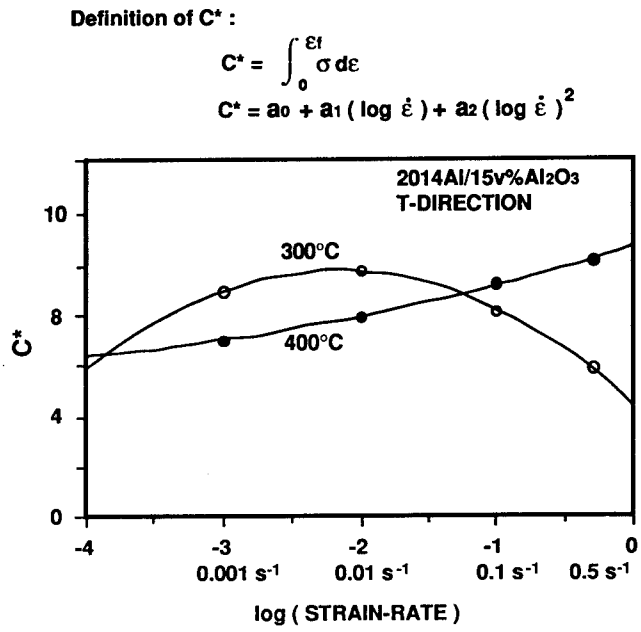


(*b*)

Fig. 10—Calculated strain-path deviation from a strain ratio ($d\varepsilon_1/d\varepsilon_2$) of −1 due to defect factors, F, for 2014 Al/15 vol pct Al_2O_3 at (*a*) 300 °C and (*b*) 400 °C. Sample will fracture before the strain ratio reaches $-\infty$.

C. *Fracture Criterion*

The forging-limit calculation for the present composite is dominated by the fracture criterion rather than flow localization. To use the Cockcroft fracture criterion, the value of C^* at a given strain rate was determined by integrating the area under the stress-strain curve up to the incipient fracture strain (as indicated on the plots in Figure 8). The relationships between C^* and strain rate at 300 °C and 400 °C were thus determined, as shown in Figure 11. The best-fit equation was generated to match the curve for each temperature. These equations and their constants were the input data for the computer calculations. Consequently, the calculated forging limits under various conditions were determined using this fracture criterion, as seen in Figures 12 through 15, where various defect factors (F) have been chosen to match the experimental results. In Figures 12 through 15, the strain paths (dashed lines) in region A were assumed to be linear. The calculated point of each strain path that intercepts the $F = 0$ forging-limit line (*i.e.*, defect absent) indicates failure by directly satisfying the



MATERIAL : 2014 Al / 15v% Al_2O_3 (Unit: MPa)

T-DIR	a_0	a_1	a_2
300°C	4.429	-4.846	-1.123
400°C	10.58	1.667	0.152

Fig. 11—A plot showing C^* *vs* strain rate. The value of C^* was based on the incipient-fracture strain using the Cockcroft criterion. The best-fit equation and its constants for each temperature were chosen for computer calculations, as seen in Table I.

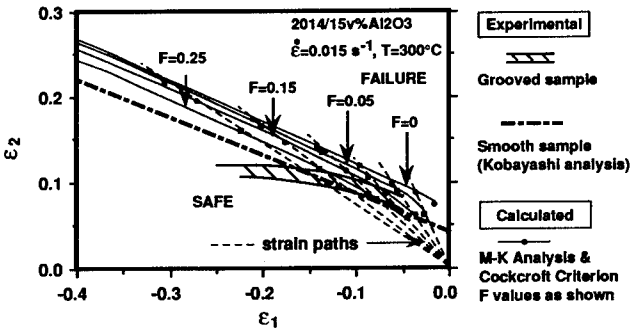


Fig. 12—Comparison of calculated and experimental forging-limit diagrams for 2014 Al/15 vol pct Al_2O_3 at 300 °C and an effective strain rate of 0.015 s^{-1}.

fracture criterion. The calculated points on the various strain paths that intercept the $F > 0$ forging-limit line (*i.e.*, defect present) indicate the strain states in region A when the corresponding strain state in region B satisfies the fracture criterion.

Two other fracture criteria, the fracture-stress-based and the constant effective fracture-strain criteria, were also examined. These are discussed in the Appendix. It was found that when the fracture-stress-based criterion was used, the slope of the calculated forging-limit line was about −1.5 for $F = 0.05$ at 300 °C and a strain

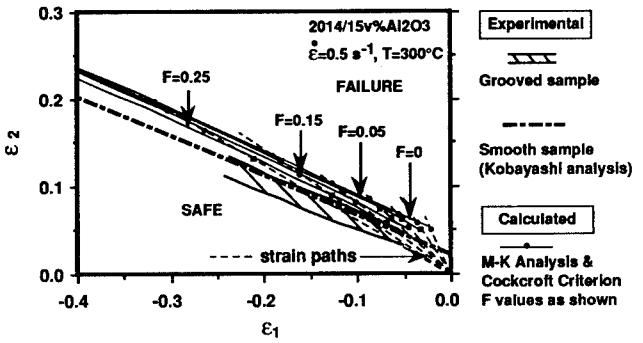


Fig. 13—Comparison of calculated and experimental forging-limit diagrams for 2014 Al/15 vol pct Al_2O_3 at 300 °C and an effective strain rate of 0.5 s^{-1}.

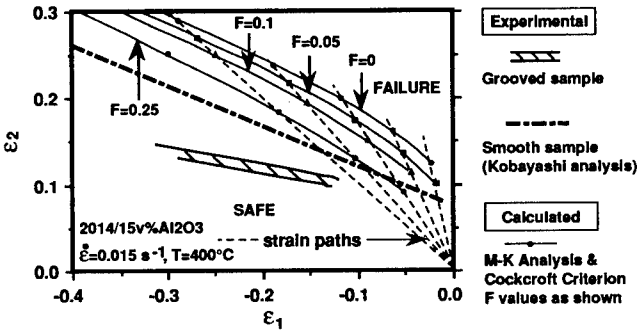


Fig. 14—Comparison of calculated and experimental forging-limit diagrams for 2014 Al/15 vol pct Al_2O_3 at 400 °C and an effective strain rate of 0.015 s^{-1}.

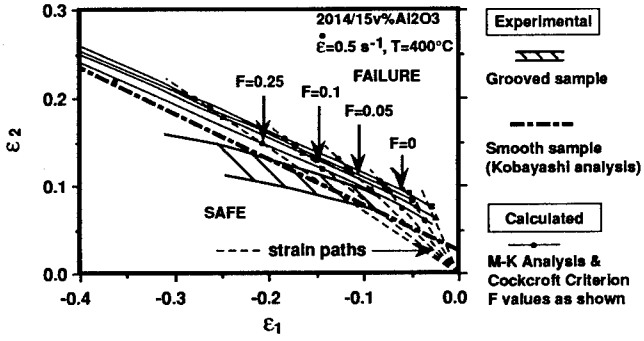


Fig. 15—Comparison of calculated and experimental forging-limit diagrams for 2014 Al/15 vol pct Al_2O_3 at 400 °C and an effective strain rate of 0.5 s^{-1}.

rate of 0.015 s^{-1} and much smaller than the experimental one (about −0.3).[1] When the constant effective fracture-strain criterion and the same forging condition were used, it was found that the predicted data decreased after some strain level ($\varepsilon_1 \sim 0.025$), which did not follow the experimental trend. Thus, the Cockcroft criterion provides a better match than the other criteria. This implies that the forging limits for 2014 Al/15 vol pct Al_2O_3 were dominated by the deformation work performed by the tensile component during the upset tests. For this reason, the calculated and experimental forging limits based on the Cockcroft fracture criterion will be discussed in Section III.

III. DISCUSSION

In comparing the experimental forging limits with the calculated forging limits in Figures 12 through 15, it is seen that at 300 °C with strain rates of 0.015 and 0.5 s^{-1} and at 400 °C, with a strain rate of 0.5 s^{-1}, the predictions were reasonably good with F values of about 0.25. Such a high value of F may be justified in a composite material that contains as much as 20 pct nondeformable and jagged particles as stress raisers. The values of the defect factors did not significantly change the calculated forging limits. However, at 400 °C and slow strain rate (0.015 s^{-1}, Figure 14), the defect factor did change the calculated forging limits, and an F value of 0.25 or greater was needed to match the experimental results. This is because diffusion creep and grain boundary sliding are active at slow strain rates. It is expected that the fracture mechanism would change as the deformation mechanism is changed.[22] Therefore, it may not be appropriate to use the Cockcroft criterion at 400 °C and a slow strain rate (0.015 s^{-1}). Moreover, the Cockcroft fracture criterion does not consider diffusion creep, grain-boundary sliding, void initiation, and the interaction between particles and the matrix, all of which play an important role during fracture in an alloy-containing inclusions.[23–32]

Examination of Eq. [4] shows that if there is no tensile stress operating but only a compressive stress, $\sigma_{2B} = 0$, fracture does not occur. A typical example is the plane-strain compression test,[1] *i.e.*, $\varepsilon_2 = 0$ in Figure 4.

In Figures 7(a) and (b) it is seen that in 2014 Al/15 vol pct Al_2O_3 the flow stresses were higher along the L direction than along the T direction at 300 °C and 400 °C, respectively. The difference in flow stress between the two testing directions is possibly due to the alignment of the Al_2O_3 particles along the extrusion direction. This aligned orientation combined with their aspect ratio (about 2:1) and possibly a crystallographic texture in the matrix produced during extrusion are responsible for the difference between the L and T directions. It is also seen in Figures 7(a) and (b) that the anisotropic behavior became less significant with increasing temperature. One of the possible reasons is that the particle-strengthening effect decreased with increasing temperature. To verify this hypothesis, tensile tests were conducted on the matrix material and on the composite at 300 °C, 400 °C, and 500 °C with a strain rate of 0.1 s^{-1}. The σ *vs* ε curves are shown in Figure 16. From this data, yield stress at 0.2 pct offset plastic strain (plotted in Figure 17 for both materials as a function of temperature). It was found that these two curves intersected at about 410 °C, indicating that for a strain rate of 0.1 s^{-1}, the Al_2O_3 particles strengthen the matrix below 410 °C, and weaken the matrix material above 410 °C. The same observation was also reported by Pickens *et al.* in 6061 Al/20 vol pct SiC and 7090 Al/20 vol pct SiC during a hot-torsion test.[33] The reason for this weakening is probably a greater contribution from interface sliding in the composite at the very high temperatures.[34,35] Since the particle-strengthening effect was reduced above 400 °C, the anisotropic behavior due to particle clusters in the composite also became less significant above 400 °C.

IV. SUMMARY AND CONCLUSIONS

Analytical calculations for forging limits have been carried out for a particulate-reinforced composite by assuming that a material geometrical defect can concentrate strains locally and grow to lead to fracture. This defect factor, which in this case arises from strain concentration between nondeformable particles and agglomerates, was found to be around 0.25. This value is considerably larger than that for unreinforced metal matrices. Simulations showed the strain-path deviation within the defect region (as suggested by MK)[2] during multiaxial deformation. To terminate the plasticity analysis, Cockcroft's fracture criterion based on constant work performed by the tensile component of stress was used and was found to be the most suitable criterion.

1. The present analysis can provide a reasonably good match to the experimental forging limits for the composite at 300 °C with strain rates of 0.015 s^{-1} and 0.5 s^{-1} at 400 °C with a strain rate of 0.5 s^{-1}. However, at 400 °C with a slower strain rate (0.015 s^{-1}), the prediction was higher than the experimental result, probably because the fracture mechanism changed with temperature, such that Cockcroft criterion was no longer valid at these temperatures.
2. The defect factor did not significantly change the forging limits at 300 °C for strain rates from 0.015 to 0.5 s^{-1} for the composite. However, the defect factor produced a large change in the calculated forging limit at 400 °C for a slow strain rate (0.015 s^{-1}).
3. The flow stresses along the extrusion direction were higher than along the transverse direction at 300 °C and 400 °C in 2014 Al/15 vol pct Al_2O_3. This anisotropic behavior became less significant at higher temperatures (above 400 °C), because the particle-strengthening effect was reduced.

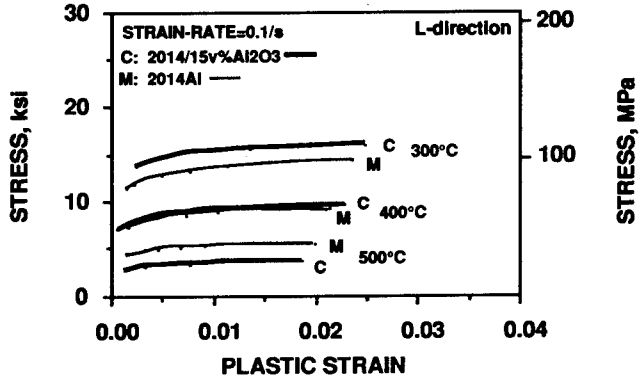


Fig. 16—The σ-ε curves of 2014 Al matrix (M) and its composite (C) at various temperatures under uniaxial tension along L direction. At 500 °C the composite was weaker than the matrix material.

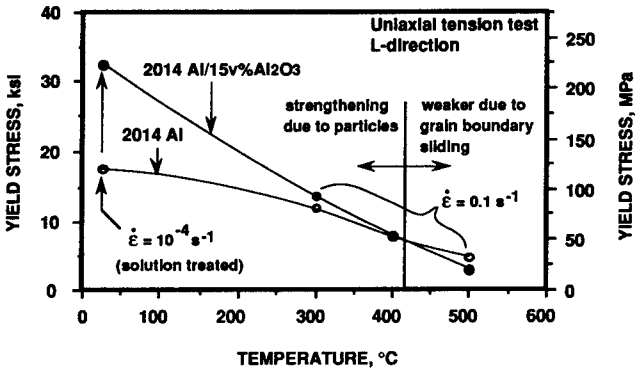


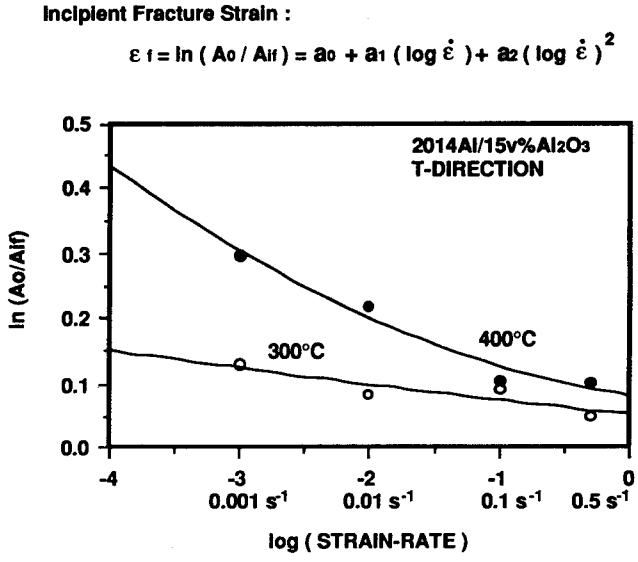
Fig. 17—Yield stress for 0.2 pct offset plastic strain *vs* temperature for 2014 Al/15 vol pct Al_2O_3 and 2014 Al. Above 410 °C the yield stress of the composite became lower than that of the matrix material.

APPENDIX

Fracture Criteria

A. *Constant Effective Fracture-Strain Criterion*

This fracture criterion states that fracture will occur when the effective strain in region *B* (Figure 1(b)) reaches a constant effective fracture strain. The relationships between incipient-fracture strain (ε_f) and strain rate at 300 °C and 400 °C were determined from the σ-ε curves for the calculation, as shown in Figure A1. The best-fit equation was generated to match the curve for each temperature. These equations and their constants were the input data for the computer calculations. Figure A2 shows the result obtained using this fracture-strain criterion at 300 °C with a strain rate of 0.015 s^{-1}. Apparently, this criterion did not match the experimental result.



MATERIAL : 2014Al/15v%Al2O3

T-DIR	a_0	a_1	a_2
300°C	0.051494	-0.022036	0.00070802
400°C	0.082146	-0.03061	0.014275

Fig. A1—A plot showing incipient-fracture strain *vs* strain rate. The best-fit equation and its constants for each temperature were chosen for computer calculations, as seen in Table I.

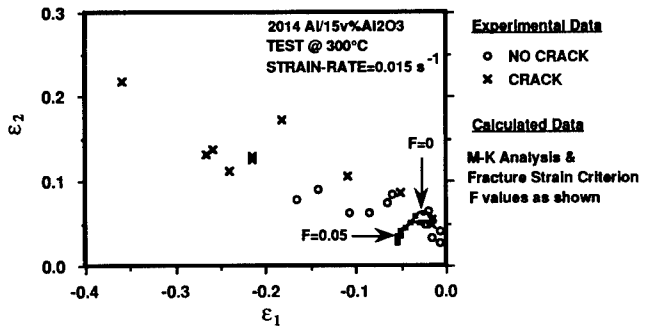
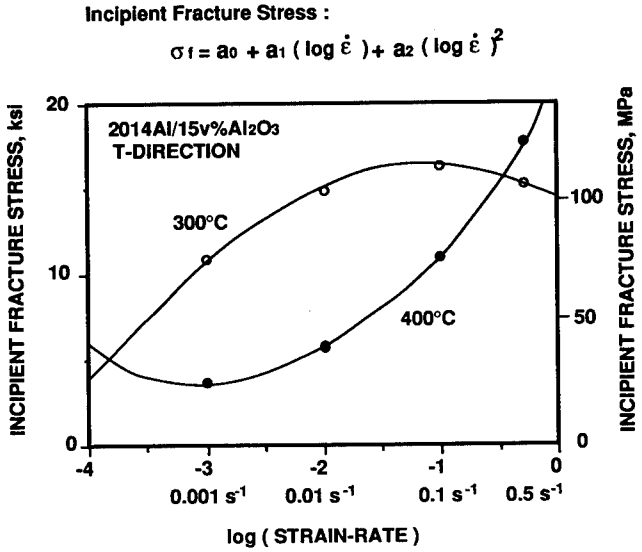


Fig. A2—Prediction of forging-limit diagram based on the MK analysis and effective-fracture-strain criterion not matching the experimental result for 2014 Al/15 vol pct Al_2O_3.



MATERIAL : 2014Al/15v%Al2O3 (Unit: MPa)

T-DIR	a0	a1	a2
300°C	102.14	-21.66	-10.22
400°C	146.43	81.22	13.67

Fig. A3—A plot showing incipient-fracture stress *vs* strain rate. The best-fit equation and its constants for each temperature were chosen for computer calculations, as seen in Table I.

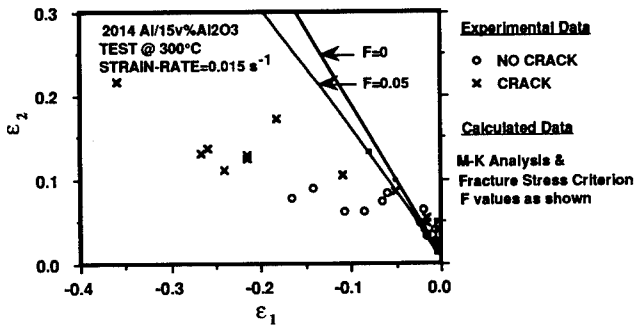


Fig. A4—Prediction of forging-limit diagram based on the MK analysis and fracture stress-based criterion not matching the experimental result for 2014 Al/15 vol pct Al_2O_3.

B. *Fracture-Stress-Based Criterion*

This criterion is based on the assumption that the growth rate of cavitation is proportional to the mean stress, and the shear strain rate is proportional to the maximum shear stress. The criterion proposes that fracture will occur when the product of the mean stress with the maximum shear stress is greater than or equal to the square of effective fracture stress (s_f, from the uniaxial-tension test) divided by 6. That is,

$$((\sigma_1 + \sigma_2 + \sigma_3)/3) \times ((\sigma_1 - \sigma_3)/2) \geq (\sigma_f/3) \times (\sigma_f/2) = (\sigma_f^2)/6 \quad [A1]$$

where $\sigma_1 > \sigma_2 > \sigma_3$. In the present case, for $\sigma_3 = 0$, the above criterion can be rewritten as

$$(\sigma_{1B} + \sigma_{2B}) \times \sigma_{max} \geq \sigma_f^2 \quad [A2]$$

where σ_{1B} and σ_{2B} are the stresses along 1 direction and 2 direction, respectively, in region B, in Figure 3, and σ_{max} is the maximum value between $(\sigma_{1B} - \sigma_{2B})$, σ_{1B}, and σ_{2B}.

The relationships between incipient fracture stress (σ_f) and strain rate at 300 °C and 400 °C were determined from the σ-ε curves for the calculation, as shown in Figure A3. The best-fit equation was generated to match the curve for each temperature. These equations and their constants were the input data for the computer calculations. Figure A4 shows the result obtained using this fracture-stress-based criterion at 300 °C with a strain rate of 0.015 s^{-1}, and it did not match the experimental result.

ACKNOWLEDGMENTS

The authors would like to thank the AES, GM Technical Center, for supporting this work. Fruitful discussions with Dr. B. Taylor of GM-AES are gratefully acknowledged.

REFERENCES

1. D.-G.C. Syu and A.K. Ghosh: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 2027-38.
2. Z. Marciniak and K. Kuczynski: *Int. J. Mech. Sci.*, 1967, vol. 9, pp. 609-20.
3. T. Christman, A. Needleman, S. Nutt, and S. Suresh: *Mater. Sci. Eng.*, 1989, vol. A107, pp. 49-61.
4. T. Christman, A. Needleman, and S. Suresh: *Acta Metall.*, 1989, vol. 37 (11), pp. 302-50.
5. J.R. Brockenbrough, S. Suresh, and H.A. Wienecke: *Acta Metall.*, 1991, vol. 39 (5), pp. 735-52.
6. D.B. Zahl and R.M. McMeeking: *Acta Metall.*, 1991, vol. 39 (6), pp. 1117-22.
7. J. Llorca, S. Suresh, and A. Needleman: *Metall. Trans. A*, 1992, vol. 23A, pp. 919-34.
8. T.L. Dragone and W.D. Nix: *Acta Metall.*, 1990, vol. 38 (10), pp. 1941-53.
9. A.K. Ghosh: *ASME J. Eng. Mater. Technol.*, 1977, July, pp. 264-74.
10. R. Sowerby and J.L. Duncan: *Int. J. Mech. Sci.*, 1971, vol. 13, pp. 217-29.
11. N.-M. Wang: in *Numerical Analysis of Forming Processes*, J.F.T. Pittman, O.C. Zienkiewicz, R.D. Wood, and J.M. Alexander, eds., John Wiley & Sons Ltd., New York, NY, 1984, pp. 117-64.

12. A.R. Ragab and A.T. Abbas: *ASME J. Eng. Mater. Technol.*, 1986, vol. 108, pp. 250-57.
13. H.A. Kuhn, P.W. Lee, and T. Erturk: *ASME J. Eng. Mater. Technol.*, 1973, Oct., pp. 213-18.
14. P.W. Lee and H.A. Kuhn: *Metall. Trans. A*, 1973, vol. 4A, pp. 969-74.
15. M.G. Cockcroft and D.J. Latham: *J. Inst. Met.*, 1968, vol. 96, pp. 33-39.
16. V. Vujovic and A.H. Shabaik: *ASME J. Eng. Mater. Technol.*, 1986, vol. 108, pp. 245-49.
17. F.A. McClintock: *ASME J. Appl. Mech.*, 1968, June, pp. 363-71.
18. Gy. Ziaja, L. Cser, and Z. Darvas: in *Advanced Technology of Plasticity 1987*, K. Lange, ed., Springer-Verlag Ltd., Berlin, vol. 2, pp. 999-1003.
19. N.L. Dung and O. Mahrenholtz: in *Advanced Technology of Plasticity 1987*, K. Lange, ed., Springer-Verlag Ltd., Berlin, vol. 2, pp. 1013-20.
20. R. Sowerby, I. O'Reilly, N. Chandrasekaran, and N.L. Dung: *ASME J. Eng. Mater. Technol.*, 1984, vol. 106, pp. 101-06.
21. W.F. Hosford: *Metal Forming: Mechanics and Metallurgy*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
22. M.F. Ashby, C. Gandhi, and D.M.R. Taplin: *Acta Metall.*, 1979, vol. 27, pp. 699-729.
23. R. Becker, A. Needleman, S. Suresh, V. Tvergaard, and A.K. Vasudevan: *Acta Metall.*, 1989, vol. 37 (1), pp. 99-120.
24. J.R. Fisher and J. Gurland: *Met. Sci.*, 1981, May, pp. 185-92.
25. J.R. Fisher and J. Gurland: *Met. Sci.*, 1981, May, pp. 1193-202.
26. J. Gurland and J. Plateau: *Trans. ASM*, 1963, vol. 56, pp. 442-54.
27. A. Needleman: *ASME J. Appl. Mech.*, 1987, vol. 54, pp. 525-31.
28. F.J. Humphreys, W.S. Miller, and M.R. Djazeb: *Mater. Sci. Technol.*, 1990, vol. 6, pp. 1157-66.
29. J.C. Earthman and W.D. Nix: *Acta Metall.*, 1987, vol. 35 (7), pp. 1475-85.
30. J.S. Wang and W.D. Nix: *Acta Metall.*, 1986, vol. 34 (3), pp. 545-55.
31. D.S. Wilkinson: *Acta Metall.*, 1987, vol. 35 (6), pp. 1251-59.
32. D.S. Wilkinson: *Acta Metall.*, 1987, vol. 35 (11), pp. 2791-9.
33. J.R. Pickens, T.J. Langan, R.O. England, and M. Liebson: *Metall. Trans. A*, 1987, vol. 18A, pp. 303-12.
34. T.G. Nieh, K. Xia, and T.G. Langdon: *ASME J. Eng. Mater. Technol.*, 1988, vol. 110, pp. 77-82.
35. M.W. Mahoney and A.K. Ghosh: *Metall. Trans. A*, 1987, vol. 18A, pp. 653-61.

Stress-State Dependence of Strain-Hardening Behavior in 2014 Al/15 Vol Pct Al_2O_3 Composite

D.-G.C. SYU and A.K. GHOSH

The purpose of this investigation was to examine if the effective stress-strain function for discontinuously reinforced aluminum (DRA) matrix composites is independent of stress state, as they are for aluminum alloys. The rationale for such work is provided by the need to develop constitutive equations for applications in metal forming and forging problems. Experimental effective σ-ε curves at room temperature were determined for a particulate-reinforced composite, 2014 Al/15 vol pct Al_2O_3, and the matrix material, 2014 Al, under a variety of stress states. The tests consisted of uniaxial tension, equibiaxial tension (bulge test), and compression tests. To eliminate the effects of prior precipitation, all samples were given a solution-heat-treatment prior to tests. It was found that for the composite the effective yield stress in uniaxial tension was higher than that in equibiaxial tension but slightly lower than that in compression. However, the effective yield stresses for the matrix material in uniaxial tension and equibiaxial tension were nearly the same. The strain-hardening rate of the composite under equibiaxial tension was higher than that under either uniaxial tension or compression. It is suggested that nondeformable dead zones can develop around the particles during deformation whose shape changes with the applied stress state, and this is partly responsible for the observed differences in behavior.

I. INTRODUCTION

FOR metallic materials, generally the strain-hardening curves are assumed to be given by a unique effective stress-strain function independent of the stress state. For isotropic materials, the effective stress and strain are assumed to follow von Mises criterion (*i.e.*, the effective yield stresses under uniaxial tension, equibiaxial tension, and compression are the same). It has been found that the effective stress-strain (σ-ε) curves in aluminum alloys, copper, and brass under these states of stress are not significantly different.[1,2,3] However, it is questionable if the above argument is true for a material reinforced with a high-volume fraction of nondeformable particles, particularly because nonuniform deformation can occur around the particles. For applications in metal forming and forging, it is essential to determine if there is a variation in the stress-strain relationships as a function of stress state.

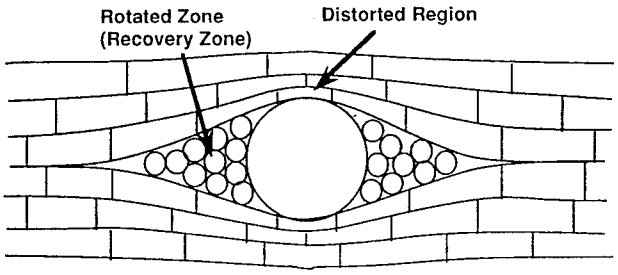
It is clear that the reinforcements can disturb the plastic flow in the matrix. For instance, the flow pattern could be a lensoid distortion of the slip and substructure development around the particles during rolling,[4] as shown in Figure 1. Although the microstructure in this figure has been recrystallized and recovered due to heat treatment, it is well documented that the particle does disturb the matrix flow pattern and result in two different microstructures during deformation. It is possible that the flow pattern around the particles could also change when the applied states of stress on the composite are varied. This could affect the effective flow stress in the matrix differently under uniaxial tension, equibiaxial tension, and compression.

D.-G.C. SYU, formerly Graduate Student, the University of Michigan, is Senior Engineer, Taipei Municipal Government, Hsintein, Taipei, Taiwan, Republic of China. A.K. GHOSH, Professor of Materials Science and Engineering, is with the University of Michigan, Ann Arbor, MI 48109-2136.

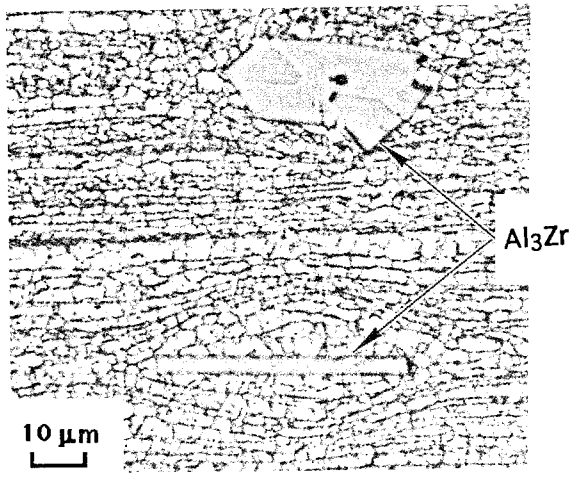
Manuscript submitted August 11, 1992.

Zahl and McMeeking[5] used the finite element method (FEM) to study the effect of residual stresses on the yielding behavior of composites comprised of elastic particles well bonded to a ductile matrix. They found that for a composite with a uniform distribution of matrix material around the particles, *i.e.*, spheres or unit cylinders with no clustering, a transient softening of the composite was predicted due to residual stresses around the particles. This transient softening was comparable under tension and compression. With an increase in the volume fraction and/or an increase in thermal straining, which results in an increase in the yielded area, the composite would have a lower effective yield stress under compression than under uniaxial tension. However, the maximum stress limit of the composite was not influenced by the residual stress. They also concluded that the effect of softening was more pronounced with angular particles than with spherical particles.

In addition to the local turbulence around the particles during deformation and the effect of residual stresses, a nonuniform particle distribution also causes the variety of complex local stress and strain states, which can change the flow behavior of the matrix. Brockenbrough *et al.*[6] used the FEM calculations and reported that the fiber distribution in a metal-matrix composite reinforced with continuous fiber could significantly influence the overall stress-strain response. Davidson[7] used the stereoimaging technique to measure the localized strains due to nonuniform particle distribution and particle spacing under uniaxial tension. He reported that while overall tensile elongations were measured at 1.6 to 2.4 pct, the localized strains could be as high as about 50 pct in a 2014 Al/15 vol pct SiCp composite due to the interaction between adjacent particles and matrix. This implies that the local stress state or plastic flow can significantly change due to the nonuniform particle distribution. Thus, it is possible that varying macroscopic stress state in DRA matrix composites can change the local flow



(*a*)



(*b*)

Fig. 1—Examples of flow turbulence during rolling. (*a*) Schematic of rotated deformation zones (recovery zones) and distortions associated with a large-round particle after rolling followed by heat treatment.[4] (*b*) Photomicrograph showing elongated Al_3Zr intermetallic and deformation zone around it, after extrusion. The primary compressive stress is normal to the direction of the particle.

pattern and eventually influence the macroscopic effective σ-ε behaviors. An investigation, therefore, was undertaken to study the interaction of the stress-state effect and nondeformable particles on the effective σ-ε curve for a soft matrix, using 2014 Al/15 vol pct Al_2O_3.*

*Manufactured by Duralcan Aluminum Composites Corporation, San Diego, CA. Produced by dispersion of Al_2O_3 particles in molten 2014 Al, then casting and extruding the ingot. The average particle size was 5.6 μm and subgrain size was 2 μm.

II. EXPERIMENTAL APPROACH

To examine if the effective stress-strain function for DRA matrix composites vary with stress state, several stress states were selected for this study. They were uniaxial tension, equibiaxial tension (bulge test), and compression. An unreinforced matrix material, 2014 Al, was also included in the study as a control material for comparison. The uniaxial tension and compression tests are commonly used and do not require extensive elaboration, although the reader can consult References 8 and 9 for pertinent details. It is worthwhile, however, to review the equibiaxial tension test. The most common method for the equibiaxial tension test is a hydraulic or pneumatic bulge test. In this test, a circular sheet diaphragm is deformed under fluid pressure, and, thus, it allows biaxial tensile deformation without frictional effects.

The determination of the effective stress-strain curve from the bulge test has been reported by many researchers.[1,10–18] Figure 2 shows a schematic diagram of the bulge test in which the pole displacement is detected by a linear variable differential transformer (LVDT). Thus, the radius of curvature, ρ, of the bulge and the angle, θ, can be determined from the following equations:

$$\rho = D/2 + (R^2/2D) \quad [1]$$

$$\theta = \sin^{-1}(R/\rho) \quad [2]$$

where R is the radius of circular die and D is the pole displacement.

The requirements and assumptions utilized for determining relevant data from the bulge test are:

(1) maintain a small ratio (about 1 ~ 2 pct) of the sheet thickness to the radius of curvature of the bulge in order to eliminate bending effects;
(2) use a pressurization rate that maintains an approximately constant strain rate in the diaphragm;
(3) assume that the curvature of the deforming sheet is nearly spherical;
(4) assume the deforming sheet maintains a constant volume and is isotropic in nature; and
(5) assume von Mises yield criterion for calculation of effective stress and strain.*

*While initial yield stress may be affected by residual thermal stresses, the effects of initial yielding are eliminated after plastic strains exceeding 2 pct.

Based on these assumptions, it is possible to calculate the overall lengthening of the sheet at any stage of deformation, but it is not possible to uniquely determine the strain at the pole of the bulge. Our experiments show that at various stages of formation of the bulge, the thickness of the entire diaphragm is nearly uniform. Thus, uniform thinning is a reasonable assumption to calculate pole strain from the bulge geometry. However, because the periphery of the sheet deforms in plane strain and the pole deforms in equibiaxial tension, a stress gradient and, therefore, a thickness gradient develop in the sheet. Typically the pole region is 20 to

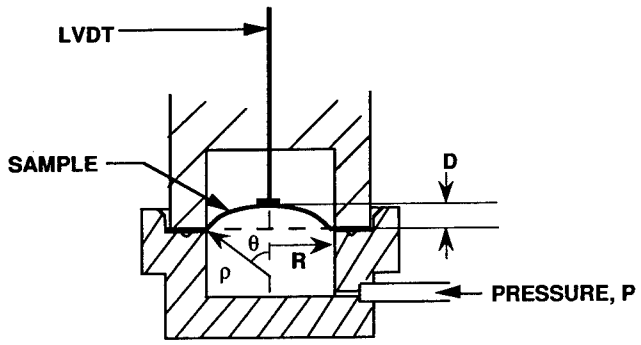


Fig. 2—Schematic diagram of the bulge test.

50 pct thinner than the edge region.[12,14] These uncertainties in strain determination do not make bulge tests ineffective, but a recognition of the source of uncertainties allows us to set bounds on the possible position of the stress-strain curve in equibiaxial tension.

Thus, the effective strain at the pole of the bulged sample was approximated by several approaches. The first approach (Method A) is based on the assumption of uniform thickness in the sample; therefore, the effective strain, ε_{eff}, can be determined from the following equations:

$$A = 2\pi\rho^2(1 - \cos\theta) = 2\pi R^2[(1 - \cos\theta)/(\sin^2\theta)] \quad [3]$$

$$\varepsilon_{eff} = -\ln(t/t_0) = \ln(A/A_0)$$
$$= \ln[2(1 - \cos\theta)/(\sin^2\theta)] \quad [4]$$

where A_0 and A are, respectively, the initial surface area and the surface area of the deformed diaphragm, and t_0 and t are the initial and current thicknesses of the sheet, respectively. Since a uniform strain in the sample was assumed, the thickness strain, ε_3, can be represented by $[-\ln(A/A_0)]$. For balanced biaxial tension, the effective strain, ε_{eff}, is equal to $-\varepsilon_3$.

The second approach (Method B) is based on uniform lengthening of the curved sheet along each meridian. Thus the lengthening (radial) strain, ε_1, can be determined from

$$\varepsilon_1 = \ln(\rho\theta/R) \quad [5]$$

since $\rho\theta$ is the length of the arc. For equibiaxial tension at the pole, the effective strain, ε_{eff}, as well as $-\varepsilon_3$ is $2\varepsilon_1$. Using Eqs. [1] to [5] for varying values of θ it is found that the effective strain from Method B is about 1.33 times higher than that from Method A.

The third approach (Method C) is based on realistic pole thinning that is greater than that at the edge. According to Jovane,[12] the actual thickness at the pole of the bulged sample is nearly 20 pct less than the mean sample thickness at large strains. For this case the effective strain at the pole is about 1.22 times higher than that obtained from Method A.

The effective stress in equibiaxial tension is equal to the hoop or radial stress (σ_1), using von Mises criterion, and, therefore, this can be given by the membrane stress equation:

$$\sigma_{eff} = \sigma_1 = \rho P/2t \quad [6]$$

where P is the pressure and the instantaneous thickness, t, can be given by

$$t = t_0 \exp(\varepsilon_3) = t_0 \exp(-\varepsilon_{eff}) \quad [7]$$

Thus for the various methods described in this section, the determination of the effective stress and effective strain for the bulge test was made by using Eqs. [1] through [7].

III. EXPERIMENTAL PROCEDURES AND RESULTS

Samples of 2014 Al and 2014 Al/15 vol pct Al_2O_3 were received in as-extruded condition.[8] The diameters of these rods were 38.1 and 50.8 mm, respectively. Several sample geometries were used to obtain the various stress states under investigation. For the uniaxial tension test, samples were machined directly from an as-extruded rod and tested both along the extrusion direction (L-direction) and along the transverse direction (T-direction). The sample geometry for this test is shown in Figure 3, and its dimensions are given in Table I. For the compression test, samples were machined directly from an as-extruded rod and tested along the L-direction (Figure 3 and Table I for details). For the equibiaxial tension, a thin-sheet sample was needed for circular bulge test. Because the diameter of the available bulge-test die was 25.4 mm, the sheet thickness needed was less than 150 μm in order to avoid bending effects. The convenient way to prepare this kind of thin sample is rolling. The method is discussed subsequently.

A. *Preparation of Thin Sample for Bulge Test*

For the composite, a 25.4-mm-thick section was cut from the extruded rod (50.8-mm diameter), then hot forged at 500 °C to 80 pct reduction. The sample was then encased in a 1.59-mm-thick stainless steel sheet and hot-rolled at 450 °C ~ 470 °C. The reduction was 20 to 30 pct in each pass, and the final thickness was between 40 and 115 μm. Cross-rolling was performed every other pass so that the resulting composite would be expected to show a planar isotropic behavior. For the matrix material, a 3.175-mm-thick slab was cut from the as-extruded rod (38.1-mm diameter) and rolled at room temperature, including cross-rolling at room temperature, to obtain a final thickness between 40 and 125 μm. Several intermediate annealing steps were performed by heating these samples to 400 °C for about 10 minutes and then quenching them in water.

The geometry of the thin samples for the bulge test is shown in Figure 3, and dimensions are given in Table I. Sample thicknesses were measured in different locations by a micrometer and by an optical microscope, before and after the test, to minimize measurement errors. The thickness variation for the thin samples were controlled under ±5 pct. For example, if the yield strength for a sample with a thickness of 50 μm is 125 MPa, the uncertainty of this value would lie between 119 and 132 MPa (±7 MPa).

To examine if the material properties have changed

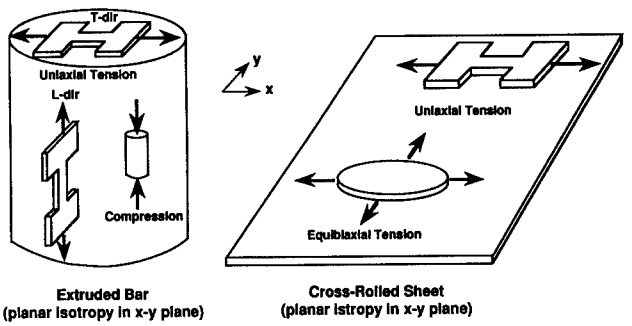


Fig. 3—Schematic diagrams showing the correlation of the sample testing directions under different stress states sectioned from the extruded bar and the cross-rolled sheet.

Table I. Sample Dimensions for Uniaxial Tension Test, Bulge Test, and Compression Test

Testing Mode	Sample Dimensions
Uniaxial tension (L- and T-direction) (as-extruded)	3.05 mm (width) × 3.175 mm (thickness) gage length = 11.11 mm
Uniaxial tension (thin sample) (as-rolled)	9.525 mm (width) × ~100 μm (thickness) gage length = 26.99 mm
Equibiaxial tension (thin sample) (as-rolled)	41.28 mm (diameter), 40 ~ 125 μm (thickness)
Compression (as-extruded)	5.08 mm (diameter), 10.16 mm (height)

after rolling, the thin samples were also tested under uniaxial tension. Figure 3 and Table I give the sample geometry details for the uniaxial tension test specimens.

Samples for testing in each stress state were given a solution-heat-treatment to eliminate the effects of prior precipitation and tested immediately after the treatment to avoid aging effects. The solution treatment involved heating the sample to 520 °C in 5 minutes, holding for 25 minutes, and then quenching in water. Microstructures of the 2014 Al matrix and the composite before and after rolling were examined after polishing and etching in Keller's reagent (H_2O:HNO_3:HCl:HF = 190:5:3:2 by volume) for 10 seconds. The optical micrographs of the matrix material and the composite are shown in Figures 4 and 5, respectively. It is seen that for the matrix alloy the average grain size in the as-extruded condition was 500 ~ 2000 μm that was refined to 10 ~ 20 μm after rolling and solution-heat-treatment. For the composite, both the grain size and particle distribution are nonuniform in the extruded condition (Figure 5(a)). Upon rolling, the particle distribution became quite uniform, but the grain size did not significantly change and remained in the range of 2 to 10 μm (Figure 5 (b)). The average Al_2O_3 particle size is 5.6 μm, and they possess an aspect ratio of approximately 2:1.

B. *Determination of σ-ε Curves under Uniaxial Tension, Equibiaxial Tension, and Compression*

The procedures and results for the various mechanical tests outlined in Section III–A are discussed in this section. All tests were conducted at room temperature and a strain rate of 10^{-4} s^{-1}. In analyzing the results, elastic strains were subtracted from total strain to develop a relationship between stress and plastic strain.

1. *Uniaxial tension test*

A chart of load *vs* displacement from the tensile tests, conducted on an Instron universal testing machine, was used to calculate the effective σ-ε curve, as discussed in Reference 8. Data was analyzed up to the point of maximum load or fracture. Figure 6 shows the comparison of effective tensile σ-ε curves between the composite and the matrix material. The Al_2O_3 particles contribute to the significantly higher flow stress of the composite and its higher initial strain-hardening rate. However, the stress saturation in the composite also causes an early attainment of instability.[4] Furthermore, the flow stress and fracture strain along the L-direction were higher than those along the T-direction in the composite. The difference of the flow stress between the two testing directions is possibly due to the oriented distribution of Al_2O_3 particles aligned along the extrusion direction. This aligned orientation combined with their aspect ratio and possibly a crystallographic texture in the matrix produced during extrusion are responsible for the difference between L- and T-direction.[8]

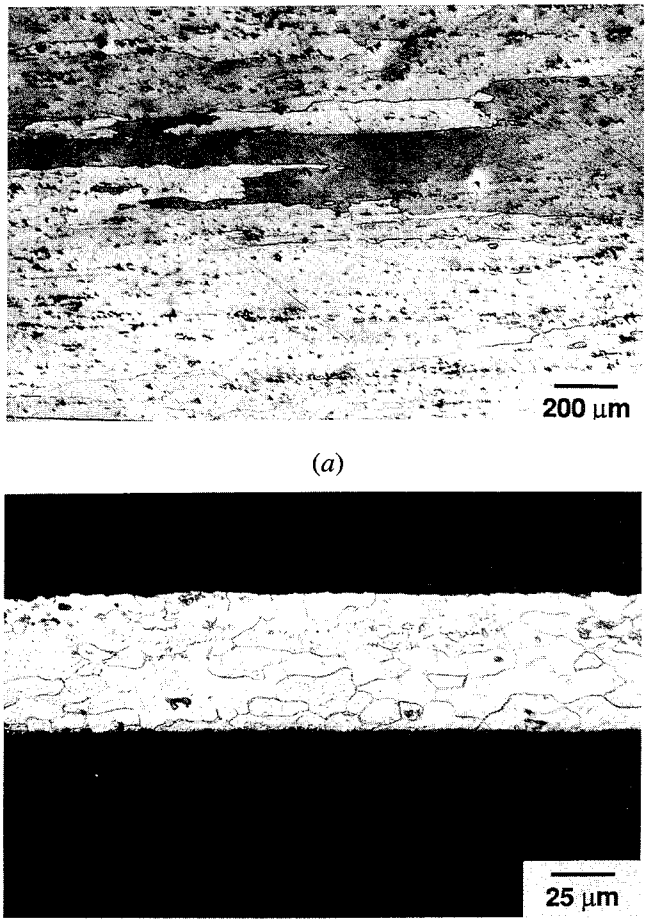
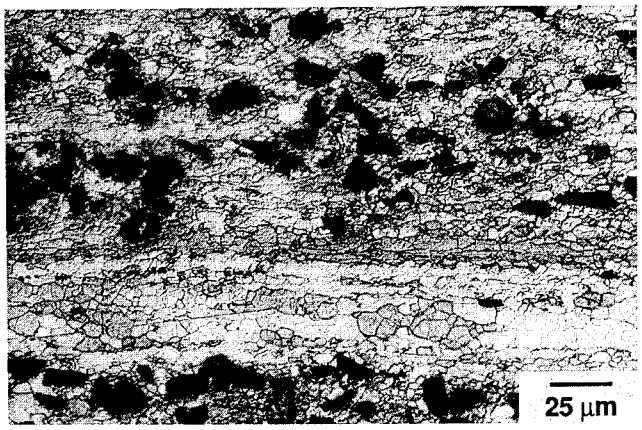


Fig. 4—Micrographs of 2014 Al. (*a*) As-extruded condition followed by solution-heat-treatment. (*b*) As-rolled condition followed by solution-heat-treatment. Samples were etched by Keller's reagent for 10 seconds. Grain sizes in the as-extruded condition were about 500 to 2000 μm and 10 to 20 μm in the rolled thin sample.

For the matrix material, the effective tensile σ-ε curves for different thin and thick samples are shown in Figure 7. It is seen that the effective tensile σ-ε curves are nearly the same in the thin samples (thickness ranges from 35 to 53 μm and grain sizes of 10 to 20 μm) and the thick sample along the L-direction (a cross-sectional



(*a*)



(*b*)

Fig. 5—Micrographs of 2014 Al/15 vol pct Al_2O_3. (*a*) As-extruded condition followed by solution-heat-treatment. (*b*) As-rolled condition followed by solution-heat-treatment Samples were etched by Keller's reagent for 10 seconds. Grain sizes in both cases were about 2 to 10 μm.

area of 3.175 × 2.286 mm and grain sizes of 500 to 2000 μm). This implies that the thickness and grain size did not significantly influence the effective tensile σ-ε curves in the matrix alloy. The earlier onset of instability in one of the thin samples is possibly a result of minor thickness variations.

For the composite, it is found that rolled thin samples with thickness ranging from 38 to 57 μm, have the similar effective tensile σ-ε curves, as shown in Figure 8. The effective tensile σ-ε curve in the thin samples was also found to be similar to that in the thick sample (a cross-sectional area of 3.175 × 2.794 mm) tested along the T-direction; however, the flow stresses in the above two cases were lower than that along the L-direction, as shown in Figure 8. Again this higher strength in the L-direction may be related to the aligned orientation of the particles.

2. *Bulge test (equibiaxial tension test)*

A schematic diagram for bulge-test apparatus is shown in Figure 9. A circumferential groove was designed into the die plate to allow sealing at both ambient and elevated temperatures. For the room-temperature test, an O-ring was placed in the groove to provide a seal, and

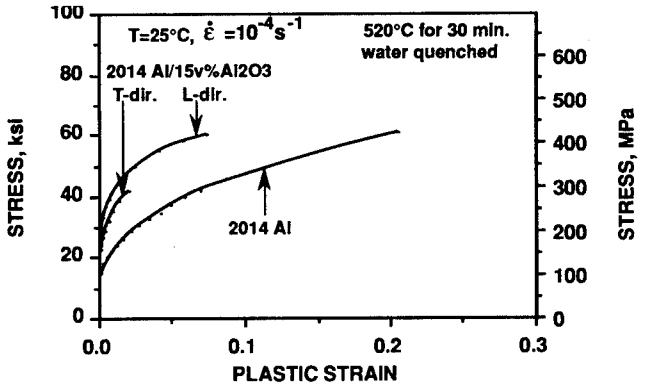


Fig. 6—Effective σ-ε curves of 2014 Al and 2014 Al/15 vol pct Al_2O_3 under uniaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water. Al_2O_3 particles in the composite significantly increased the flow stress and accelerated the onset of stress saturation and flow instability.

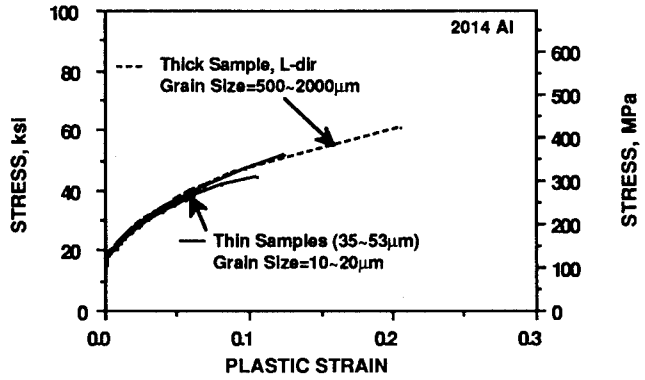


Fig. 7—Effective σ-ε curves of 2014 Al under uniaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water. The sample thickness and grain size did not influence the effective σ-ε curve.

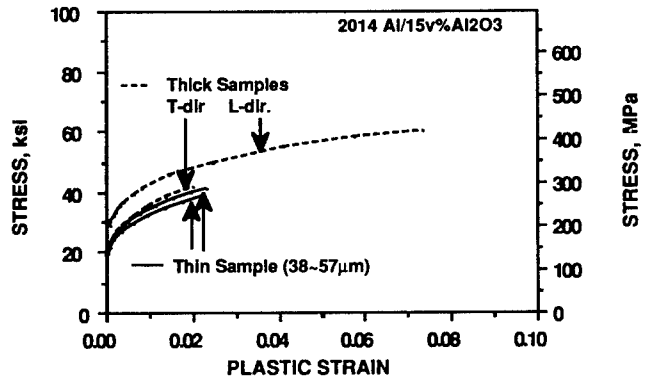


Fig. 8—Effective σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under uniaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water. The sample thickness did not influence the effective σ-ε curve.

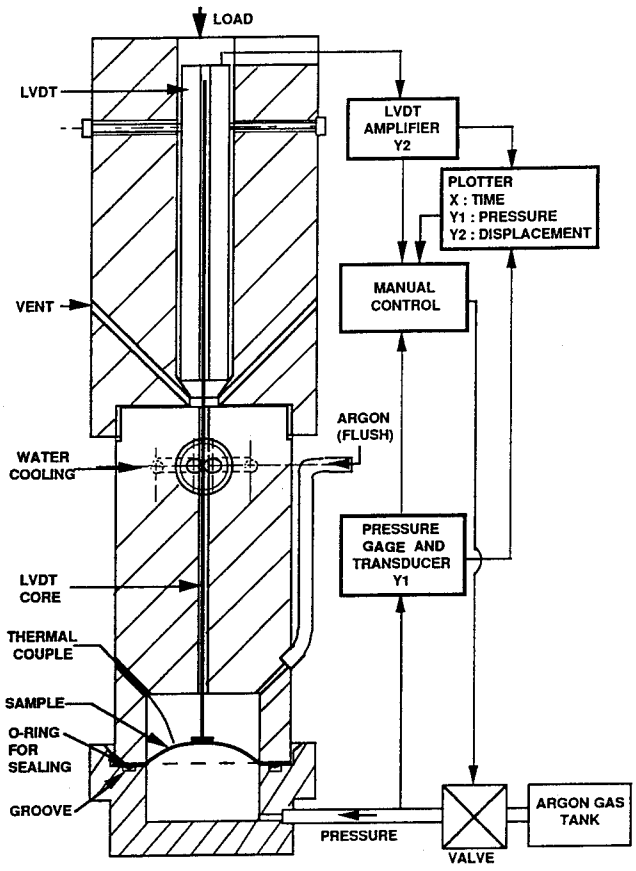


Fig. 9—Schematic diagram of the bulge test and the control system.

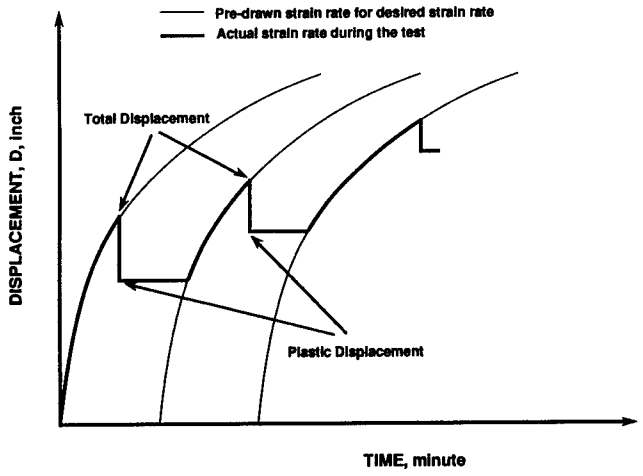


Fig. 10—Schematic of the displacement (strain) *vs* time during the bulge test. Plastic displacements were determined by releasing pressure during the test.

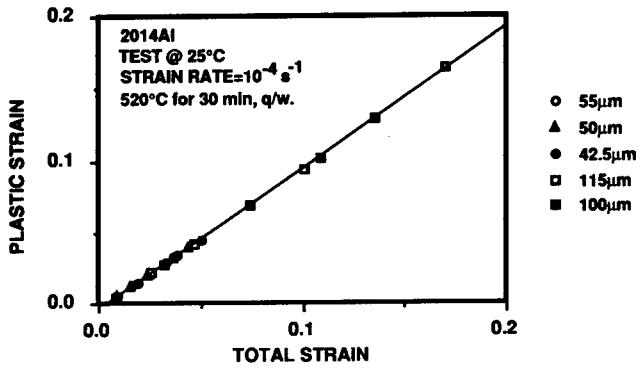


Fig. 11—Plastic strain is proportional to total strain for 2014 Al in the bulge test. Sample thickness did not influence the relationship between total strain and plastic strain.

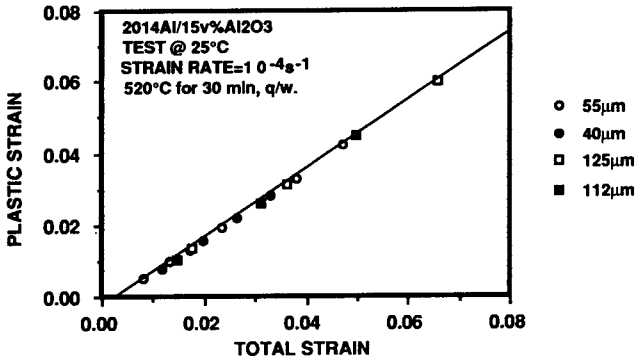


Fig. 12—Plastic strain is proportional to total strain for 2014 Al/15 vol pct Al_2O_3 in the bulge test. Sample thickness did not influence the relationship between total strain and plastic strain.

for the high-temperature test, a soft aluminum ring could be used. To maintain a nearly constant strain rate, the rate of pole displacement of the bulge needed to be controlled. To do this, the relationship between D and ε_{eff} was first determined from Eqs. [1] to [3] and [5] by computer calculation for a given strain rate (*e.g.*, 10^{-4} s^{-1}), and the resulting D *vs* time was determined. Next, several D *vs* time curves were drawn parallel to each other on the chart paper before the test, as shown in Figure 10. To determine the plastic strain rather than the total strain, tests were conducted in an interrupted manner by releasing forming pressure to record the plastic displacement. The test was restarted by returning to the forming pressure after a few seconds to avoid recovery in the sample. The external pressure was controlled to follow the D *vs* time curve predrawn on the chart during the test. The pressure *vs* time curve was simultaneously recorded on the same recorder.

Figures 11 and 12 show the relationships between total strain and plastic strain in 2014 Al and 2014 Al/15 vol pct Al_2O_3, respectively. In both cases it was found that the plastic strain was proportional to the total strain and independent of the sample thickness. It was also found that when the plastic strain approached zero, the total strain in the matrix material (about 0.005) was slightly larger than that in the composite (about 0.004), probably because the elastic modulus of the matrix alloy is lower than that in the composite.

Figure 13, shows the effective σ-ε curves of 2014 Al under equibiaxial tension at a strain rate of 10^{-4} s^{-1} using computational Methods A, B, and C, respectively, as discussed in Section II. It is seen that the sample thickness has a small effect on the stress-strain curves, the stresses being lower for the thicker sheets in the thickness range of 42.5 to 115 μm. Similar curves for Methods A, B, and C are shown in Figure 14 for the

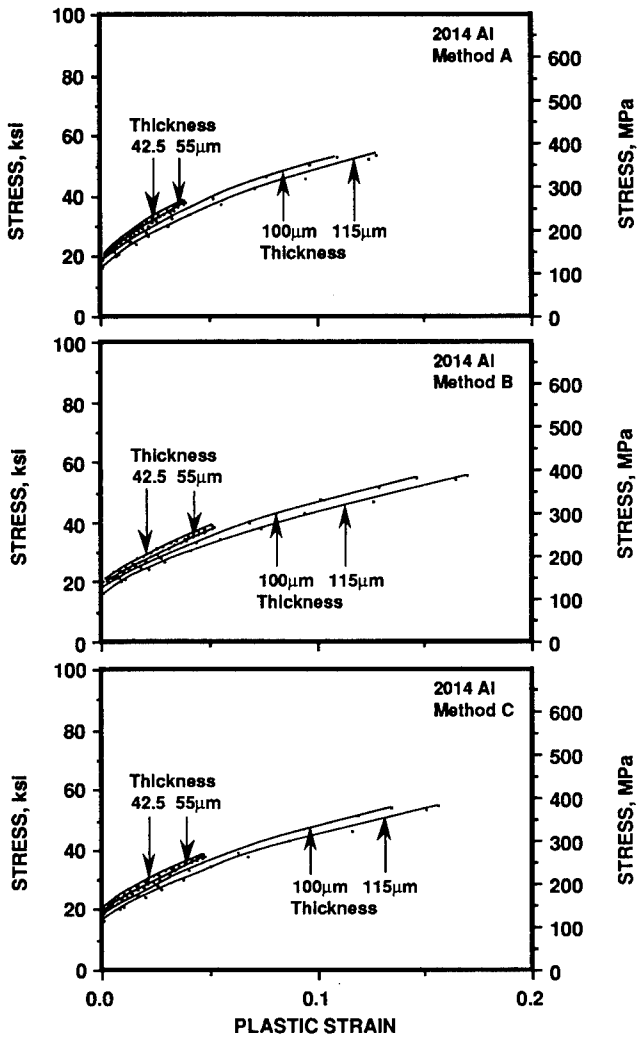


Fig. 13—Effective σ-ε curves of 2014 Al under equibiaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. The effective strains were calculated based on methods A, B, and C. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water.

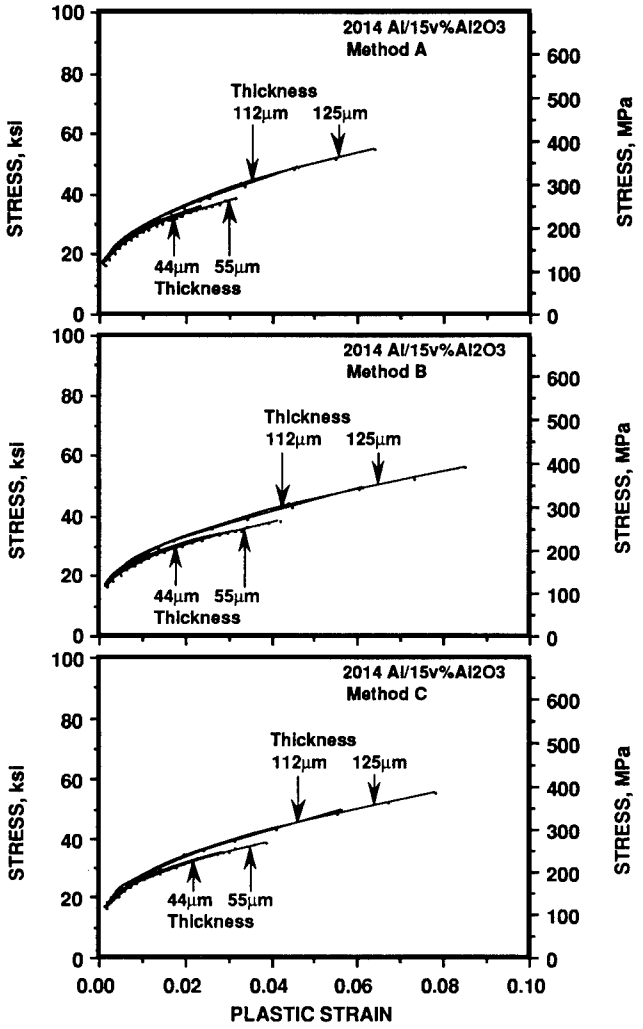


Fig. 14—Effective σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under equibiaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. The effective strains were calculated based on Methods A, B, and C. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water.

composite. The yield strength was the same for all thicknesses in the range from 44 to 125 μm, as shown in Figure 14; however, the thinner sheets showed an earlier onset of instability possibly due to variations existing in sheet thickness.

3. *Uniaxial compression test*

Compression tests were carried out on samples of matrix as well as composite material lubricated with graphite. The tests were terminated at a low-strain level (about 5 pct) to avoid the barreling effect or plastic instability.[19] To calculate the stress and plastic strain at the stage just after yielding in the compression test, the same method was used as in the uniaxial tension test. After the test was completed, the diameter of the deformed sample was measured at different positions to examine the form of the "barrel" shape. It was found that the difference in diameter between the center and the ends of the sample was less than 1 pct in both materials, *i.e.*, barreling could be ignored. Thus, the final strain was determined by using twice the final diametrical true strain. The final stress was determined by the final load divided by the final cross-sectional area. Effective σ-ε curve calculated from the load, and displacement plot was combined with σ-ε values from the final loading point to obtain the entire σ-ε curve.

It should be noted that the final stress at end of test provides a correct measure of effective stress (free of friction effects), since barreling was essentially absent. Thus, the as-tested data is a close approximation to the true stress-strain curve. However, in order to determine a lower bound of σ for possible influences due to friction, a calculation for friction correction was made using slab analysis.[20] According to this, the average pressure, P_{ave}, is related to stress, σ, on the sample by

$$P_{ave} = (\sigma/2)\,(h/\mu R')^2\,[\exp(2\mu R'/h) - (2\mu R'/h) - 1] \quad [8]$$

where P_{ave} is determined from the chart of load *vs* displacements and so are h and R', the instantaneous height and radius of the deformed sample. The friction coefficient, μ, was chosen as 0.3 for the calculation.

Figures 15 and 16 show the compressive σ-ε curves of 2014 Al and the composite, respectively, for a strain rate of 10^{-4} s^{-1} at room temperature, with several tensile curves for comparison. It is seen for both materials that the flow stresses are higher in compression than in uniaxial tension; however, after correcting for frictional effects, the difference becomes considerably less. Since it is not clear if friction correction is needed here, the actual stress level may lie between the data for as-tested and $\mu = 0.3$ case.

C. *Comparison of Effective σ-ε Curves under Various Stress States*

For the matrix alloy, the effective yield stress under compression, corrected for friction using $\mu = 0.3$, were about 24 MPa higher than that under uniaxial tension along the L-direction, as shown in Figure 15. The possible reason for this will be discussed in Section IV. For the composite, it was found that the compressive yield stress, corrected for friction using $\mu = 0.3$, was about 16 MPa less than the tensile one along the L-direction but about 32 MPa higher than that along the T-direction. Even though sticking-friction assumption lowers the effective stress for compression, we do not believe that this assumption is realistic for well-lubricated tests such as those conducted here. In fact, $\mu = 0.3$ is believed to be too high in this case. Thus, it is concluded that, in general, the compression data is somewhat higher in stress level than tension data.

Figure 17 shows a comparison of the effective σ-ε curves for matrix material under uniaxial tension and equibiaxial tension (equibiaxial tension data from the 100-μm-thick sample) and calculated using Methods A, B, and C. It is seen that the effective σ-ε curve was nearly the same under uniaxial tension and equibiaxial tension.

Figure 18 shows a comparison of the effective σ-ε curves for the composite material under uniaxial tension and equibiaxial tension, where the effective strains under equibiaxial tension were chosen from the sample with 125-μm thickness and calculated from Methods A, B, and C. It was found that the average effective yield stress in the thin sample under uniaxial tension was about 42 (using Method A) to 46 MPa (using Method B) higher than that under equibiaxial tension. This difference is much higher than any possible error (±7 MPa) due to the thickness variation in the sample as mentioned in Section III–A. The same observation was also found

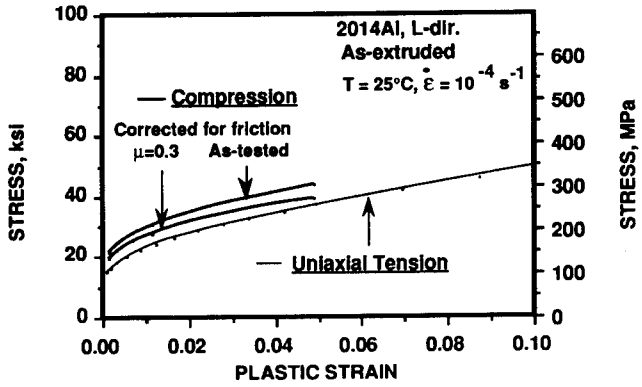


Fig. 15—Effective σ-ε curves of 2014 Al under uniaxial tension and compression at room temperature and a strain rate of 10^{-4} s^{-1}. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water.

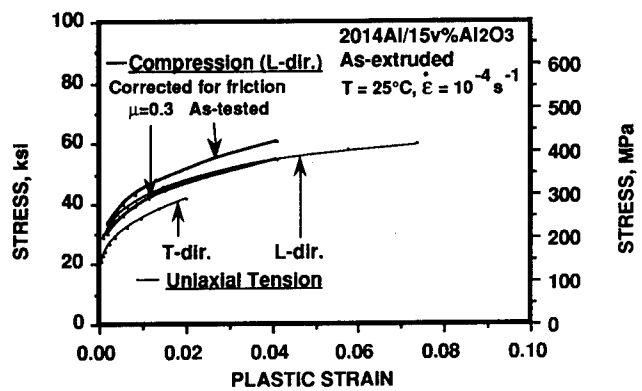


Fig. 16—Effective σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under uniaxial tension and compression at room temperature and a strain rate of 10^{-4} s^{-1}. Samples were solution-heat-treated at 520 °C for 30 minutes then quenched in water.

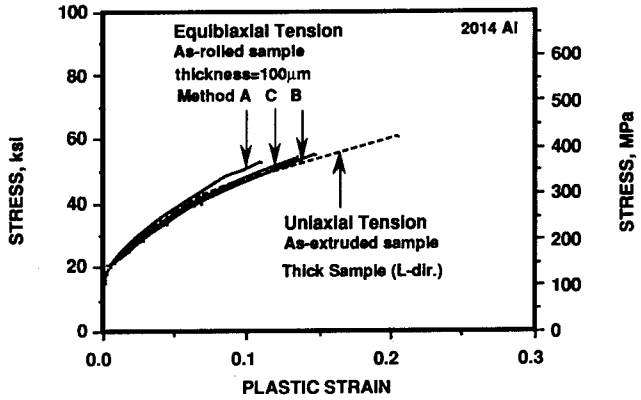


Fig. 17—Comparison of effective σ-ε curves of 2014 Al under uniaxial tension and equibiaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. The effective strains under equibiaxial tension were calculated based on Methods A, B, and C.

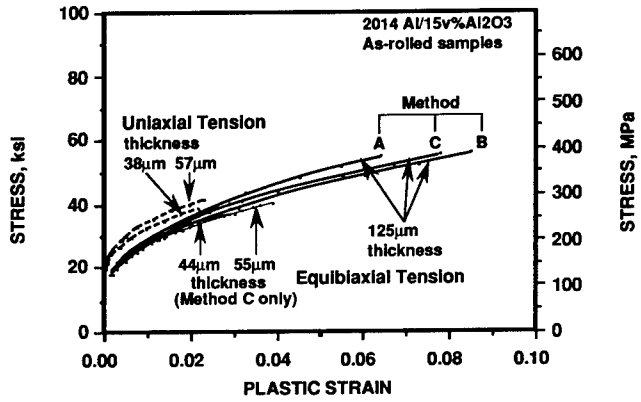


Fig. 18—Comparison of effective σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under uniaxial tension and equibiaxial tension at room temperature and a strain rate of 10^{-4} s^{-1}. The effective strains under equibiaxial tension were based on Methods A, B, and C.

at a higher strain rate (10^{-3} s^{-1}), shown in Figure 19, where the effective strain was determined from Method A.

Table II is the summary of the effective yield stresses of 2014 Al and 2014 Al/15 vol pct Al_2O_3 under the three testing modes at ambient temperature and a strain rate of 10^{-4} s^{-1}. It is seen that the effective yield stress of the composite is higher than that of matrix alloy under either uniaxial tension or compression. However, the effective yield stress of the composite is not different from that of the matrix material under equibiaxial tension. The strength coefficient (K) and strain-hardening exponent (n) for the composite are listed in Table III, according to $\sigma = K\varepsilon^n$. It is seen that the composite had a higher strain-hardening exponent under equibiaxial tension ($n = 0.251 \sim 0.303$) than that under either uniaxial tension ($n = 0.184$) or compression ($n = 0.196$). The reason for this is unknown; however, one of the possible explanation is a nondeformable area (dead-zone) developing around the particle that will change with the applied-stress-states change and result in the change of the stress-strain function. This will be discussed in the Section IV.

It was also found that for the composite the tensile yield stress along the T-direction in the extruded bar was similar to that along the rolling direction in the cross-rolled sheet, as shown in Figure 8. This implies that in Figure 3 the yielding behavior of the x-y plane in the extruded bar is similar to that in the cross-rolled sheet for the composite material.

IV. DISCUSSION

A. *Stress-State Dependence of σ-ε Curve*

For the matrix material, 2014 Al, the grain size was refined due to rolling, as seen in Figure 4; however, the tensile effective σ-ε curves for these thin samples were not significantly different from the thick as-extruded samples, as shown in Figure 7. The flow stress of the matrix under uniaxial tension was less than that under compression but was equal to that in equibiaxial tension.

For the composite, the flow stress under equibiaxial tension was less than that under uniaxial tension for all levels of strain. Again, the uniaxial tension data was somewhat lower in stress level than compression. As previously mentioned, these general trends are not masked by possible errors due to thickness variation, curvature measurements, *etc.* The data were determined statistically, and parallel tests were done under the same conditions for both the composite and the matrix material. No significant difference between the σ-ε curves is found between uniaxial tension and equibiaxial tension for the matrix alloy. In addition, the effective strains in the bulge test were calculated by three different methods to check the bounds of possible errors. These conclusions are thus believed to be well founded.

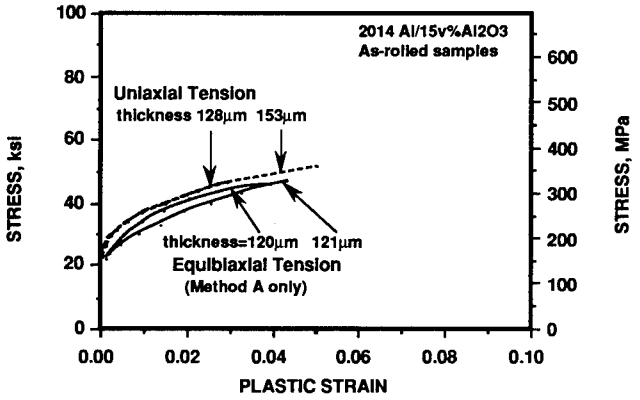


Fig. 19—Comparison of effective σ-ε curves of 2014 Al/15 vol pct Al_2O_3 under uniaxial tension and equibiaxial tension at room temperature and a strain rate of 10^{-3} s^{-1}. The effective strains under equibiaxial tension were calculated based on Method A.

Several factors can be responsible for the observed differences in effective stress-strain function, such as the different microstructures in the matrix and the areas surrounding the nondeformable particles[4] in the composite, residual stresses,[5] particle shape,[5,6] and nonuniform distribution of the particles.[6,7] One might also suggest that the development of hydrostatic tension in the equibiaxial tension case can assist void initiation and yield a lower flow stress. This would be expected to lower the fracture strain in equibiaxial tension. However, it was found that the fracture strain in the equibiaxial case was higher than that in the uniaxial one, and, also, voids were not observed in the deformed sample to support this possibility. It is suggested in this work that the formation of nondeformable matrix zone around the particles may provide an improved understanding of the problem. This is discussed in Section IV–C.

B. *Effect of Particles on Strength*

In Figure 6 the strengthening effect of the Al_2O_3 particles on the 2014 Al matrix is seen. Composite-strengthening theories have been reported by several researchers.[21–28] For the fiber-reinforced metal-matrix composites, the basic principle of the strengthening is the volume-fraction rule (VFR):[21]

$$\sigma_{cy} = \sigma_{my}(1 - f) + \sigma_{fy}f \qquad [9]$$

where σ_{cy}, σ_{my}, and σ_{fy} are the yield stresses of the composite, matrix, and fiber, respectively, and f is the volume fraction of the reinforcement. Here the equal-strain approximation between matrix and fibers is assumed. For the whisker-reinforced metal-matrix composite, the VFR can be modified based on the simple transfer of shear across the interface between the matrix and the whiskers:[21]

$$\sigma_{cy} = \sigma_{my}(1 - f) + \tau_{my}(L_w/d_w)f \qquad [10]$$

where τ_{my} is the shear yield stress of the matrix, L_w and d_w are the length and diameter of the whiskers, respectively. This is the so-called shear-lag model.

The modified shear-lag model[22] predicts the yield stress of the whisker-reinforced metal-matrix composites using the shear-lag model, Eq. [10], and taking into account the load transfer across the ends of the whisker:

$$\sigma_{cy} = \sigma_{my}(1 - f) + \tau_{my}(L_w/d_w)f + \sigma_{my}f$$
$$= \sigma_{my} + \tau_{my}(L_w/d_w)f \qquad [11]$$

Table II. Effective Yield Stresses (0.2 Pct Plastic Strain) of 2014 Al* and 2014 Al/15 Vol Pct Al_2O_3* under Uniaxial Tension, Equibiaxial Tension (Bulge Test), and Compression Tests at Ambient Temperature and a Strain Rate of 10^{-4} s^{-1}

	As-Rolled Samples (Thin Samples)		As-Extruded Samples (Thick Samples)	
Materials	2014 Al	2014 Al/15 vol pct Al_2O_3	2014 Al	2014 Al/15 vol pct Al_2O_3
Effective yield strength	Average σ_{my},** MPa	Average σ_{cy},** MPa	σ_{my},** MPa	σ_{cy},** MPa
Uniaxial tension	126.4	170	117.2 (L-direction)	224.1 (L-direction) 176.5 (T-direction)
Equibiaxial tension	A: 139.6† B: 135.7 C: 137.2	A: 128.4† B: 123.7 C: 125.6	N.A.	N.A.
Compression	N.A.	N.A.	(L-direction) 155.8 (μ = 0) 141.3 (μ = 0.3)	(L-direction) 230.3 (μ = 0) 208.2 (μ = 0.3)

*Solution-heat-treated at 520 °C for 30 min and then quenched in water.
**σ_{my} and σ_{cy} are the effective yield stresses of the matrix alloy and the composite, respectively.
†The effective strains were calculated based on Methods A, B, and C.

Table III. Strength Coefficient (K) and Strain-Hardening Exponent (n)* under Various Stress States for 2014 Al/15 Vol Pct Al_2O_3 at Ambient Temperature and a Strain Rate of 10^{-4} s^{-1}**

2014 Al/15 Vol Pct Al_2O_3	Strength Coefficient, K (MPa)	Strain-Hardening Exponent, n
Uniaxial tension (L-direction)	649.7	0.172
Uniaxial tension (T-direction)	613.7	0.198
Uniaxial tension (thin sample)	538.6	0.184
Equibiaxial tension (thin sample)	A: 853.3 B: 783.1 C: 806.5	0.251 to ~0.303
Compression (μ = 0.3)	694.3	0.196

*Based on stress-strain relationship following $\sigma = K\varepsilon^n$.
**Solution-heat-treated at 520 °C for 30 min and then quenched in water.

The enhanced dislocation density model,[23,24,25] developed from the modified shear-lag model, also takes into account the difference in dislocation densities between the reinforced and unreinforced materials.

Although these models have predicted the experimental data in some cases,[22–25] they are not directly applicable for particulate-reinforced composites, and they do not include any stress-state effect.

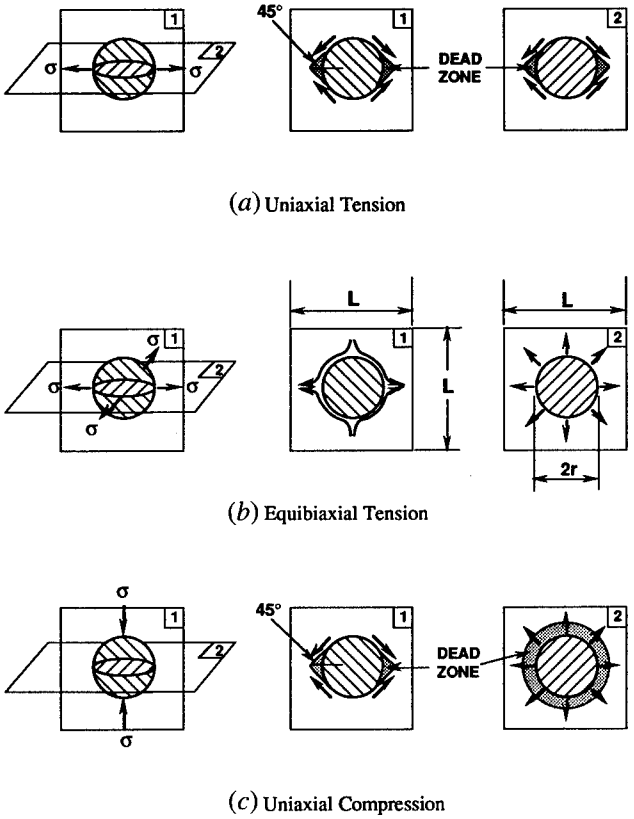
Christman *et al.*[26,27] used the FEM instead of the analytical methods to predict the yield strength of DRA matrix composites from the mechanical behavior of the constituent materials. They found that the constraint induced by the reinforcements on the adjoining matrix and the consequent development of matrix triaxiality influences the apparent flow strength of the composite. Levy and Papazian[28] also used the FEM to predict the tensile stress-strain behavior in short fiber-reinforced metal-matrix composites. They used elastoplastic analysis for both transversely aligned fibers and staggered end fibers and concluded that the above models can provide an accurate prediction of the overall composite stress-strain response. They also found that the strain-hardening rate increases with increasing fiber-volume content and fiber-aspect ratio. Their explanation was that the increased strain-hardening rate was a consequence of the increase of plastic constraint introduced by the presence of the nondeformable SiC dispersoids.

In addition to these strengthening theories, the hardening mechanism should also be related to the matrix itself and the particle volume fraction, size, shape, and spacing. In addition, the quench rate, which causes the residual stress and dislocation punch-out, is also an important strengthening parameter. For example, in 2014 Al/15 vol pct Al_2O_3, the coefficient of thermal expansion for Al_2O_3 (9×10^{-6} K^{-1}) is lower than that for Al (22×10^{-6} K^{-1}). One would expect dislocation density to increase around the particles during quenching from 520 °C to room temperature in the present case. It is believed that this increase in dislocation density can retard the plastic flow in the matrix,[24] and thereby increase the yield stress in the composite material. The above discussions explain why in Figure 6 the Al_2O_3 particles strengthen the 2014 Al matrix. However, the effects of initial yielding are less important at larger plastic

strains, where the nondeformable particles act as obstacles for plastic flow of the matrix.

C. *Proposed Concept for Stress-State Effect in DRA Composites*

A possible reason is suggested here to explain the differences in stress-strain behaviors for the composite material. Our hypothesis is that zones of nondeforming matrix develop around the Al_2O_3 particles as evidenced by lensoid zones in the deformed microstructure (Figure 1).[4] These zones that result from turbulence are behind the particles where shear flow cannot reach. Consequently, the action of multiaxial forces and the resulting plastic flow in various directions can change the shape of these zones. To describe the various flow patterns and dead zones around the particles, we first assume a cubic-unit cell containing a spherical particle at the center. The dead zones that form under uniaxial tension, equibiaxial tension, and compression may be represented as shown in Figures 20(a), (b), and (c), respectively. For instance, the dead zones under uniaxial tension are two areas with conical end caps (apex angle = 90 deg) along the axial direction, as shown in Figure 20(a). The apex angle of 90 deg is suggested by the slip-line field theory.[20] The dead zone under equibiaxial tension does not exist, because the matrix is pulled equally around the particle in tension in the plane 2, as shown in Figure 20(b). In compression test, the dead zone is a ring (or doughnut shape) with a triangular cross section (apex angle = 90 deg), as shown in Figure 20(c).



(*a*) Uniaxial Tension

(*b*) Equibiaxial Tension

(*c*) Uniaxial Compression

Fig. 20—Schematic diagrams showing the different flow patterns and the development of the dead zones under (*a*) uniaxial tension, (*b*) equibiaxial tension, and (*c*) compression.

Considering the difference of dead-zone geometries for different stress states, the total volume fraction of the nondeformable areas can be calculated for the zones shown in Figure 20 when the spherical particle volume fraction is 15 pct, the total dead-zone volume fraction, f_d, will be 16 and 20 pct under uniaxial tension and compression, respectively. For the equibiaxial tension, a dead zone does not form, and f_d remains 15 pct. Because Al_2O_3 particles and the dead zone around these particles do not participate in plastic deformation, the applied strain must be accommodated by the remaining matrix. The higher matrix strain can be determined from

$$\varepsilon_m = \varepsilon_c/(1 - f_d) \qquad [12]$$

where ε_m denotes matrix strain, and ε_c denotes composite strain. This higher matrix strain at composite yielding ($\varepsilon_c = 0.002$) is given in Table IV by ε^c_{my}.

The tensile stress-strain curve of the matrix material is shown in Figure 7, which provides data to calculate flow stresses for composite. The matrix has a yield strength of 126.2 MPa. Using the early part of the matrix stress-strain curve (Figure 7), the matrix flow stress, σ^c_{my}, is calculated for each value of ε^c_{my}.

In addition to increasing the effective yield stress in the matrix, the nondeformable area also increases the resistance to plastic flow *via* shear resistance between the matrix and the nondeformable areas. This additional force, whose component, F_1, acts along the axial-stress direction is schematically illustrated in Figures 21(a), (b), and (c). The shear stress, τ_{my}, acts on the surface of the particle and dead zone, and its axial components can be integrated to calculate the additional force, F_1, as indicated in Figures 21(a), (b), and (c) for the various states of stress. Since von Mises yield criterion gives $\tau_{my} = \sigma_{my}/\sqrt{3}$, F_1 can be expressed as

$$F_1 = f_g \pi r^2 \tau_{my} = f_g \pi r^2 (\sigma^c_{my}/\sqrt{3}) \qquad [13]$$

where f_g is a geometric factor for additional shear component determined from Figures 21(a), (b), and (c), and r is the radius of the spherical particle. It should be noted that r is related to particle volume fraction, f, by the relation $f = [(4\pi r^3/3)/L^3]$, whose unit-cubic-cell volume is L^3 (Figure 20). For $f = 0.15$, r is equal to $L/3.034$. The values of f_g under uniaxial tension, equibiaxial tension, and compression are then calculated from the scheme shown in Figures 21(a), (b), and (c), respectively, and listed in Table IV. The additional resistance to flow, $\Delta\sigma$, is then

$$\Delta\sigma = F_1/L^2 \qquad [14]$$

Thus, the effective yield stress of the composite, σ_{cy}, is obtained as

$$\sigma_{cy} = \sigma^c_{my} + \Delta\sigma \qquad [15]$$

Table IV shows the values of $\Delta\sigma$ and σ_{cy} under the three different tests. It is seen that the calculated effective yield stress of the composite under uniaxial tension is slightly less than that under compression ($(\sigma_{cy})_c/(\sigma_{cy})_u$ is about 1.002) but greater than that under equibiaxial tension ($(\sigma_{cy})_b/(\sigma_{cy})_u$ is about 0.885). The same trend is also

Table IV. Parameters for Calculating Composite-Yield Stress Compared with Experimental Effective-Yield Stress under Uniaxial Tension, Equibiaxial Tension, and Compression Tests

2014 Al/15 Vol Pct Al_2O_3	Uniaxial Tension	Equibiaxial Tension	Compression
Effective volume fraction of nondeformable area, f_d	0.16	0.15	0.20
ε^c_{my}	0.00238	0.00235	0.00250
σ^c_{my} MPa	128	127.9	128.2
Factor for additional shear component, f_g	1.785	1	1.785
$\Delta\sigma$ due to interfacial shear	45	25.2	45.1
Calculated σ_{cy} MPa	173	153.1	173.3
Calculated $(\sigma_{cy})_i/(\sigma_{cy})_b$	1	0.885	1.002
Rolled sample (thin sample)			
Experimental σ_{cy} MPa	170	126.1 (average)	NA
Experimental $(\sigma_{cy})_i/(\sigma_{cy})_u$	1	0.742	NA
Extruded sample (thick sample)			
Experimental σ_{cy} MPa	200.3 (average)	NA	208.2 (μ = 0.3)
Experimental $(\sigma_{cy})_i/(\sigma_{cy})_u$	1	NA	1.039

*Uniaxial matrix yield stress is 126.2 MPa at a plastic strain of 0.002.

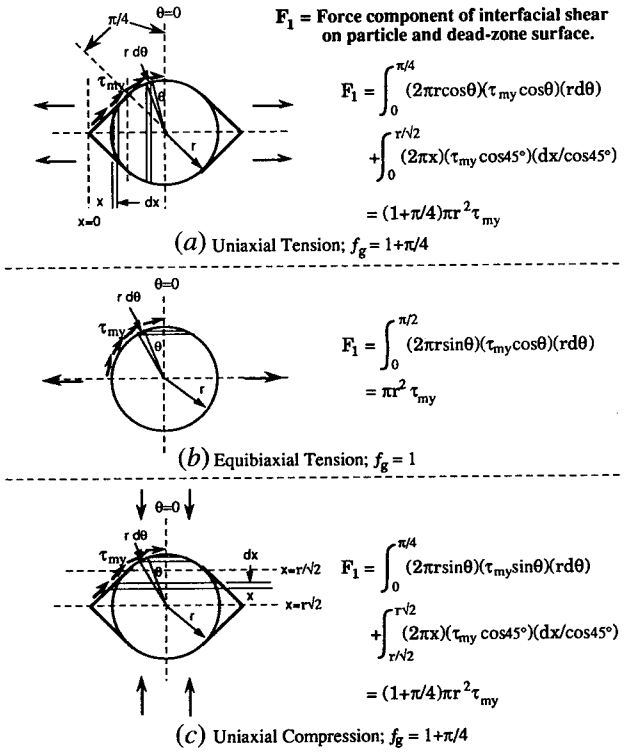


Fig. 21—Schematic diagrams and corresponding equations showing the calculation of the F_1 due to shear stress on the interface of the matrix and nondeformable area under (*a*) uniaxial tension, (*b*) equibiaxial tension, and (*c*) uniaxial compression; *r* represents the radius of spherical particle.

seen in the experimental observation; $(\sigma_{cy})_c/(\sigma_{cy})_u$ and $((\sigma_{cy})_b/(\sigma_{cy})_u$ are about 1.039 and 0.742, respectively, as shown in Table IV. It should be noted that the effective tensile yield stresses for the thick samples along the L-direction and the T-direction have been averaged to compare with compression data.

Although the values of the predicted effective yield stresses using the dead-zone model do not exactly match the experimental ones, the calculated data do follow a generally correct trend. The present hypothesis is a simple concept and does not consider the variation of the particle size, spacing, and shape that can influence the yielding behavior of the composite.[29] In addition, the dead zones can overlap each other within particle clusters. This will cause the flow pattern of the matrix to be more complicated and influence the accuracy of the calculated data using the present model. Other parameters, such as the local stress concentration and bonding strength between matrix and particles, *etc.*, could also affect the yielding behavior of the composite under different tests.

V. SUMMARY AND CONCLUSIONS

1. The stress-strain responses of a DRA matrix composite, 2014 Al/15 vol pct Al_2O_3, and its matrix alloy, 2014 Al, were investigated under a variety of stress states by conducting mechanical tests under uniaxial tension, equibiaxial tension, and compression conditions.
2. For the composite, at a strain rate of 10^{-4} s^{-1}, the von Mises effective yield stress under uniaxial tension was about 40 MPa higher than that under equibiaxial tension. On the other hand, the average tensile yield stress (L- and T-direction) was about 30 MPa lower than that under compression (μ = 0.3). For the matrix material, the effective yield stresses were nearly the same under uniaxial tension and equibiaxial tension; however, they were less than that under compression.
3. The strain-hardening rates were higher in the composite than in the matrix material under the three different stress states examined here. For the composite,

the strain-hardening exponent under equibiaxial tension ($n = 0.251 \sim 0.303$) was higher than that under either uniaxial tension ($n = 0.184$) or compression ($n = 0.196$) at a strain rate of $10^{-4}\ s^{-1}$. For the matrix material, the slopes of the stress-strain curves were nearly the same under the three test conditions.

4. Several possible sources responsible for the observed differences in effective stress-strain function were examined. It is suggested in this work that the formation of nondeformable matrix zone around the particles can be an important factor. These zones that result from flow turbulence are behind the particles where shear flow can not reach. Consequently, the action of multiaxial forces and the resulting plastic flow in various directions can change the shape of these zones. This dead-zone model is rather simple; however, it can predict a generally correct trend for the yield stress of the composite under different stress states. More detailed experimental and analytical work should be done in the future to develop this concept further.
5. It is possible that such nonuniqueness of stress-strain relationship based on the von Mises criterion will persist at elevated temperature; therefore, the development of such information could be important for applications to hot-forging problems.

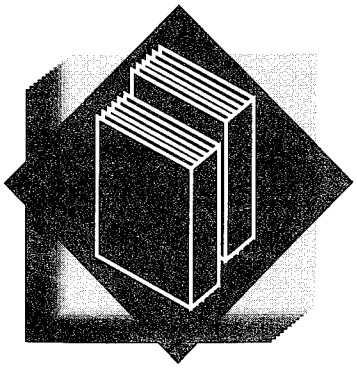
ACKNOWLEDGMENTS

The authors would like to thank the AES, GM Technical Center for partially supporting this work. Fruitful discussions with Dr. B. Taylor of GM Technical Center are gratefully acknowledged.

REFERENCES

1. C. Vial, W.F. Hosford, and R.M. Caddell: *Int. J. Mech. Sci.*, 1983, vol. 25 (12), pp. 899-915.
2. J. Woodthorpe and R. Pearce: *Int. J. Mech. Sci.*, 1970, vol. 12, pp. 341-47.
3. R.H. Wagoner: *Metall. Trans. A*, 1980, vol. 11A, pp. 165-75.
4. F.J. Humphreys, W.S. Miller, and M.R. Djazeb: *Mater. Sci. Technol.*, 1990, vol. 6, pp. 1157-66.
5. D.B. Zahl and R.M. McMeeking: *Acta Metall.*, 1991, vol. 39 (6), pp. 1117-22.
6. J.R. Brockenbrough, S. Suresh, and H.A. Wienecke: *Acta Metall.*, 1991, vol. 39 (5), pp. 735-52.
7. D.L. Davidson: *Metall. Trans. A*, 1991, vol. 22A, pp. 113-23.
8. D.-G.C. Syu and A.K. Ghosh: *Metall. Mater. Trans. A*, 1994, vol. 25A, pp. 2027-48.
9. T.C. Hsu: *Mater. Res. & Stand.*, 1969, vol. 9 (12), pp. 20-25.
10. C.C. Chu and A. Needleman: ASME *J. Eng. Mater. Technol.*, 1980, vol. 102 (F-28), pp. 249-56.
11. W. Johnson and J.L. Duncan: *Sheet Met. Ind.*, 1965, April, pp. 271-75.
12. F. Jovane: *Int. J. Mech. Sci.*, 1968, vol. 10, pp. 403-27.
13. A.R. Ragab and A.T. Abbas: ASME *J. Eng. Mater. Technol.*, 1986, vol. 108, pp. 250-57.
14. N.-M. Wang and M.R. Shammamy: *J. Mech. Phys. Solids.*, 1969, vol. 17, pp. 43-61.
15. N.-M. Wang: *Numerical Analysis of Forming Processes*, J.F.T. Pittman, O.C. Zienkiewicz, R.D. Wood, and J.M. Alexander, eds., John Wiley & Sons Ltd., New York, NY, 1984, pp. 117-64.
16. M.F. Ilahi, A. Parmar, and P.B. Mellor: *Int. J. Mech. Sci.*, 1981, vol. 23, pp. 221-27.
17. W.F. Brown, Jr., and G. Sachs: *Trans. ASME*, 1948, vol. 70, pp. 241-51.
18. W.F. Brown, Jr., and F.C. Thompson: *Trans. ASME*, 1949, vol. 71, pp. 575-85.
19. J.K. Banerjee: ASME *J. Eng. Mater. Technol.*, 1985, vol. 107, pp. 138-44.
20. W.F. Hosford: *Metal Forming: Mechanics and Metallurgy*, Prentice-Hall, Englewood Cliffs, NJ, 1983, pp. 122 and 146.
21. T.H. Courtney: *Mechanical Behavior of Materials*, McGraw-Hill, New York, 1990.
22. V.C. Nardone and K.M. Prewo: *Scripta Metall.*, 1986, vol. 20, pp. 43-48.
23. Y. Flom and R.J. Arsenault: *Mater. Sci. Eng.*, 1985, vol. 75, pp. 151-67.
24. R.J. Arsenault, L. Wang, and C.R. Feng: *Acta Metall.*, 1991, vol. 39 (1), pp. 47-57.
25. R.H. Jones, C.A. Lavender, and M.T. Smith: *Scripta Metall.*, 1987, vol. 21, pp. 1565-70.
26. T. Christman, A. Needleman, S. Nutt, and S. Suresh: *Mater. Sci. Eng.*, 1989, vol. A107, pp. 49-61.
27. T. Christman, A. Needleman, and S. Suresh: *Acta Metall.*, 1989, vol. 37 (11), pp. 3029-50.
28. A. Levy and J.M. Papazian: *Metall. Trans. A*, 1990, vol. 21A, pp. 411-2.
29. J.J. Lewandowski, C. Liu, and W.H. Hunt, Jr.: Paper presented at 1987 AIME Annual Meeting, Denver, CO, Feb. 24, 1987.

Comprehensive refractory metal technology from TMS



Evolution of Refractory Metals and Alloys

E.N.C. Dalder, T. Grobstein, and C.S. Olsen, editors

This book provides a review of the development of refractory metal alloys (those based on vanadium, niobium, tantalum, molybdenum, tungsten, and rhenium) from the 1950s to the present. The technology of many of the leading reserchers in this area has been recorded and preserved as a reference for current and future researchers.

Coverage includes:

- Vanadium Alloy Development
- Columbium Alloy Development at Westinghouse
- Evolution of Tantalum Alloy Development
- The History of Development of Molybdenum Alloys for Structural Applications
- A History of Tungsten and Molybdenum-Base Alloys
- Rhenium—A Unique Rare Metal

A collection of papers from the 1993 TMS Annual Meeting held in Denver, Colorado, February 21–25, 1993.

ISBN: 0-87339-234-5
220 pp., illus., index, hardcover

Order No. 518
TMS Member $39 Student Member $29 List $52

Klaus Schulze Symposium on Processing and Applications of High Purity Refractory Metals and Alloys

P. Kumar, H.A. Jehn, and M. Uz, editors

This volume presents the latest advances in the processing and applications of high-purity refractory metals and alloys as well as a review of previous developments. The topics of discussion cover a wide range that includes extraction, recovery, and ultrapurification of refractory metals; physical metallurgy, chemical metallurgy, and kinetics of alloys; and the latest analysis and characterization techniques. Current and potential applications, as well as the advantages and limitations of refractory metals and alloys, are also covered.

A collection of papers from Materials Week '93 held in Pittsburgh, Pennsylvania, October 17-21, 1993.

ISBN: 0-87339-228-0
269 pp., illus., index, hardcover

Order No. 510
TMS Member $64 Student Member $47 List $88

You may order these or any TMS publication by contacting:

The Minerals, Metals & Materials Society
Attn: Book Order Department
420 Commonwealth Drive
Warrendale, PA 15086-7514 U.S.A.

Telephone: 1-800-759-4867 (U.S.A. only)
or (412) 776-9000, ext. 244

Fax: (412) 776-3770

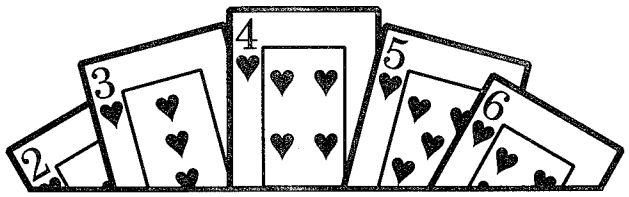
e-mail: bookorder@tms.org

or

TMS Minerals • Metals • Materials OnLine

Order TMS publications any time day or night via TMS OnLine. This computer-based, modem/telnet-accessed service provides you with free, instant access to detailed information on all TMS publications as well as 24-hour online publications ordering. For free access to TMS OnLine, via modem dial (412) 776-2040 or via Internet: Telnet to online.tms.org.

TMS
Minerals • Metals • Materials



LAS VEGAS

124th TMS Annual Meeting

The Minerals, Metals & Materials Society's 1995 Annual Meeting and Exhibition

February 12–16, 1995
Las Vegas Convention Center – Las Vegas, Nevada

Science and industry's premier forum for state-of-the-art processing, fabrication, and design technology

Plan now to join more than 3,000 of your colleagues for 1995's most anticipated presentation of minerals, metals, and materials technology. Many of the world's foremost professionals from industry and research will gather in Las Vegas to disseminate and gather technical expertise through a program of more than 200 informative technical sessions.

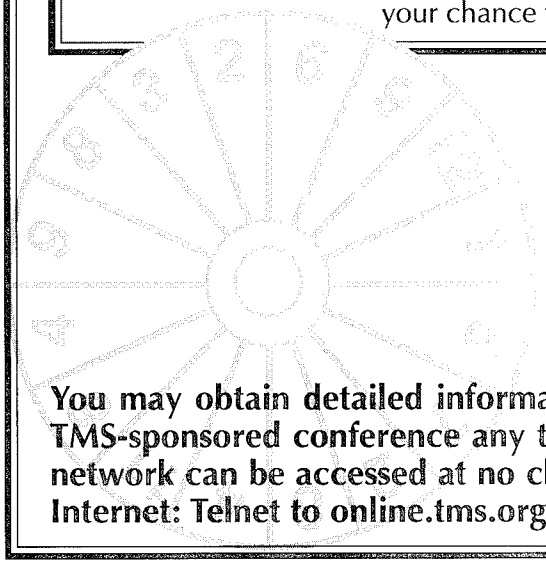


Featuring programming by the:

- Electronic, Magnetic and Photonic Materials Division
- Extraction & Processing Division
- Light Metals Division
- Materials Design and Manufacturing Division
- Structural Materials Division

Don't miss the 1995 TMS Annual Meeting Exhibition.

The latest new products, processes, and services offered by more than 130 international metals and materials organizations will be on display for you in the Las Vegas Convention Center. It's your chance to put technology to work for you.



For more information, contact:

TMS
Meeting Services Department
420 Commonwealth Drive
Warrendale, PA 15086-7514 U.S.A.
Telephone: (412) 776-9000, Ext. 232
Fax: (412) 776-3770
E-mail: karl@tms.org

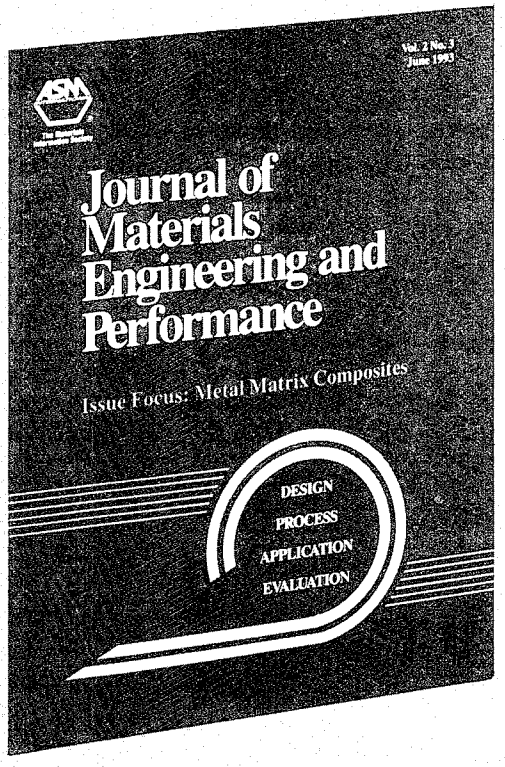
You may obtain detailed information about the 1995 TMS Annual Meeting or any TMS-sponsored conference any time day or night via TMS OnLine. This computer network can be accessed at no charge via modem by dialing (412) 776-2040 or via Internet: Telnet to online.tms.org

Journal Of Materials Engineering And Performance

ASM's Premier Presentation Of Engineering Applications

This journal presents high-level materials science information useful for solving day-to-day engineering challenges in real-time applications.

All aspects of materials selection, design, characterization, processing, testing and performance used in engineering applications that typically result in components for larger systems are emphasized.



Presents you with a clear understanding of the relationships between materials selection, processing, applications and performance. Improvement of materials properties through processes such as casting, forming, heat treating, surface modification and coating, fabrication and the control of these processes is presented. Testing and characterization are demonstrated by mechanical and physical tests, NDE, metallography, failure analysis, corrosion resistance, chemical analysis, surface characterization and microanalysis of surfaces, features and fractures.

Case studies, examples and illustrations using these technologies to achieve innovative applications of engineered materials are highlighted.

900+ pages a year packed with information presented in a variety of features including: technical notes and short subjects; reports from national labs; news from international manufacturing/research centers; guest comments and editorials; corporate newsline; calendar entries; new product and brochure offerings; literature abstracts and news items of recent developments on environmental issues; and archival **peer-reviewed** articles. Advertising rate card available on request.

Journal of Materials Engineering and Performance
• 6 issues • 1 year • ISSN: 1059-9495 • CODEN: JMEPEG

Order No: JMEP-PER

ASM Member	**$50**
ASM Student Member	$25
Non-Member/Multi-User*	$395

- ☐ **Archival, peer-reviewed articles.**
- ☐ **Reviewed technical notes and short subjects.**
- ☐ **Reports from national laboratories.**
- ☐ **News from manufacturing research centers.**
- ☐ **Guest comments and editorials.**
- ☐ **Newsline and calendar features.**
- ☐ **New product and literature announcements.**
- ☐ **Subject index per volume.**

* Institutions where single copy is used by more than one person, and individual non-members.

ATTENTION: LIBRARIANS

For A FREE sample copy of any ASM periodical, call the ASM Member Services Center: (216) 338-5151, ext. 700

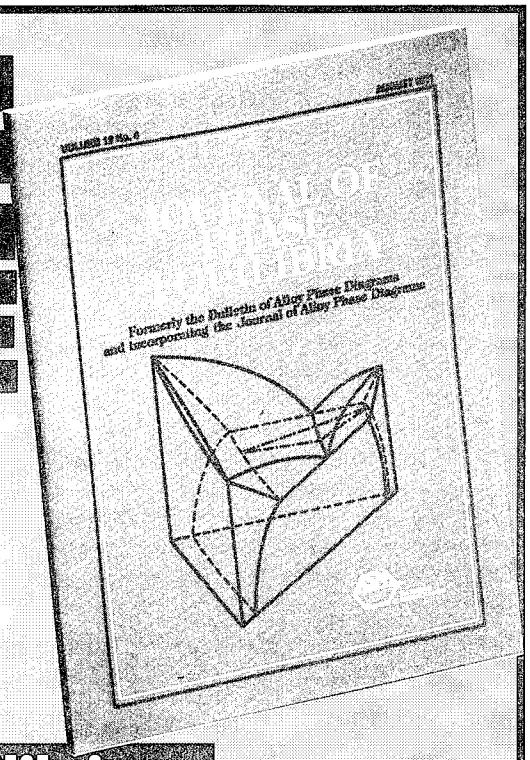
HOW TO ORDER:
(Please provide complete order number.)

PHONE
216-338-5151
EXT. 700

FAX
216-338-4634
ATTN: Member Services Center

MAIL
Member Services Center
ASM International®
Materials Park, OH 44073-0002

FREE SAMPLE COPY



Call Today:
216-338-5151
Ext. 700

Journal of Phase Equilibria
(formerly Bulletin of Alloy Phase Diagrams)

Published bi-monthly by ASM International®
Editor: Jack F. Smith
Editor, Phase Diagram Updates: Hiroaki Okamoto

- Archival, peer-reviewed articles
- Reliable research, applications, updates
- Critical phase diagram evaluations authored by experts

ASM INTERNATIONAL®
The Materials Information Society

Original Research, Applications and Updates

The Bulletin was retitled and expanded in February 1991 to include original research on the generation and application of data to attain or prevent phase equilibria. Content includes:

- Theoretical and experimental research on phase equilibria coupled with other fields in both solidification and oxidation
- Basic and applied phase diagram calculations
- Articles detailing uses of phase diagram data when designing and predicting performance and reliability of alloys for high-tech applications

Please send: **Journal of Phase Equilibria**

☐ FREE Sample Copy ☐ Subscription Information ☐ Author Instructions

NAME _____ COMPANY _____

STREET _____ CITY _____

COUNTRY _____ POSTAL CODE _____

SIGNATURE _____ DATE _____

WRITE: ASM International, MSC/Journals, Materials Park, OH 44073-0002

PHONE: 216-338-5151, Ext. 700

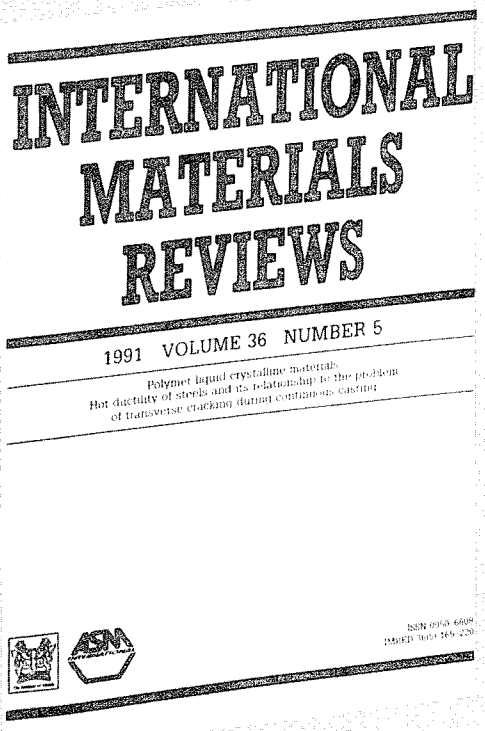
FAX: 216-338-4634, Attn: MSC/Journals

The World's Best Source of Reliable Alloy Phase Diagram Data

Jac493

International Materials Reviews

In-Depth Topical Reviews By International Authorities



Each issue is devoted to in-depth treatment of specific topics ranging from the theory and practice of extraction, production, fabrication, properties and behavior of materials to actual usage.

This unique journal is served by editorial panels in the U.S. and the U.K. who commission international authorities to write the reviews, each of which is dedicated to a specific subject.

International Materials Reviews contains diagrams, tables, charts, photographs and formulas that provide an overview of technical information on a wide variety of subjects in the world of metallurgy and materials science. Each subject is examined thoroughly through highly readable and comprehensively referenced text. Progress reports of recent advances in wider fields of metallurgy are also published. **Peer-reviewed**.

International Materials Reviews
• 6 issues • 1 year • An ASM/IOM Publication
ISSN: 0950-6608 • CODEN: IMRED

Order No: IMR-PER

ASM Member	**$205**
ASM Student Member	$103
Non-Member/Multi-User*	$455

- ☐ **Archival, peer-reviewed articles.**
- ☐ **Reviews authored by international experts.**
- ☐ **Editorial panels in the U.S. (ASM International) and U.K. (The Institute of Materials).**
- ☐ **Diagrams, tables, charts and photographs.**

* Institutions where single copy is used by more than one person, and individual non-members.

ATTENTION: LIBRARIANS

For A FREE sample copy of any ASM periodical, call the ASM Member Services Center: (216) 338-5151, ext. 700

HOW TO ORDER:
(Please provide complete order number.)

PHONE
216-338-5151
EXT. 700

FAX
216-338-4634
ATTN: Member Services Center

MAIL
Member Services Center
ASM International®
Materials Park, OH 44073-0002

STYLING OF REFERENCES

1. A *journal* reference should be readily available on subscription and included in most library collections. Use journal abbreviations as given in the current listing of *Chemical Abstracts Service Source Index*. Article titles are not to be included.
   Example: Author: *Journal*, year, ser., vol., pp.
   R.M. Horn and Robert O. Ritchie: *Metall. Trans. A*, 1978, vol. 9A, pp. 1039-53.
2. References to *books* should include the title and pages within the book.
   Example: Author: *Book*, edition, publisher, place, date, pages.
   George E. Dieter: *Mechanical Metallurgy*, 2nd ed., McGraw-Hill Book Co., New York, NY, 1976, pp. 160-65.
3. A "*private communication*" or "*unpublished research*" may be referenced when required to give proper credit. The citation must include the affiliation and address of the person involved, as well as the year. *Papers Presented at Meetings but not Published* fall under this category.
   Example: J.J. Doe: AAA Company, Washington, DC, unpublished research, 1988.
4. References to *Internal Reports* and other publications of limited availability* are not desirable. However, they will be permitted when the use results in a saving of page space or is required for proper recognition. In these matters, the author's judgment must be supported by the review committee and editor. The report should be available on request and include the source from which a copy may be obtained.
   Example: J.J. Doe: Report No. 738, AAA Company, Washington, DC, January 1988.
5. *In Press* references must include the name of the journal. Balance of reference should be supplied when available. This may be done on the proofs.
6. References such as *submitted for publication* and *to be published* are not acceptable. If the item is still undergoing review, use same format as in 3 above.
7. References must be numbered throughout the manuscript and presented in consecutive numerical order on Reference page.
8. No other references will be published.

*Not available by subscription.

INFORMATION ON PAGE CHARGES AND REPRINTS

METALLURGICAL AND MATERIALS TRANSACTIONS, like many other scientific and technical journals, depends on author page charges for a significant part of its support. Therefore a charge of $75.00 per printed page is levied for all papers published. This charge is $70.00 per printed page for manuscripts submitted on computer diskette (see *Metall. Trans. A*, 1985, vol. 16A, p. 2082). One hundred reprints will be supplied without additional charge for every paper on which the payment of these charges is authorized. It is expected that each author or sponsoring institution will contribute fairly to the support of METALLURGICAL AND MATERIALS TRANSACTIONS through payment of the page charges.

However, in special cases in which the author would have to pay the charges personally, such charges may be waived upon written request. Instructions and forms regarding page charges, waivers, and reprints will be forwarded from the editorial office and appropriate action should be taken at that time. These and any other special matters related to page charges should be referred to METALLURGICAL AND MATERIALS TRANSACTIONS, Carnegie Mellon University, Schenley Park, Pittsburgh, PA 15213.

PUBLICATION OF SYMPOSIUM PAPERS IN METALLURGICAL AND MATERIALS TRANSACTIONS

There are special situations when it is desirable to publish a group of papers concurrently in METALLURGICAL AND MATERIALS TRANSACTIONS. Typically, these arise when a technical committee has organized a limited meeting program of two or three half-day sessions on a particular theme. The material must be original, of high quality, and closely correlated but not sufficient to justify an independent monograph or proceedings-type publication. A high percentage of these would normally be published as METALLURGICAL AND MATERIALS TRANSACTIONS papers, but the authors and committees feel there is an additional benefit from concurrent publication.

To obtain approval, the chairman of the symposium program should write to the editor outlining the symposium scope, the papers to be considered, with abstracts if available, and the plans for payment of page charges. On receipt of this information, the editor will submit a request to the Joint Commission for approval. All manuscripts in the symposium will be reviewed according to METALLURGICAL AND MATERIALS TRANSACTIONS standard review procedure.

An example of a symposium published by METALLURGICAL AND MATERIALS TRANSACTIONS: "Symposium on the Mechanical, Microstructural and Fracture Processes in Superplasticity," *Metall. Trans. A*, 1982, vol. 13A, pp. 688-743.

COPYRIGHT POLICY

METALLURGICAL AND MATERIALS TRANSACTIONS is subject to the U.S. Copyright Law which became effective on January 1, 1978. Accordingly, the signed copyright transfer form must be received in the Editor's Office before the article can be processed for publication. This form, published in the February issue of METALLURGICAL AND MATERIALS TRANSACTIONS A and METALLURGICAL AND MATERIALS TRANSACTIONS B, must be *signed by each author* unless excepted as noted under Part A, "work made for hire" or Part B which pertains to U.S. Government employees. Other copyright transfer forms may not be substituted for this form.

This copyright transfer formalizes the author-publisher relations of professional societies and becomes effective when the manuscript is accepted for publication. If the article is not published in METALLURGICAL AND MATERIALS TRANSACTIONS, the copyright transfer will not take effect. As the copyright owners for articles in METALLURGICAL AND MATERIALS TRANSACTIONS, ASM and TMS-AIME will continue to promote the widest dissemination of the technical information printed in these journals.

ANNUAL AWARDS

TMS—**Champion H. Mathewson Award**: Established to honor Dr. Champion H. Mathewson, President of the Institute in 1943, award is made for a paper which represents a notable contribution to metallurgical or materials science. Paper must have been published in *Metallurgical and Materials Transactions A* or *B* or any other TMS publication within the three-year period preceding the award.

ASM—**Marcus A. Grossmann Young Author Award**: Established to honor Dr. Marcus A. Grossmann, President of ASM in 1944, award is made to suitable author whose age is less than 40 years of age on 1 January of year in which paper was published in *Metallurgical and Materials Transactions*. Paper must have been published in the calendar year preceding that in which award is to be made.

ASM—**Henry Marion Howe Award**: Established to honor Dr. Henry Marion Howe, this is the oldest of the ASM awards and metals. Paper must have been published in the calendar year preceding that in which award is to be made.

POLYMERS/CERAMICS/COMPOSITES ALERT

Get Advanced Materials News in 30 Minutes or Less!

Published monthly, by Materials Information, each 20-page Alert contains around 250-300 summaries, an editorial which highlights the key issues of the month and a loose-leaf titles listing for quick scanning.

Our editors monitor over 2500 trade, business, and technical journals, press releases and newspapers worldwide to assure you comprehensive coverage of advanced materials in a bulletin that can be scanned in 30 minutes or less.

Categories include: ***(Subject sections are divided into polymers, ceramics and composites)***

- Fuel
- Recycling
- Control and Testing
- Environmental Issues
- Applications
- Management
- Marketing
- Resources and Reserves
- Energy Usage
- Plant Developments and Descriptions
- Product and Process Developments
- Waste Treatment
- Competitive Materials
- Training
- Economics
- World Industry News
- Raw Materials
- Engineering
- Machinery
- Health and Safety
- Substitution
- Regulations
- Statistics
- Company Information

Some recent titles include: ***(From February '93 issue)***

Transparent Ceramic is Made From Compacted Ultra-Fine Particles
Isuzu's Ceramic Engine
Advances in Superconducting Magnetic Bearings - an Enabling Technology for Applications of Bulk High Temperature Superconductors
Push for New European Recycling Code
Dow Opens and Closes Plants
Why Not Rebuild Blow Molding Machines?
Composite Surfacing Films for Aerospace Applications
Nanocomposite Materials
EPA Targets Composites

For your
FREE SAMPLE
CALL
216-338-5151 (U.S.)
Ext. 532
or
071-839-4071 (U.K.)

ASM INTERNATIONAL®
The Materials Information Society

Materials Information
ASM International®
Materials Park, Ohio
44073-0002 U.S.A.

The Institute Of Materials

Materials Information
The Institute of Materials
1 Carlton House Terrace
London SW1Y5DB U.K.

Materials Information is a joint service of ASM International and The Institute of Materials