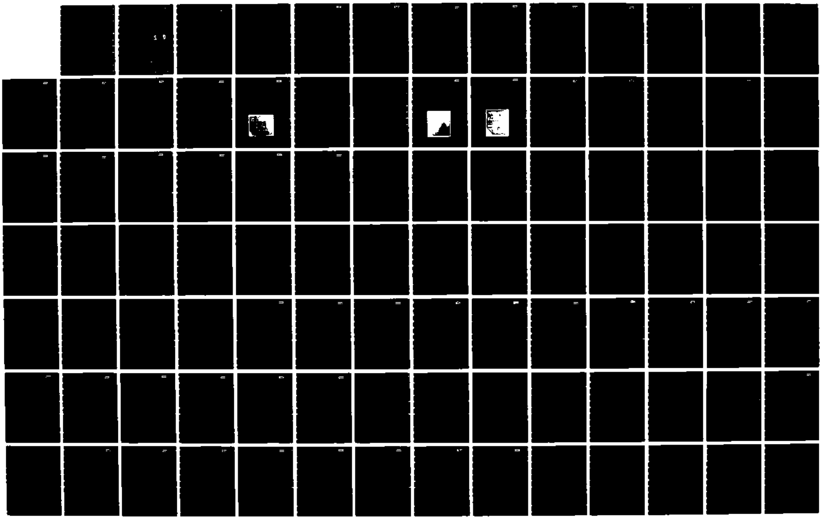
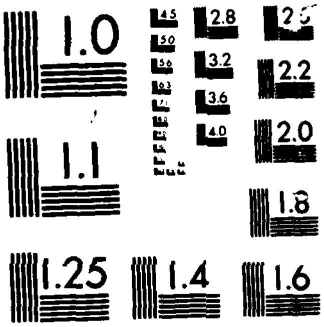
AD-A166 794 HGCDTE SURFACE AND DEFECT STUDY PROGRAM(U) SANTA 1/3
BARBARA RESEARCH CENTER GOLETA CALIF J A WILSON ET AL.
JUL 85 SBRC-60070 MDA903-83-C-0108

UNCLASSIFIED F/G 20/12 NL





MICROCOPY CHART

(2)

HgCdTe SURFACE AND DEFECT STUDY PROGRAM

J. A. Wilson and V. A. Cotton
Santa Barbara Research Center
Goleta, CA 93117

J. A. Silberman, G. P. Carey, A. K. Wahl
D. J. Friedman, C. K. Shih, K. A. Bertness, I. Lindau, and W. E. Spicer
Stanford Electronics Laboratories
Stanford, CA 94305

A. Sher and A.-B. Chen
SRI International
Menlo Park, CA

July 1985

FIFTH INTERIM TECHNICAL REPORT

Contract No. MDA-903-83-C-0108

The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official Department of Defense position, policy, or decision, unless so designated by other official documentation.

Prepared for:
DARPA
1400 Wilson Blvd.
Arlington, Virginia 22209
Attention: Dr. James Murphy
Defense Science Office

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

1378

86/09339

86 4 22 183

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A166 794

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT

4. PERFORMING ORGANIZATION REPORT NUMBER(S) RPT60070

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION SANTA BARBARA RESEARCH CENTER

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION DARPA/DSO

6c. ADDRESS *(City, State and ZIP Code)*
75 Coromar Drive
Goleta, CA 93117

7b. ADDRESS *(City, State and ZIP Code)*
1400 Wilson Blvd.
Arlington, VA 22209

8a. NAME OF FUNDING/SPONSORING ORGANIZATION DARPA/DSO

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS *(City, State and ZIP Code)*
1400 Wilson Blvd.
Arlington, VA 22209

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE *(Include Security Classification)* HgCdTe Surface and Defect Study Program (U)

12. PERSONAL AUTHOR(S) Wilson, J.A.; Cotton, V.A.

13a. TYPE OF REPORT INTERIM TECHNICAL

13b. TIME COVERED FROM 12/31/84 TO 7/1/85

14. DATE OF REPORT *(Yr., Mo., Day)* JULY 1985

15. PAGE COUNT

16. SUPPLEMENTARY NOTATION Authors of papers included in this report: Silberman, J.A.; Carey, G.P.; Wahi, A.K.; Friedman, D.J.; Shih, C.K.; Bertness, K.A.; Lindau, I.; Spicer, W.E.; Sher, A; Chen, A.-B.; and

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)* HgCdTe, II-VI alloys, surfaces, defects, electronic structure, interface, storage time, minority carrier lifetime, channeled implants, surface damage.

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Cont'd

Interface state structures for the HgCdTe PHOTOX© SiO_2 interface was determined by complex admittance spectroscopy while compositional variations were measured by low temperature (77K) electroreflectance. Results are also presented of the dependence of capacitance-voltage behavior on planar defect proximity. Deep states due to ion implant damage were examined with DLTS and atomic profiles measured with SIMS are presented for Al, Mg, Na, Si, P, As, H, Br and Cu. Photoelectron spectroscopy results are presented of the Ag/HgCdTe interface, e beam induced Hg desorption from a cleaved surface and of the surface that results from Br_2 based chemical treatments. Several recent theoretical studies of the local structure of semiconductor alloys are also presented.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL

22b. TELEPHONE NUMBER *(Include Area Code)*

22c. OFFICE SYMBOL

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

(THIS PAGE INTENTIONALLY LEFT BLANK)

SUMMARY

Correlations are made between capacitance-voltage measurements of MIS structures formed on horizontal zone melt grown HgCdTe (x = 0.3) wafers and the devices proximity to grain boundaries. Large changes in flat-band voltage and hysteresis are seen though the density of interface traps is relatively unaffected. The type and amount of changes seen are sufficient to account for the wafer to wafer inconsistancies often seen in C-V measurements. Similar fluctuations in surface properties over epitaxially grown wafers can be attributed to the low angle grain boundaries, which are present acting as undesired interfaces in a similar manner.

Cont'd Pg 1

Low temperature (77K) electroreflectance measurements of the HgCdTe/PHOTOX™ SiO_2 interface reveal a second set of E_1, $E_1 + \Delta_1$ peaks representing a region with effective mole fraction of 0.1 in addition to the x = 0.30 peak representing the bulk alloy. This lowered x value region is interpreted to be distributed as islands near the interface and result from interaction of the processed HgCdTe surface with the deposited PHOTOX™ SiO_2 layer.

The complex admittance of an n-type HgCdTe/PHOTOX™ SiO_2 interface, with x = 0.3, has been examined for frequencies ranging between 1 MHz and 4 MHz. The conductance method is used to decompose the total interface state density into three types of components: a valence band tail, a conduction band tail, and some well resolved discrete states. The fixed charge density is low and there is no statistical broadening. The surface valance and conduction band edges are both found to be shifted upward in energy relative to their respective bulk values; moreover, the surface has converted to p type. The energy variation of the valence band tail states' response times follow a pattern characteristic of Shockley-Reed recombination centers with a constant capture cross section, but the behavior of the conduction band tail states is more complicated. Evidence is presented that the interface region has a higher Cd concentration than the bulk.

Depth profiles of a variety of ions implanted into HgCdTe surfaces of random orientations and channeling into the (110) surface are presented. The ions reported are: channeled Al, Na, Mg, Si, P, As, H, Br and Cu; and random B, Cu and P. Though some of this data has appeared in various other reports on this study it is compiled here because of the range statistics information it contains.

Ion implantation, commonly used to create n^+ on p junctions for photovoltaic detectors in HgCdTe, is known to create a high level of donors due to lattice damage. Deep level transient spectroscopy (DLTS) has been used to study traps in n-type $Hg_{0.7}Cd_{0.3}Te$, associated with the implant induced lattice damage. A new trapping center that is distinctly due to implant damage is found at $E_c - 0.19$ eV. The electron trapping characteristics of this level are unusual, showing a very small capture cross section of $\sigma_n = 5 \times 10^{-21}$ cm^2 and a large barrier potential to electron capture. The large barrier height indicates a weak electron trap and implies a possibly strong hole trapping nature.

The room temperature Ag/(Hg,Cd)Te interface has been studied in ultrahigh vacuum with X-ray and ultraviolet photoelectron spectroscopy. The Ag evaporated onto (Hg,Cd)Te diffuses 10^2 to 10^3Å into the bulk of the semiconductor, displacing Hg as it does so. The bands bend 0.05 to 0.1 eV upward at low coverage. No significant chemical reaction is observed.

Auger electron spectroscopy (AES), X-ray photoemission spectroscopy (XPS), low-energy electron diffraction (LEED), and angle-resolved ultraviolet photoemission spectroscopy (ARPES) were used to study the electron beam induced Hg desorption from a cleaved (110) HgCdTe surface and the electronic structure of the Hg-depleted surface. Solid state recrystallized HgCdTe single crystals were used. It was found that the electron beam heating dominated the electron beam-induced Hg desorption on HgCdTe. At the electron beam energy used, the electron beam heating extended several thousand angstroms deep. However, the Hg depletion saturated after a few monolayers were depleted of Hg atoms. At the initial stage of Hg loss (only 3%), the surface bands bend upward (more p-type). The ARPES spectrum showed the loss of some E versus K dispersion after 22% Hg atoms were removed from the surface region, and no dispersion was observed after 43% Hg atoms were removed. These results have important implications for the electronic structure of the surfaces and for interfaces whose stoichiometry is altered.

Using previous photoelectron spectroscopy measurements of the cleaved HgCdTe surface as a baseline, the extension of our investigations to technologically important processed surfaces of HgCdTe that have been cleaned by chemical etching is important in enabling us to understand the complexities of these "device" surfaces. Composition changes are seen to occur upon etching of the material with bromine based etches. This etchant has shown anisotropic reaction etch rates on HgCdTe, with Cd removal being the fastest. The goal of this research is to determine the composition, degree of contamination, and the production of oxides on these technologically important surfaces and to determine the pinning of the fermi level at the surface ($E_{f,s}$).

Several recent theoretical studies of the local structure of semiconductor alloys are summarized. First, dilute limit calculations of local bond lengths and mixing enthalpies are discussed. These calculations include effects due to both bond length and bond-angle distortions, as well as local chemical rearrangements. Then, a new statistical theory of concentrated alloys is described. Deviations from random alloy distributions (microclusters) are predicted.

Harrison's bonding theory, the valence force field (VFF), and an elastic continuum are combined in a study of the substitution energies Δ_s and local (first-shell) bond lengths d_1 of isoelectronic impurities in semiconductors. Explicit expressions for Δ_s and d_1 are derived, which enable us to absorb measured electric constants into the calculation and to study the chemical effects arising from differences in the covalent radii and polarities. Several models based on VFF alone are also derived for comparison. The full theory and at least five VFF models are found to produce impurity bond lengths in excellent agreement with experiment. The substitution energies are shown to provide good estimates of the mixing enthalpies Ω of pseudobinary alloys and to predict miscibility gaps properly. The chemical shifts in Ω are found to be negative for most cation alloys but positive for anion substitutions.

We have identified the principal microscopic phenomena controlling dislocation densities in bulk grown semiconductors. Based on this understanding, a strategy for selecting materials to reduce dislocation densities is offered. The relevant quantities are calculated from an extension of Harrison's bonding theory, which, with our improved accuracy relates properties of the solids to the constituent atoms' valence electron energy states and wave functions. We report on the alloy composition variation of bond energies, bond lengths, charge redistribution among constituents, vacancy formation energies, dislocation energies, and hardness. Several III-V and II-VI compound semiconductors are treated including GaAs, GaInAs, HgCdTe, and ZnHgTe.

The bond length and energy changes of the constituents of alloys relative to their pure crystal values are calculated from an extension of Harrison's method. It is demonstrated that the already weak HgTe bonds are destabilized by adjacent CdTe, HgS, or HgSe, but are stabilized by ZnTe. It is also argued that dislocation energies and the hardness of semiconductors vary as a high inverse power of the bond length of the constituents. Hence, the shorter ZnTe bond as an additive should improve the structural properties of HgTe and CdTe. Experiments that support these predictions are noted. The electronic transport properties of 0.1 eV bandgap HgZnTe are about the same as those of HgCdTe, and the structural properties of the Zn compound are superior; thus, we conclude that HgZnTe is likely to be the better material for IR devices.

The sensitivity of defect energy levels in semiconductors to the host band structures and impurity potentials has been studied for approximately 30 impurities in CdTe using four different band-structure models. The discrepancies in the defect levels between two different sets of band structures and impurity potentials are found to range from less than 0.1 eV to the whole bandgap (1.6 eV). The band-structure effects are analyzed here in terms of detailed partial densities of states. Examples of contradictory predictions from different band structures are illustrated, and ways to improve the theory are suggested.

The dislocation energies and hardness of semiconductors are calculated by an extension of Harrison's method. It is demonstrated in agreement with experiment that dislocation energies per unit length are proportional to d^{-3}-d^{-9}, where d is the bond length and hardness is proportional to d^{-5}-d^{-11}. The hardness is related to the interaction energies among dislocations. It is argued that dislocation densities of semiconductors will be reduced if they are alloyed with a second constituent that has a shorter bond length. Experimental evidence supporting this strategy is noted.

The bond-length relaxation in pseudobinary alloys can be predicted by a simple radial force model. In tetrahedral structure alloys the bond-length deviation of the solute in the dilute solution is a quarter of the bond-length difference between the two components. This result agrees with the experimental work done on the GaInAs system performed by Mikkelsen and Boyce.

A formula for alloy-scattering-limited electron mobility in semiconductors is obtained for indirect gap systems with multiple band minima. all the input parameters needed are defined explicitly. The drift mobility of SiGe which has a dip at x ~ 0.13 and a broader minimum at x ~ 0.5 is calculated by adding alloy scattering to other scattering mechanisms and correlates well with the measured Hall mobility.

A calculation is presented to explain the anomalous experimental behavior of the Si 2p core-excitation binding energy and linewidth in SiGe alloys. The observed minimum in the linewidth near x ≈ 0.15 can be explained as the result of a competition between intrinsic broadening due to screening and extrinsic alloy broadening. For pure Si, the binding energy is estimated to be 0.15 ± 0.05 eV and the width is shown to be smaller than that observed at x ≈ 0.15.

Starting from realistic band structures of the constituent materials, the electronic structure of SiGe alloys are obtained in the coherent potential approximation (CPA). Various quantities, including the bowing parameter of the fundamental gap and the energies of several optical gaps, the masses, and the linewidths of the E_0 and E_1 transitions, are calculated on the basis of both diagonal and off-diagonal CPA. All of the band-energy and line-width predictions are in good agreement with experiments. Furthermore, the theory yields alloy-scattering-limited electron-drift mobility in qualitative agreement with experimental results.

ACKNOWLEDGMENTS

We gratefully acknowledge the many helpful discussions with T.N. Casselman and C.E. Jones. We especially acknowledge the technical assistance of R.E. Cole and C.R. Curtis for bulk crystal growth and M.S. Langell, V.A. Harper and R.H. Stallings for sample preparation and M.E. Boyd and J.F. Santarosa for technical assistance in recording capacitance data.

We are also appreciative of the many collaborations which have aided this work, specifically; W.A. Harrison (Stanford University), A.B. Chen (Auburn University), F. Pollak (Brooklyn College) and S. Perkowitz (Emory University). Finally, we are grateful for the support and encouragement of the program sponsors R.A. Reynolds and J. Murphy of DARPA/DSO.

CONTENTS

(THIS PAGE INTENTIONALLY LEFT BLANK)

INTRODUCTION

This report presents technical results of the DARPA-supported HgCdTe Surface and Defect Study program (contract number MDA-903-83-C-0108) for the six-month period ending July 1, 1985. The purpose of the program is to study the physics of the alloy system which determine how interfaces are formed, in general, and passivated interfaces, in particular. This knowledge will then act as a basis for developing improved interface technologies. The focus has been on defect formation and the resulting alteration of electronic structure. The study examines HgCdTe surfaces from the atomically clean surface produced by cleaving in ultrahigh vacuum, and the initial stages of interface formation to detailed examination of the final device interface. To probe the device interface we have relied on that PHOTOX™ SiO_2/HgCdTe interface formed in an MIS device. Since we are interested in the effects due to the HgCdTe alloy itself some examination of the technologies involved has been necessary so that the general effects due to fundamental properties, general to the alloy itself, can be isolated.

The program has relied on two principal types of HgCdTe material both grown at SBRC. These are horizontal zone melt and solid state recrystallized ingots. The zone melt process produces very pure ingots due to the sweeping out of impurities by passage of the melt zone through the ingot. N-type doping levels can be in the mid $10^{13} cm^{-3}$ range producing state-of-the-art metal insulator semiconductor (MIS) and photoconductive devices. The solid state recrystallization technique is used to produce the large single crystal sections needed for cleaving in order to produce atomically clean surfaces for examination of the undamaged electronic structure, and oriented wafers for channeled implantation. Measurements of the cleaved surface form a necessary basis for all further determination of alterations of electronic structure. Due to the small (approximately 1 cm) cross sections of the ingots produced by these two techniques, SBRC first developed a liquid phase epitaxial (LPE) growth capability in order to produce larger photovoltaic arrays and has gone on to develop a capability in vapor transport and interdiffusion epitaxy (VTIE) and metalorganic chemical vapor deposition (MOCVD) technologies and very recently has begun a molecular beam epitaxy (MBE) capability in cooperation with the research laboratories at the Hughes Malibu facility. These newer growth techniques are used as a source of a small fraction of the material examined in this study; however, the more mature bulk growth techniques provide a larger self-consistent and more completely characterized material base.

The picture now emerging is that the alloy's interface properties are determined predominantly by the extreme susceptibility of the alloy to defect formation. Defect formation of one kind or another determines heterojunction and passivated interface properties as well as the bulk response to ion implantation. The surface is damaged by prepassivation chemical and mechanical processes, or by application of over layers, such as insulators, or by forming metal contacts. Results of the investigations which contribute to this view are presented in the following four sections of this report. Each section covers results principally from one of the four main tasks of the study: Interface Study task, Channeled Ion Implantation task, Surface Study and Oxidation Chemistry and Alloy Bonding study.

Section I

INTERFACE STUDY

A. INTRODUCTION

This section presents results of the investigation of the interface electronic structure in MIS devices formed on mid-wavelength infrared (MWIR) (x = 0.3) HgCdTe with PHOTOX™ SiO_2 as the applied insulator. The focus is on the properties of HgCdTe and how it forms interfaces in a general sense. Properties of MIS devices made specifically with PHOTOX™ SiO_2 are important as this structure is our primary method of probing the semiconductor surface region. The effects due to the HgCdTe are expected to be general to all passivated interfaces on HgCdTe with only minor and possibly insignificant variations.

The proximity of grain boundaries to MIS devices in polycrystalline wafers of horizontal zone melt grown HgCdTe shows a strong effect on device properties. Virtually all the properties measured by C-V analysis show some variation when the device is made directly on a boundary. Grain boundary proximity is taken as being essentially totally responsible for all of the significant nonuniformities seen in this study, from device to device across a given wafer or from wafer to wafer. Taking grain boundaries as a general representation of the class of all two-dimensional lattice defects we can expect low-angle grain boundaries to be similarly active and to have a corresponding effect on the uniformity of interface properties on epitaxially grown wafers where low-angle grain boundaries are known to exist.

The structure of the HgCdTe in the region of the passivated interface in MIS devices which are entirely free of the effects due to grain boundaries, was investigated using low temperature, 77K electroreflectance (ER) through collaboration with Professor F. Pollak's group at Brooklyn College. These devices were specially prepared with a transparent gate consisting of 60Å of Ti over the actual device, with electrical contact made through a thicker Au-coated lead line contacting the gate. Two sets of E_1, $E_1 + \Delta_1$ structures have been detected. One set corresponds to the transition expected for the bulk crystal; however, the other set is shown to represent a surface region which is laterally inhomogenously distributed in bandgap. It appears that the region nearest the interface with the PHOTOX™ SiO_2 consists of islands of more Hg-rich material. These islands have an effective mole fraction of about

0.1. Examination of the pre-PHOTOX™ SiO_2 coated surface shows no corresponding effect suggesting these lower X value islands are a result of the PHOTOX™ SiO_2 process itself.

Also presented in this section are results of a detailed examination of the electronic state structure at the PHOTOX™ SiO_2 interface using admittance spectroscopy. These measurements were jointly supported by this program and a program from the Air Force Material Research Laboratory. Admittance measures both capacitance and conductance and allows separations of contributions to the density of electronic states according to the states' response time. Thus, distributions of states due to the same source can be identified. Contributions have been identified that are due to conduction and valence band tailing into the surface bandgap and resulting in a surface bandgap effectively narrower than the bulk value. Also seen are well resolved discrete states believed to be due to bonding between Te and some rare species at the interface. Evidence is also presented that the interface region has a higher Cd concentration than in the bulk in contrast to the photoemission measurements of the unpassivated surface which indicate a severe Cd depletion.

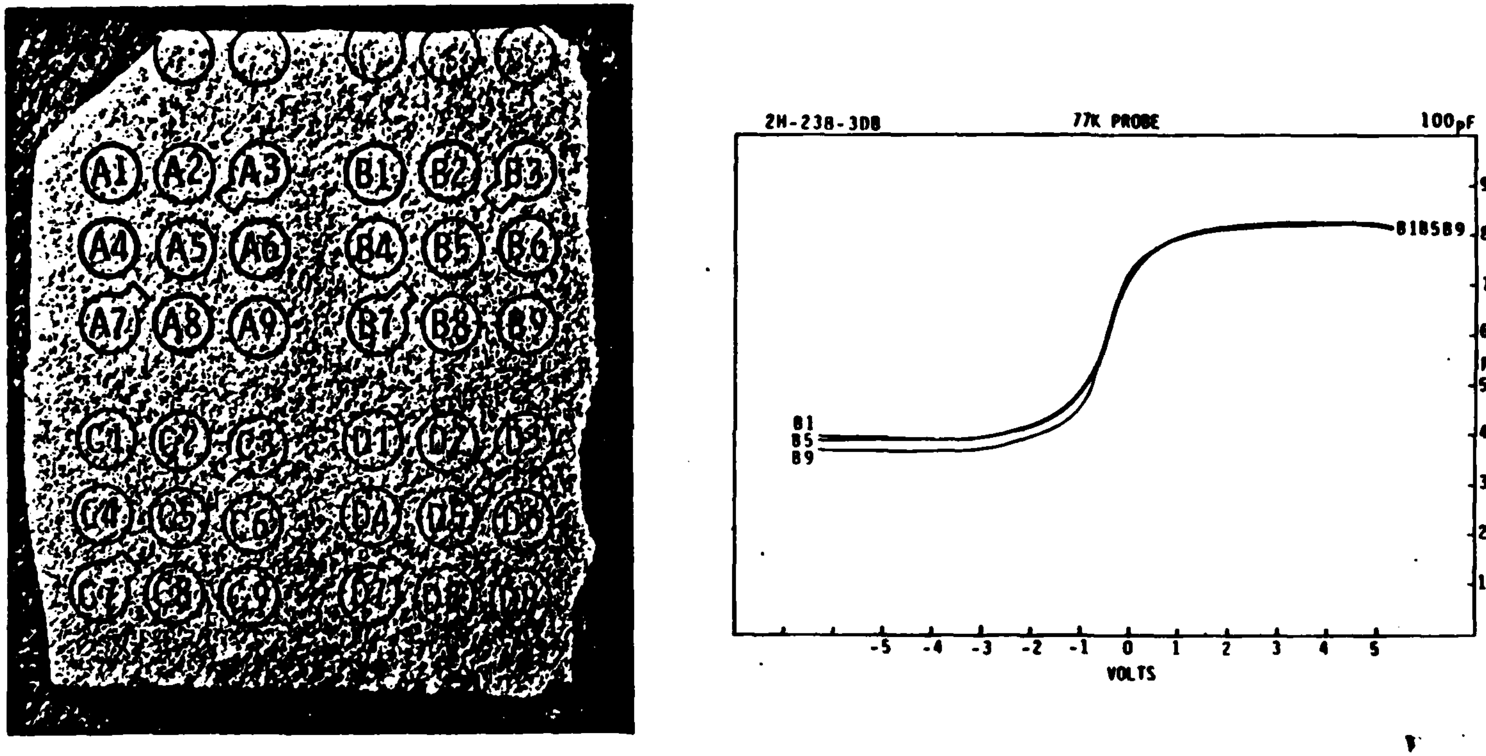
1. INTERFACE CONSISTENCY

Interfaces produced on HgCdTe have been seen to fluctuate in quality from device to device and wafer to wafer in a variety of ways, even when the devices incorporating these interfaces were produced in nominally identical processes. Fluctuations in properties of HgCdTe devices due to apparently uncontrolled parameters of the material or process are seen to be a common thread running through all the different technological approaches to device production. Isolating the specific degrees of freedom unique to HgCdTe is one of the principal tasks of this effort. In MIS devices measured by C-V analysis, for example, wafers cut from some ingots have been shown to produce above average interface properties more often than wafers from other ingots. These properties include a lower or better shaped density (free of discrete defect states) of interface traps, lower hysteresis, more uniform distribution of flat-band bias, or longer storage times.

The devices used in this study have been from the beginning the MIS devices formed using a simple two-step process to clean and polish the surface followed by deposition of PHOTOX™ SiO_2 as the insulator, and a metal gate. No photoresist processes or metal etches are used. This process has been applied consistantly in order to provide a uniform data base to separate fundamental material effects due to the HgCdTe from processing or device effects. The material used has consistently been the highest quality horizontal zone melt grown HgCdTe ($x=0.3$) available at SBRC. Due to the zone melt process, the material is of the highest purity and frequently produces n-type doping in the 10^{13} cm^{-3} range. However, this process does not lend itself well to consistent production of single crystal ingots so that the wafers used frequently incorporate sections of more than one crystallite. The grain boundaries between crystallites can be considered interfaces with properties that combine with or alter those of the passivated MIS interface.

For this study, a number of wafers were selected that had produced regions of MIS devices whose uniformity of properties among neighboring devices ranged from very uniform to very nonuniform. The MIS devices were measured and their positions recorded; then the gates and insulator were removed and the wafer was examined for grain boundaries. A very good correlation was found between proximity of a grain boundary and size of the fluctuation of the interface properties. It was also noted that device properties were very consistent with each other when no grain boundary was present.

Two wafers each were selected from five different ingots. These were cut from the section near the middle of the ingot in order to avoid the slightly higher density of impurities which can occur near the ingot ends. The wafers were all isothermally annealed in Hg vapor which resulted in n-type activity at about 1×10^{15} cm^{-3}. Flat-band bias (V_{FB}), hysteresis (V_{hyst}) and doping density were derived from high frequency (1 MHz) C-V measurements. An average of 40 devices per wafer was surveyed. High-frequency capacitance-voltage curves and wafer photographs for representative wafers from ingots ZH-238, ZH306 and ZH-258 are shown in Figures I-1, I-2, and I-3 to illustrate the association between grain boundaries and C-V curve distortion.



(a)

(b)

Figure I-1. (a) Crystal Surface Showing No Grain Boundaries
(b) Corresponding C-V Curves; Good Uniformity, No Rise in Inversion

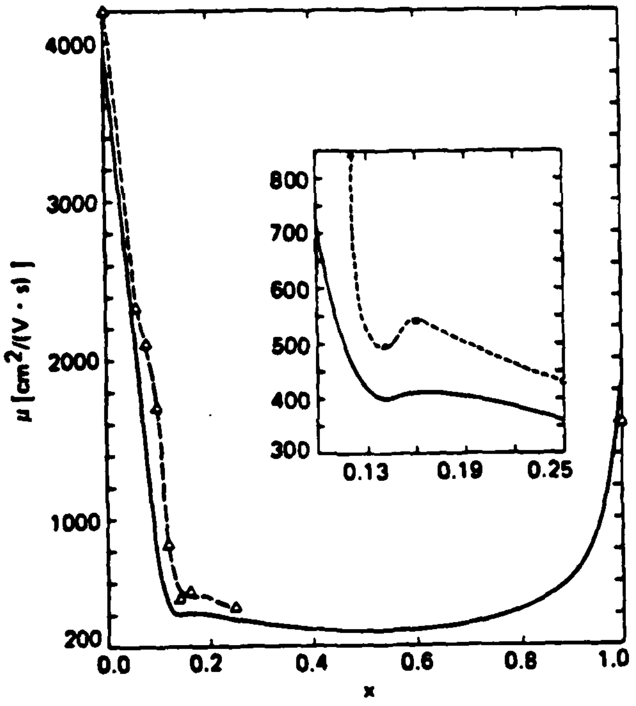

μ [cm^2/(V · s)]
4000
3000
2000
1000
200
0.0
0.2
0.4
0.6
0.8
1.0
x
800
700
600
500
400
300
0.13
0.19
0.25

FIGURE 7

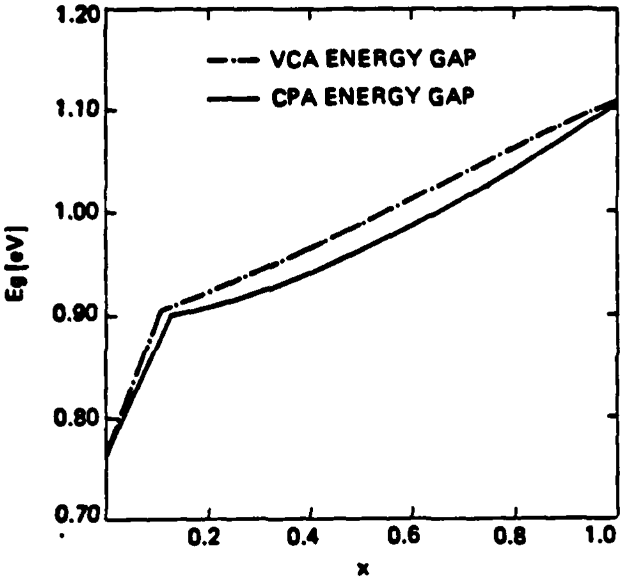
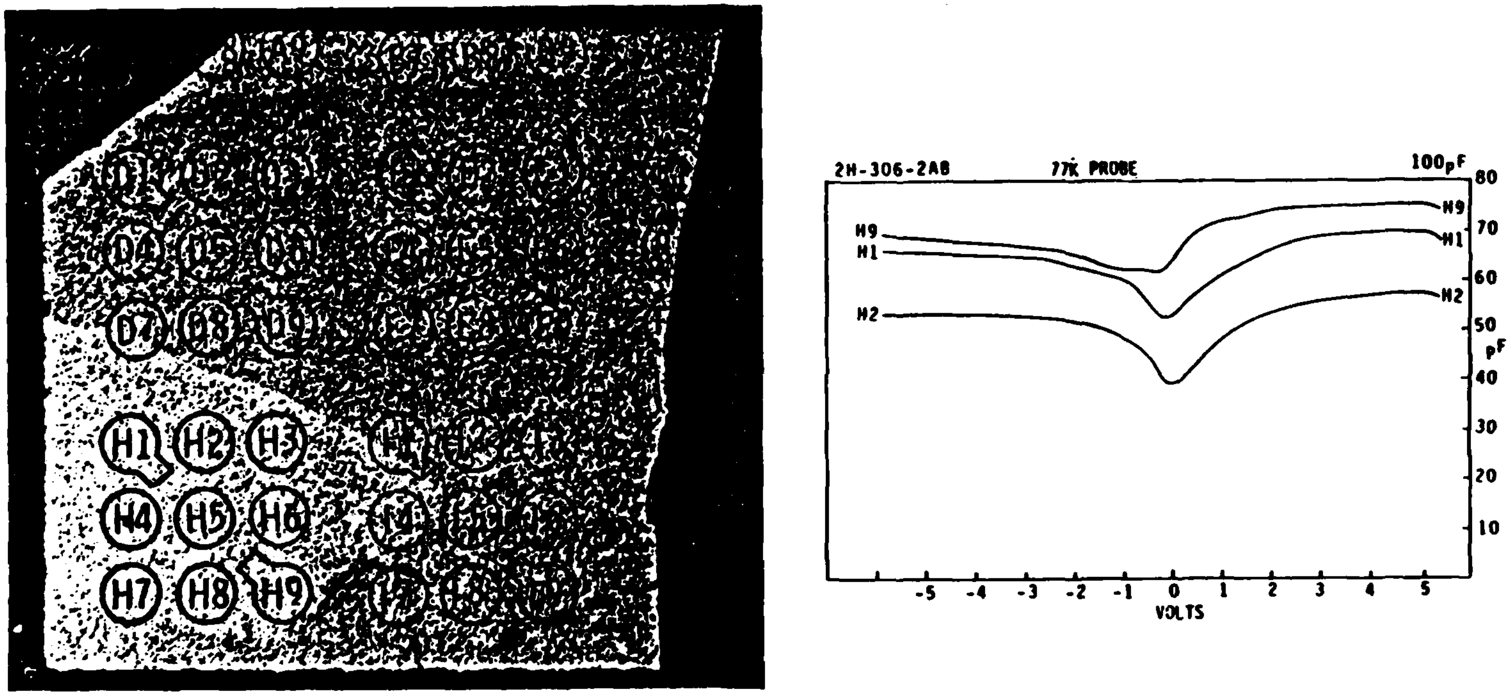

1.20
1.10
1.00
0.90
0.80
0.70
Eg [eV]
VCA ENERGY GAP
CPA ENERGY GAP
0.2
0.4
0.6
0.8
1.0
x

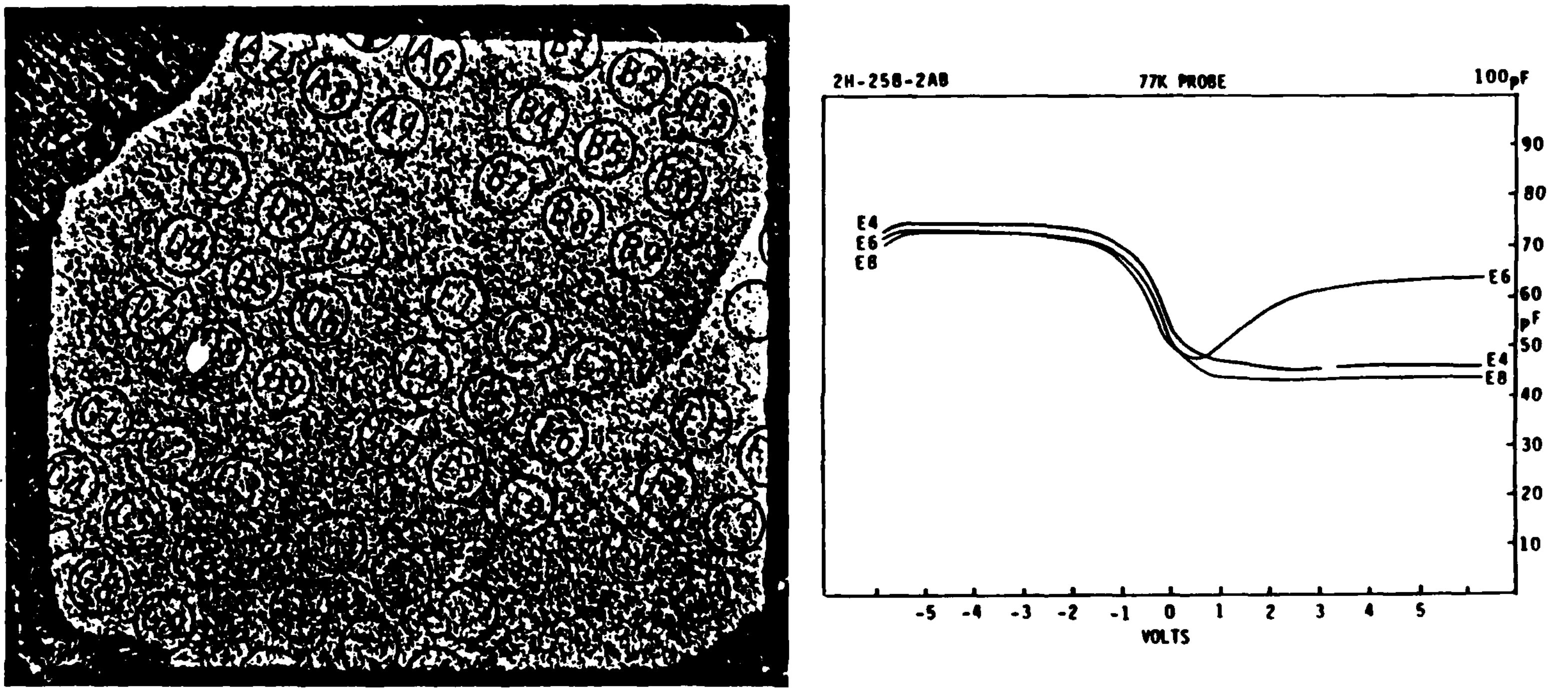
FIGURE 6



(a)

(b)

Figure I-2. (a) Wafer With Multiple Grain Boundaries
(b) Corresponding C-V Curves Showing Low Frequency Type Characteristics and Large Scatter in C_{ins}



(a) (b)

Figure I-3. (a) Wafer Photograph Showing a Single Grain Boundary Through Sections E & H
(b) C-V Curve of MIS E6 Located on the Grain Boundary Compared With Adjacent Devices

It was found that samples with no visible grain boundaries showed good electrical characteristics. As can be seen in Figure I-1 for ZH-238-3DB, the crystal quality on a macroscopic scale is very good with no visible defects or grain boundaries. The corresponding C-V curves are clearly n-type with good uniformity with respect to C_{ins}, V_{FB} and (N_D-N_A). Such uniformity indicates uniform lateral distribution of residual donor impurities and any charge sources within ~ 1 μm of the surface (maximum depletion width for this doping level).

In contrast, samples that showed a multitude of grain boundaries, such as ZH-306 shown in Figure I-2, had low-frequency type C-V characteristics at 1 MHz. There is also significant scatter in V_{FB} and C_{ins}. The low-frequency type characteristics indicate the presence of a charge source that is able to respond rapidly enough to follow modulations of the charge in the inversion layer at 1 MHz. C_{ins} is also seen to be lowered in these instances.

Frequently wafers show large regions of good quality HgCdTe between grain boundaries and it is possible to identify those devices formed on a grain boundary from C-V results alone. Figure I-3 shows the mapping for ZH-258-2AB as an example. It can be seen that capacitor E6 lies directly on the grain boundary. The C-V curve for this device shows a rise in the inversion capacitance not seen in adjacent devices. This again indicates the presence of a rapid generation charge source. A summary of results of the average value of C_{ins}, V_{FB} and ΔV_{FB} is given in Table I-1.

The overall picture that emerges is that the presence of excess electrons and holes, due to the change in lattice orientation and hence incomplete bonding which occurs at a grain boundary, acts as an interface whose characteristics are combined with those of the device interface. This unwanted interface contributes minority carriers in both n- and p-type samples. It is likely that the lowering of C_{ins} is due to a component of the grain boundary capacitance in series with the passivated interface's insulator capacitance.

Similar distortions are observed on oriented epitaxial layers grown on single-crystal substrates which are free of grain boundaries; however, these effects can be expected to result from other examples of two dimensional lattice defects such as low-angle grain boundaries which definitely are present both near the substrate interface and in the bulk of the layer.

Table I-1. Summary of MIS Capacitance Data

Sample	Crystallinity	C_{ins} (pF)	V_{FB}(V)	ΔV_{FB}(V)	Comments
ZH-306-2AA	Multiple grain boundaries	50 ± 5	+0.5 ± 0.5	0.3	low-frequency type curves poor uniformity
ZH-306-2AB	Several grain boundaries	65 ± 2	-0.2 ± 0.1	0.2	n-type uniform
	in single region	64 ± 2	-0.5 ± 0.1	0.2	p-type uniform
	in polyregion	50 ± 4	+0.5 ± 0.5	0.3	low-frequency type curves.
ZH-258-2AA	Several grain boundaries 5 distinct regions	60 ± 10	-1.5 ± 0.5	0.4	mostly p-type many devices with capacitance rise in inversion apparent high N_{ss}
ZH-258-2AB	One grain boundary two regions	70 ± 2	-1.0 ± 0.2	0.3	p-type; uniform devices on or near grain boundary show rise in inversion
ZH-238-3DA	One grain boundary two regions of orientation				
	single crystal region	53 ± 3	-1.0 ± 0.2	0.3	fair C-V uniformity
	near grain boundary	42 ± 6	-1.0 ± 0.3	0.3	low-frequency curves near boundary and near wafer edge
ZH-238-3DB	Good, no grain boundaries	75 ± 3	+0.2 ± 0.1	0.2	n-type very uniform
ZH-245-3BA	One grain boundary in corner	75 ± 1	+0.1 ± 0.1	0.4	very uniform n-type
	corner section	70 ± 4	+0.1 ± 0.2V	0.3	fair, less uniform

Table I-1. Summary of MIS Capacitance Data (Continued)

Sample	Crystallinity	C_{ins} (pF)	V_{FB}(V)	ΔV_{FB}(V)	Comments
ZH-245-3BB	One grain boundary	65 ± 2	+0.2 ± 0.2	0.3	n-type very uniform
ZH-296-3HA	Small off-orientation edge otherwise good single crystal	79 ± 2	+0.1 ± 0.05	0.3	n-type excellent uniformity and low N_{ss}
ZH-296-3HB	Four grain boundaries -small off orientation section in center	78 ± 3	+0.2 ± 0.2	0.3	n-type fairly uniform more dispersion going into inversion and slight rise in inversion near grain boundaries and in center section

As reported in Section I of the previous (Fourth) Interim Technical Report for this program, a fundamental limit to MIS storage time was concluded to occur because of bulk crystal or lattice damage on devices and material similar to those surveyed here. Those samples, however, were selected to avoid the type of C-V characteristics indicating grain boundaries, and indeed no such grain boundaries were noticed on the polished wafer. Certainly grain boundaries between crystallites can, and have been demonstrated[1] to affect minority carrier lifetime and therefore storage time. Low-angle grain boundaries may also be effective and may indeed prove to be the limiting defect, but this hypothesis remains to be investigated.

2. Metal-Insulator-Semiconductor Electroreflectance Study of HgCdTe at 77K.

We have measured electroreflectance (ER) from HgCdTe in the metal-insulator-semiconductor (MIS) configuration at 77K as a function of gate voltage. These are the first ER measurements of HgCdTe in an actual device structure, in a "dry" configuration (i.e., nonelectrolyte)[2,3,4] and at low temperatures. The amplitude of the modulating voltage was 1.6V peak-to-peak at a frequency of 220 Hz.

The devices used in this study were MIS devices made on n-type HgCdTe with a bulk x-value of 0.30 grown by the horizontal-zone melt technique. The resulting wafers were then etched and polished to remove saw damage and isothermally annealed in Hg vapor to reduce the cation vacancies left from the high temperature (600°C) growth. This lowers the native defect acceptor level below that of the residue donor impurities and converts the wafers outer 4 or 5 mils to n-type, in this case about 8×10^{14} cm^{-3}.

The wafers were chemomechanically polished in a solution of Br_2 in DMF (dimethylformamide) and then dip etched in a solution of Br_2 in ethylene glycol to prepare the surface for PHOTOX™ SiO_2 deposition. MIS structures were formed by deposition of 1500Å of PHOTOX™ SiO_2 followed by a 60Å thick Ti gate metal. Electrical contact was made through a 4000Å thick Au bonding pad on a 400Å thick Ti interfacial layer adjacent to the gate and connected through a thin lead line to the gate. The gate diameter is 40 mil (1 mil=25 μm) and results in a measured insulator capacitance of about 180 pF. Only shadow mask techniques were used to fabricate the device structure. No photoresist techniques were employed.

Shown in Figure I-4 by the dotted line is the experimental ER spectrum of sample ZH-245-3GB(5) in the region of the E_1, $E_1 + \Delta_1$ optical features at zero bias voltage. It will be shown that the spectrum actually consists of two sets of E_1, $E_1 + \Delta_1$ structures originating in different regions of the sample. To fit our data we have used the generalized functional form (GFF) proposed in references 2, 3, and 5 to describe the electrolyte electroreflectance (EER) spectrum of the E_1, $E_1 + \Delta_1$ peaks of HgCdTe and related materials. The lineshape L(E) is given by:[2,3,5]

$$L(E) = C_j E^{-2}\left[(\hbar\Omega_j)^3 L(E,5/2) - 4\Delta\sigma^2 L(E,3/2) \right. \\ \left. -4\ \Delta E_j L(E,1/2)\right] \qquad (1a)$$

where $E = \hbar\omega$, p is the number of spectral features to be fit and

$$L(E,N/2) = \cos\,(\theta_j - n\ \phi_j/2)\left[(E-E_j)^2 + \Gamma_j^{\,2}\right]^{n/4} \qquad (1b)$$

with

$$\phi_j = \tan^{-1}[\Gamma_j/(E-E_j) \qquad (1c)$$

Also in equation (1a)

$$(\hbar\Omega_j)^3 = \frac{e^2\hbar^2}{8\ \mu_{11}}\ \varepsilon_j^2 \qquad (2)$$

In the above equations E_j, Γ_j, C_j, θ_j are the energies, broadening parameters, amplitudes and phases of the j^{th} optical peak. The quantity ε_j is the magnitude of the modulating electric field and μ_j is the reduced interband effective mass in the direction of $\vec{\varepsilon}$. The other quantities in Equation (1a) are defined in references 2,3, and 5. The first term in Equation (1a) is the third-derivative functional form (TDFF). Thus, the total amplitude of the TDFF, i.e., $C\ (\hbar\Omega)^3$, is a measure of the modulating electric field.

The solid line in Figure I-4 is a least-squares fit of Equation (1a) to the experimental data using $p = 4$, i.e., two sets of E_1, $E_1 + \Delta_1$ transitions. The two sets of features are denoted ${E_1}^A$, $(E_1 + \Delta_1)^A$ and ${E_1}^B$, $(E_1 + \Delta_1)^B$, respectively, and their energies are indicated by arrows in the figure. In order to demonstrate the need to use two sets of transitions we have tried to fit the experimental data using only one set of E_1, $E_1 + \Delta_1$ features, i.e., $p = 2$. This is shown by the dashed line in Figure I-4. It is evident that two sets of features are required to fit the spectra, particularly the small peak at about 2.3 eV (${E_1}^A$) and the region between ${E_1}^B$, $(E_1 + \Delta_1)^B$. Note that the small peak at about 2.9 eV, which corresponds to $E_1 + {\Delta_1}^A$, is not quite accounted for by the fit, even though its presence must be included.

Since ${E_1}^A$ and ${E_1}^B$ are clearly resolved we will concentrate on these features in the remainder of the discussion. We have studied the dc bias dependence of the ER spectra of these two structures. Shown by the dotted lines in Figure I-5 are the experimental spectra in the region of the ${E_1}^A$, ${E_1}^B$ transitions for bias voltages of -3, -2.5, -2.0, -1.5, -1.0, 0, and + 3V. The solid lines are the least-squares fit to Equation (1a). Listed in Table I are the energies, broadening parameters, and phases for the ${E_1}^A$ and ${E_1}^B$ features at the different bias voltages. Also displayed are the relative amplitude of the third-derivative components of the E_1^B [i.e., C_B $(\hbar\Omega_B)^3$] to E_1^A [i.e., C_A $(\hbar\Omega A)^3$] features, where we have arbitrarily taken C_B $(\hbar\Omega_B)^3 = 1$.

There are a number of interesting features in Table I-1. Based on a temperature dependence of -4×10^{-4} eV/K we can estimate the room temperature positions of E_1^A and E_1^B in order to evaluate the composition x corresponding to these two structures. Thus at room temperature it is expected that $E_1^A \approx 2.21$ eV and $E_1^B \approx 2.43$ eV, corresponding to $x^A = 0.1$ and $x^B \approx 0.37$, respectively.[4,6] It should be noted that although the values of x^A and x^B depend on the assumption about the temperature dependence mode above the difference $(x^B - x^A)$ should be relatively independent of temperature. The value of x^B is in reasonably good agreement with the composition of the bulk material as deduced from the growth conditions.

From Table I-2 $\Gamma_B = 65$ meV and $\Gamma_A = 90$ meV at 77K. Since no systematic study of the temperature dependence of Γ for this material has been reported we cannot make a rigorous comparison with room temperature measurements. We can, perhaps, make an estimate based on Schottky barrier ER[7] work on E_1 and $E_1 + D_1$ of GaAs at 77K and 300K. From reference 7 there is about a 20 to 30 meV broadening of these features between these two temperatures. Thus, we might expect $\Gamma_B \approx 85$–95 meV and $\Gamma_A \approx 110$–120 meV at room temperature. This value of Γ_B is comparable to the best reported Γ for HgCdTe.[2]

The phases of the ER signals can give information about electric field distributions.[8,9] As discussed above, the amplitude of the TDFF is a measure of the modulating electric field $\mathcal{E}$. Also an inversion of the phase θ_B of this term [see Equation (1b)] indicates a change in sign of the nature of the SCR, which can be used as a measure of the flat-band condition.[8,9] Lastras-Martinez et al.,[4] have reported the use of this phase inversion in the EER signal of HgCdTe to determine carrier type changes. Note from Table I-2 that between -1.5V and -1.0V dc bias there is a distinct change in phase for the B feature. This phase inversion is clearly evident in Figure I-5. On either side of this bias range the phase is fairly constant within experimental error. This value of the flat-band voltage is in good agreement with C-V measurements. Although there are some variations in the phase of the A peak there is no clear evidence for a phase inversion.

Table I-2 also shows that there is a pronounced change in the relative amplitudes of the A and B ER features. There is a pronounced increase in this ratio at just those bias voltages (between -1.5V and -1.0V) where the B feature goes through flat band. Since at flat band the electric field in a SCR is zero the bias dependence of the relative amplitudes yields information about the relative electric fields in the two regions.

The above experimental results clearly indicate that there are two distinct regions in these type of MIS HgCdTe samples. There are two ER peaks corresponding to regions of $x \approx 0.1$ and $x \approx 0.37$ with different dependences on the dc bias. The $x \approx 0.37$ region goes through flat-band voltage at about -1.5V to -1.0V while the other region does not go through flat band in the range of dc biases used in this experiment. Since $x \approx 0.1$ HgCdTe is a semimetal it is not surprising that it does not go through flat-band voltage.

One of the main questions to be addressed is how these two regions are distributed, i.e., topographically or laterally. Reference 4 reports that a Br_2/methanol etch may leave a Cd depleted region extending laterally about 600Å into the material. However, the reported variation is only $\Delta x \approx 0.02$ over this region, a factor of ten smaller than the x variation between peaks A and B. Also a semimetallic region of HgCdTe in contact with the oxide would probably short out the device. Thus, the evidence seems to indicate a topographical distribution of the two types of material, i.e., islands of $x \approx 0.1$ HgCdTe dispersed in the $x \approx 0.37$ bulk material.

There have been no reports of such a two-phase system in the 300K EER of Br_2/methanol treated HgCdTe material. Even with the increased Γ at room temperature the separation between E_1^A and E_1^B ($\approx$200 meV) should permit these structures to be resolved in EER. Therefore, the presence of these two distinct regions of different composition may be a consequence of the PHOTOX™ SiO_2 process. This aspect is the subject of future studies.

In conclusion we report the first ER measurements on HgCdTe in a dry configuration and at low temperatures. We observe two distinct sets of E_1, $E_1 + D$, features corresponding to composition of $x \approx 0.1$ and $x \approx 0.37$, the quality of the latter material being quite good. We observe that the $x \approx 0.37$ region has flat-band potential of about -1.0V to -1.5V. There is some indirect evidence that the two phases may be caused by the PHOTOX™ SiO_2 process.

REFERENCES

1. R.G. Pratt, J. Hewett, P. Capper, C.L. Jones, and M.J. Quelch J. Appl. Phys. 54 (1983), 5152.
2. P.M. Raccah et al., J. Vac. Sci. Technol. A3, 138 (1985).
3. P.M. Raccah et al., P.R.L. 53, 1958 (1984).
4. A. Lastras-Martinez et al., J. Vac. Sci. Technol. 21, 157 (1982).
5. P.M. Raccah et al., P.R.L. 55, 1323 (1985).
6. A. Moritani et al., J. Phys. Soc. Japan 34, 79 (1973).
7. M. Chandresahkar and F.H. Pollak, Phys. Rev.
8. D.E. Aspnes in Handbook on Semiconductors.
9. F.H. Pollak, in Photochemistry: Fundamental Processes and Measurement Technique, ed. by W.L. Wallace, A.J. Nozik; S.K. Deband, R.H. Wilson (The Electrochemical Society, Pennington, 1982) p. 608 and references therein.

Table I-2. Bias Dependence of Electroreflectance Structure

	E_1^A FEATURE				E_1^B FEATURE			
Bias (Volts)	E_1^A (eV)	Γ_A (meV)	Phase θ_A (degrees)	Relative Amplitude	E_1^B (eV)	Γ_B (meV)	Phase θ_B (degrees)	Relative Amplitude
-3.0	2.26 ±0.03	99 ±20	-46 ±20	0.39 ±0.2	2.52 ±0.01	61 ±10	38	1
-2.5	2.31 ±0.03	87 ±20	-64 ±20	0.40 ±0.2	2.53	77 ±10	53	1
-2.0	2.30 ±0.03	86 ±20	-58 ±20	0.5 ±0.2	2.52	70 ±10	15	1
-1.5	2.29 ±0.03	94 ±20	-79 ±20	1.48 ±0.2	2.55	67 ±10	23	1
-1.0	2.30 ±0.03	96 ±20	-112 ±20	1.56 ±0.2	2.53	54 ±10	-63	1
0.0	2.26 ±0.03	93 ±20	-180 ±30	0.42 ±0.2	2.54	58 ±10	-75	1
+3.0	2.29 ±0.03	76 ±20	-135 ±30	0.09 ±0.2	2.54	56 ±10	-100	1

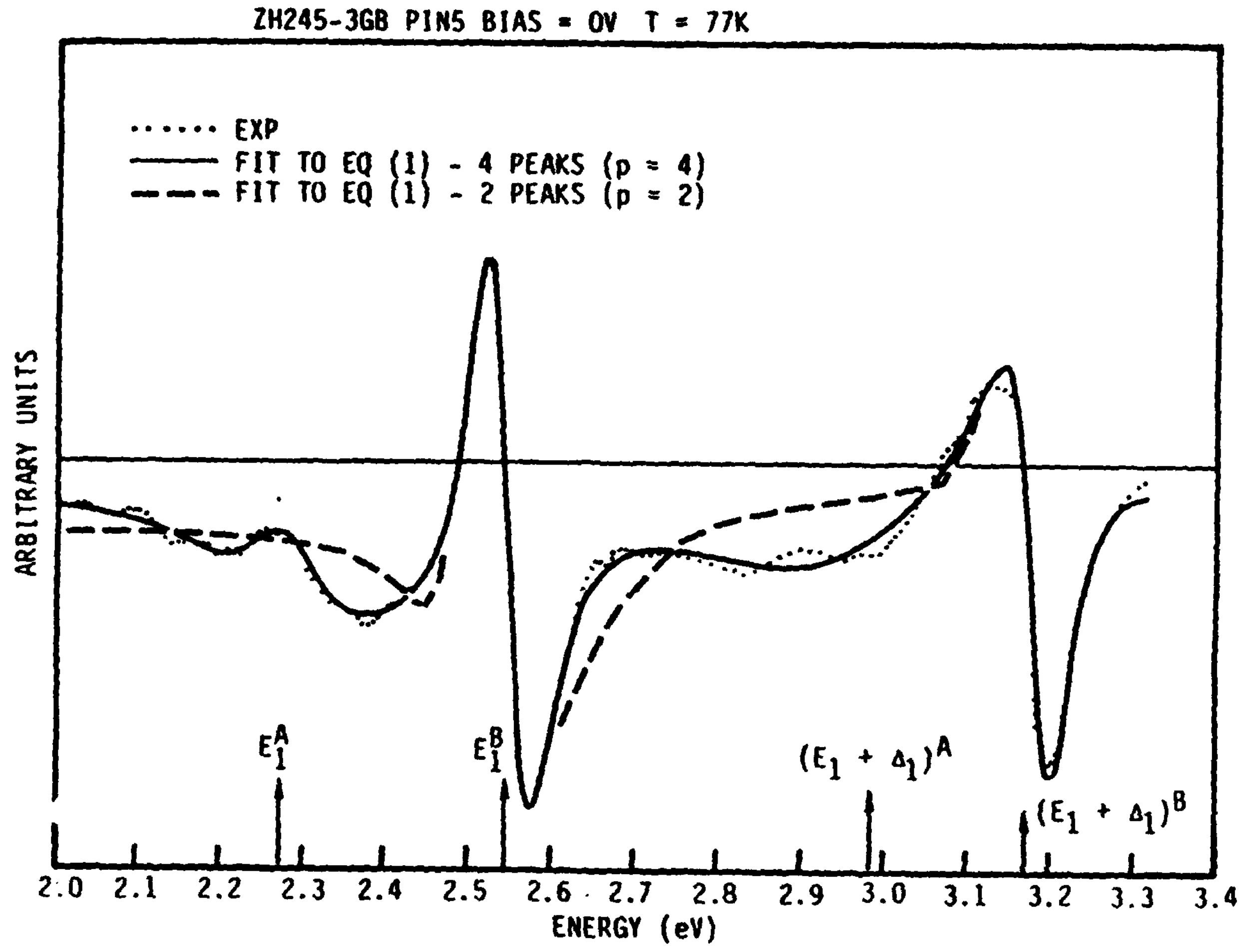


Figure I-4. Comparison of Measured and Calculated Electroreflectance Spectrum for the HgCdTe/PHOTOX™SiO_2 Interface at 77K. The dashed line is calculated using only the bulk mole fraction value of 0.3 while the more accurate solid line uses the bulk mole fraction plus a surface component with x = 0.1. The measured spectrum was recorded at zero applied bias.

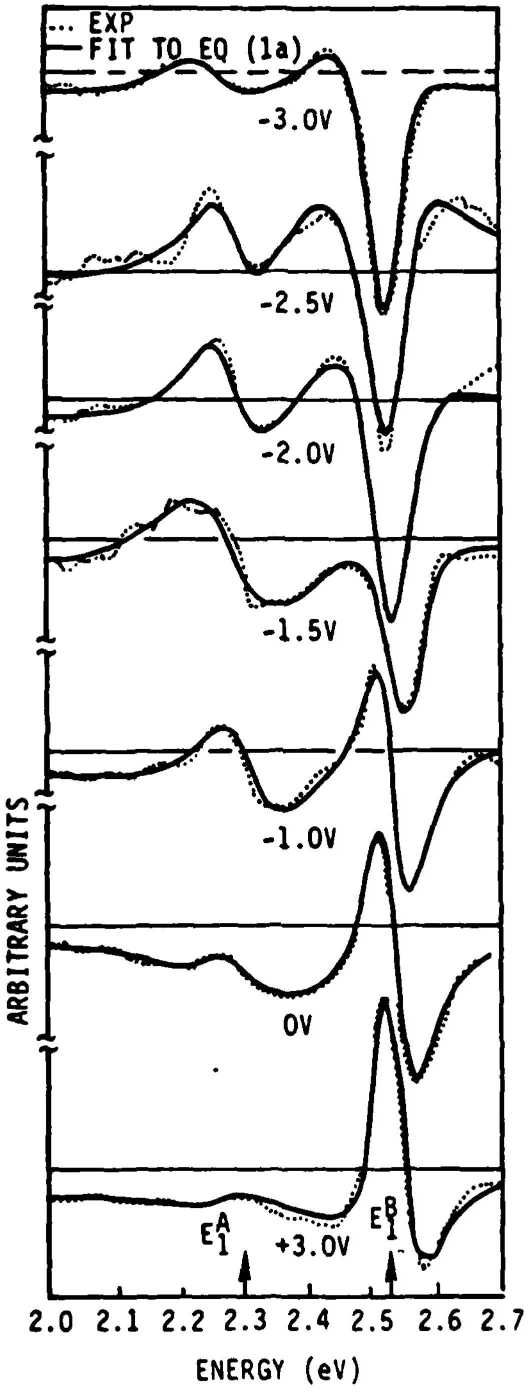


Figure I-5. Comparison of Measured and Calculated Spectra as a Function of Gate Bias. The phase of the peaks representing the bulk alloy can be made to reverse indicating the surface can be inverted. The smaller peak due to the low x surface does not invert at these biases.

(a) Low-frequency Admittance Measurements on the HgCdTe/PHOTOX™ SiO_2 Interface

G.H. Tsau, Stanford University, Stanford, California 94305

A. Sher and M. Madou, SRI International, Menlo Park, California 94025

J.A. Wilson, V.A. Cotton, and C.E. Jones,
Santa Barbara Research Center, Goleta, California

Abstract

The complex admittance of an n-type $Hg_{1-x}Cd_xTe$/Photox SiO_2 interface with $x = 0.3$ has been examined for frequencies ranging between 1 mHz and 4 MHz. The conductance method is used to decompose the total interface state density into three types of components: a valence band tail, a conduction band tail, and some well-resolved discrete states. The fixed charge density is low and there is no statistical broadening. The surface valance and conduction band edges are both found to be shifted upward in energy relative to their respective bulk values; moreover, the surface has converted to p-type. The energy variation of the valence band tail states response times follow a pattern characteristic of Shockley-Reed recombination centers with a constant capture cross section, but the behavior of the conduction band tail states is more complicated. Evidence is presented that the interface region has a higher Cd concentration than the bulk.

I. Introduction

Interface traps located at the $Hg_{1-x}Cd_xTe$/Photox™ SiO_2 interface are one of the more important factors influencing metal/insulator/semiconductor (MIS) infrared detector performance and the effectiveness of this interface to passivate junction devices. These traps interact with free carriers to limit carrier lifetimes, contribute to 1/f noise, and modify threshold gate voltages. This paper characterizes these interface states in greater detail and over a wider frequency range than previously reported.

Many methods have been used to investigate the interface properties of MIS structures; the Si/SiO_2 system that has been studied most extensively.[1-3] While the interface properties of HgCdTe MIS structures have been studied, these efforts have concentrated on the moderate (1 kHz) to high frequency (10 MHz) range.[4-7] In this work, we report the detection of interface traps that have lower emission rate constants than those previously studied. The states examined are distributed throughout the band gap. These slow states are not detectable with conventional high-frequency C-V measurements. Quasi-static voltage sweep methods, which entail voltage changes that are large compared to kT/e, usually involve some current injection into the insulator and are conducted on systems whose interface response times vary over ten or so decades across the band gap. Consequently, even these methods often have inevitable thermal equilibrium problems that prevent accurate determination of the slow states. We report here results of a method that avoids these problems by stepping the gate voltage in discrete increments and allowing the sample to equilibrate before measurements are made at each voltage. The complex impedance is then determined at a set of frequencies, varying from 1 to 4 MHz, a ten-decade frequency range. While measurements were made down to 1 mHz, the onset of ionic conduction in the SiO_2 set a lower frequency limit of 1 Hz to the data reported here.

II. Experimental Procedure

Wafers of $Hg_{1-x}Cd_xTe$ were prepared from horizontal zone melt ingots, with x = 0.3, and isothermally annealed in Hg vapor to reduce the native defect acceptor concentration. This results in n-type wafers with a carrier concentration of $6 \times 10^{14}/cm^3$ at 77 K, measured by high-frequency C-V. A Photox™ SiO_2 layer[9] with a thickness of approximately 1660 A was deposited on the substrate, followed by deposition of a gate metal to form the MIS device. The gate consists of 4000 A of Au on a 400 A layer of Ti and has an area of 0.2 cm^2. Because the material has a small band gap, all experiments were performed at liquid nitrogen temperature to ensure low intrinsic carrier concentrations and to avoid high field dielectric breakdown across the SiO_2 layer.

A Solartron 1250 frequency response analyzer capable of functioning from 10^{-4}Hz to 50 kHz and a Keithly current amplifier whose response is flat to 1 kHz were used in the low-frequency measurements. The device under test and current amplifier were connected in series. Because the input impedance of the current amplifier is very small, the applied (dc or ac) voltage appears mainly across the device. The high-frequency impedance measurements were accomplished using an HP 4061A semiconductor testing system, with four steps/decade from 10 kHz up to 10 MHz.

For both low- and high-frequency measurements, dc voltages were stepped and time allotted for thermal equilibrium to be established before ac signals were imposed to measure the impedance vs frequency.

An MIS device with interface traps can be represented by the equivalent circuit shown in Fig. 1,[8] where R_i and C_i are associated with the interface trap time constants, R_{inv} and C_{inv} model the inversion layer response, $C_{d/a}$ is the depletion/accumulation layer space charge capacitance, C_{ins} is the bulk dielectric insulator capacitance, $C_{ins} = C_{ins1} + C_{ins2}$, R_{ins1} models a Debye type of change in the bulk insulator is high-frequency response, and C_{ion} and R_{ion} are incorporated to account for the ionic conduction induced frequency dependence of the observed insulator capacitance at very low frequencies. For measurements in the frequency range reported here (1 Hz to 4 MHz), R_{ion} can be treated as an open circuit. This is true only when the frequency is greater than 1 Hz.

Inspection of Fig. 2, in which the series capacitance vs gate voltage is plotted for different frequencies (each point is the average of 100 measurements), indicates that in accumulation (positive gate voltages) the insulator capacitance varies from 41.5 nF/cm^2 at 1 Hz to 11 nF/cm^2 at 4 MHz. In the frequency interval from 1Hz to 50 kHz, the capacitance is nearly constant and falls roughly as C_{ins1} $(1 + \omega^2\tau_{ins1}^2) + C_{ins2}$, $(\tau_{ins1} = R_{ins1}C_{ins1})$ between 50 kHz and 4 MHz. For the measured insulator thickness (1660 A), the 1-Hz dielectric constant κ is 7.8, while at 4 MHz it is 2.1.

Therefore, this Photox SiO_2 layer exhibits a dielectric constant that brackets the 3.9 value of a good thermal oxide[3] grown on a silicon substrate. The fact that κ falls below 3.9 at high frequency is consistent with the fact that the Photox-SiO_2 is known to be porous and less dense than thermally grown SiO_2. The higher κ value at low frequency is evidently due to the incorporation of a more polar species. The data fit to a Debye frequency response with a characteristic frequency of 2 MHz. Since these measurements were made at 77°K, this effect is unlikely to be related to the spatial reorientation of a polar molecule.

Figure 2 also shows the dependence of the series capacitance vs gate voltage on measurements made at low frequency. The insulator capacitance that best fits the data is $C_{ins} = 41.5\ nF/cm^2$. The steps in the 1-Hz curve at negative voltage are real and reflect interface state properties. The Berglund method[10] was used to determine the surface potential $\Psi_s(V)$ variations with gate voltage V to within a constant $\psi_s(V_o)$. The constant was fixed (see Fig. 3) using the intercept of the square of the reciprocal of the depletion layer capacitance vs $\psi_s(V)-\psi_s(v_o) + kT/e$ curve method. The depletion layer capacitance was deduced from the high-frequency series capacitance with the insulator contribution removed. The resulting surface potential, $\psi_s(V)$, is plotted in Fig. 4. Notice that the flat band voltage for this sample is -0.10 V. The carrier concentration (N_D-N_A), of the sample is found from the slope of the C_d^{-2} vs $\psi_s(V)-\psi_s(V_o) + kT/e$ plot,[11] and is $n = 6.2 \times 10^{14}/cm^3$. Using an intrinsic carrier concentration[12,13] of $n_i = 1.15 \times 10^9/cm^{-3}$ at T = 77 K and electron and hole effective masses $m_e^* = 0.0182m_o$ and $m_h^* = 0.55\ m_o$, respectively, where m_o is the free electron mass, the bulk Fermi energy is $E_F = 0.226$ eV above the valance band edge. The band gap is 0.24 eV at 77 K. The Fermi energy at the surface (interface) is $E_{FS} = E_F + e\psi_s$.

Following the analysis method in Ref. (8), Eqs.(8) and (9) from that paper are used to relate the measured capacitance and dissipation factor at each frequency to C_p and G_p/ω (with the insulator impedance removed). The solution to the equivalent circuit shown in Fig. 1 is

$$C_p = \sum_i \frac{C_i}{1 + \omega^2 \tau_i^2}$$

and

$$\frac{G_p}{\omega} = \sum_i \frac{C_i \omega \tau_i}{1 + \omega^2 \tau_i^2} .$$

Since Eqs. (1) and (2) satisfy the Kramers-Kronig relations, both contain the same information and we have elected to extract most information by examining G_p/ω.[2] Exceptions will be introduced presently. Figure 5 has plots of G_p/ω vs frequency for different gate voltages in two frequency intervals. The low-frequency data taken on the Solartron extends over the range of 1 to 10^3 kHz. The data in the range from 10 kHz to 4 MHz were taken on a Hewlett-Packard network analyzer. No G_p/ω data are reported in the range 1 kHz to 10 kHz. The dashed curves sketched in that frequency interval are tentative results. The problems in this frequency interval stem from a combination of three circumstances. The first is that because the frequency response of the current amplifier we use rolls off above 1 kHz, we must switch to another circuit arrangement for frequencies from 1 to 50 kHz, and determine the current by measuring the voltage across a standard placed in series with the unknown. Because the "standards" have a slightly uncontrolled frequency variation, in part arising from stray wiring capacitance and inductance, this is a less satisfactory arrangement than the current amplifier. Second, the real part of the impedance is small in this frequency interval, making it susceptible to noise. Finally, as will be seen in Fig. 7, the interface states that respond in this frequency interval have very sharp peaked energy dependences and are spaced by only ~ 5 mV. This restricts the maximum ac amplitude that can be imposed and decreases the signal-to-noise ratio. This combination of circumstances has caused the G_p/ω data in the frequency interval 1 to 50 kHz to be noisy and somewhat unreliable. Therefore, the trap states responding in this frequency interval are deduced from the variation of C_p, as discussed below.

The curves in Fig. (5) were fit to Eq. (2) and values for C_i and τ_i deduced. The fits are quite good with the difference between the experimental results and the fits being of the order of 5%. The differences could be attributed to additional smaller interface trap features, but our current measurement and theoretical accuracy is not good enough to quote them with confidence. Once we have characterized the insulator dispersion well enough that meaningful information can be extracted, we will extend the measurement to lower frequencies. Indications of the existence of trap activity at lower frequencies can be seen at all gate voltages in Fig. 5, since the slopes of nearly all the G_p/ω curves at the lowest frequency are negative.

The state densities deduced from these measurements are summarized in Fig. 6, where they are plotted as a function of E_{FS}. The curve marketed total is deduced from $C_p = N_{ss}/e$, where $C_p + C_d$ is determined from the low frequency C vs V behavior, and C_d from the high-frequency response. In both cases, the insulator capacitance has been removed prior to the subtraction. The curves labeled τ_1 and τ_2 are deduced from the data in Fig. 5. Because the τ_1 and τ_2 curves resemble continuous state distributions rather than discrete traps, we have used the appropriate factor relating C_1 deduced from Eq.(2) and the interface state density. This relation[8] is $(N_{ss})_1 = 1.24$ C_1/e. If this 1.24 correction factor is not applied, the quantitative results are modified slightly, but none of the trends are affected. The difference curve is the total minus the curves labeled τ_1 and τ_3. Most of the sharp features in the data are present in this difference curve. While (for the reasons discussed previously) we do not have sufficiently reliable data to report them directly, this collection of features appears to occur with response frequencies in the interval 10^3 to 5×10^4 Hz, which we have identified with response times labeled τ_2 and τ_2'.

In Fig. 7, the response times deduced from the data are plotted on a logarithmic scale against the surface Fermi energy E_{FS}. The slow response times τ_1 are seen to follow a simple activated behavior varying as $e^{-E_{FS}/kT}$, as would be expected from Shockley-Reed recombination[10] with a constant capture cross section. The capture cross section deduced from the

τ_1 curve is 1.2 x 10^{-15} cm^2. However, the response times τ_3 do not follow this type of Shockley-Reed behavior.

III. Results

The interface trap density profile in Fig. 6 obviously exhibits a wealth of features. A major result of this study is the fact that the experimental method we are reporting allows them to be observed so clearly. The features cannot be identified uniquely with the phenomena giving rise to them from this study of a single sample; however, the shapes of the curves are suggestive and some tentative conclusions can be drawn.

A. Flat Band

The flat band voltage shift V_{FB}, the work function difference between the Au gate metal and the semiconductor ϕ_{ms}, the insulator capacitance per unit area C_{ins}, and the interface charge density Q_o are related by the expression[11]

$$V_{FB} = \frac{\phi_{ms}}{e} - \frac{Q_o}{C_{ins}} .$$

The effective work function difference between x = 0.30 $Hg_{1-x}Cd_xTe$ and the Ti gate in the types of MIS structures examined here has been measured to be very near zero, $\phi_{ms} = -0.03$ eV.[9] The measured flat band voltage for this device is -0.10 V, and $C_{ins} = 41.5$ nF/cm^2. Thus, the interface charge density is $Q_o = 3 \times 10^9$ coul/cm^2 or 2×10^{10} charges/cm^2. This is consistent with the small range of interface charge previously seen for this interface,[9] but is lower than that reported for other insulator interfaces on HgCdTe.[4,5,7] The value of the effective HgCdTe work function found to apply to the Photox^{TM} interface (4.3 eV) is slightly lower than that estimated for the bare cleaved surface (5.3 eV).[15] This may be due to surface alterations resulting from chemical treatment, or merely a difference in crystal orientation and by itself cannot be taken as significant.

B. Total Trap Density

The general U-shape of the total trap interface density is that typically recorded[3]; however, in this case the minimum is located on the conduction band side of mid-gap, and the curve rises much more steeply toward the conduction band edge than toward the valence band edge. Moreover, the surface Fermi energy in accumulation actually penetrated ~ 15 meV above the conduction band edge. This is possible in HgCdTe because the low-conduction band effective mass $m_c^* = 0.018\ m_o$ causes the state density at the band edge to be quite low. The surface valence band edge lies 105 meV above the bulk value, and the surface conduction band edge appears to lie 25 meV above its bulk value. Alternative interpretations of this behavior will be introduced, but the simplest one is that the band gap at the surface has narrowed from the bulk value of 240 meV to a surface value of 150 meV. The conduction band edge may be considerably higher than its "apparent" value, since the low conduction band effective mass means these electrons can tunnel through thick layers. We have no measure of the thickness of the surface region, except it must be thin compared to the depletion layer thickness for the space-charge analysis to have fit theory. Moreover, the effective heterojunction alignment is such that the surface is more p-type than the bulk, since the Fermi level at the surface is at most 75% of the way between the valence and conduction band edges compared to 94% in the bulk.

C. Band Tails

The states in Fig. 6 labeled by response times τ_1 and τ_3 are deduced from the G_p/ω data. These curves have relatively smooth variations with energy and appear to be valence and conduction band tails, respectively. The energy variations of the corresponding measured response times τ_1 and τ_3 are presented in Fig. 7.

The states labeled τ_1 follow the exponential behavior of a Shockley-Reed center with cross section 1.2×10^{-15} cm^2, hence have the same origin. These dominate as the surface valence band edge is approached and are evidently tail states.[8] The integrated area under this state density from 105 to 185 meV is ~ 3×10^{11} states/cm^2. The average surface atom density is 3.2×10^{14} atoms/cm^2; hence, the total valence band tail corresponds to only one state every 10^3 atoms. The corresponding number for the conduction band tail, the states labeled τ_3, is 7×10^{10} states/cm^2 or one state every 2×10^4 atoms. Because the valence band effective mass is much larger than that of the conduction band, one expects a larger valence band tail.

The conduction band tail states have two noteworthy features. First, the time constants τ_3 do not follow the functional form of a Shockley-Reed center with a constant captive cross section. We have no satisfactory model to explain this observation. However, the effect may be related to the fact that conduction band electron from the bulk must tunnel through a potential barrier at the surface to communicate with the conduction band tail states. Second, the slope of the conduction band tail changes abruptly at E_c; the expected bulk conduction band edge. Below E_c the slope is large and the tail points toward a high value at E_c, a new tail then appears to develop for $E > E_c$ with points directed toward the state density identified with the surface conduction band edge. Since the two-dimensional state density is a constant $\sigma_c(E-E_c) = m^*/\pi h^2 = 4.18 \times 10^{14}$ (m^*/m_o) states/eV-cm^2 independent of the energy $E-E_c$, the observed behavior $\sigma_c = 2 \times 10^{12}$ states/eV-cm^2 is consistant with $m^* = 0.005\ m_o$. This value is smaller than the $0.018_o m$ value determined from conductivity and Hall measurements. However, we are sensitive to somewhat different band structure quantities than those dominating transport properties. In addition, a thin region of wide gap material separating the bulk from the interface would lower the effective σ_c at the interface. A detailed theory of the surface is needed to sort through these questions properly.

D. Discrete States

The curves labeled τ_2 and τ_2' are the differences between the total interface density curve and those labeled τ_1 and τ_3. These features have frequency responses in the range 1 to 50 kHz. They have peaked structures characteristic of localized discrete states.[8]

The lower energy structure, labeled τ_2, consists of four sharp peaks. The lowest peak has a height of 7×10^{11} states/eV-cm^2 and a full width at half maximum of ~ 5 meV. Thus, the area under this peak is ~ 4×10^9 states/cm^2, corresponding to about one state every 10^5 surface atom sites. The peaks are nearly equally spaced, with a separation of ~ 15 meV. The third peak appears to be wider than the others. The fact that there are four peaks, their separation is ~ 15 meV, and they are quite narrow suggests their origin. We speculate that these lines are due to dangling Te bonds or bonds to some rare species connected to Te atoms at the interface. The other three Te bonds are attached to either Hg or Cd cations. The four peaks correspond to Te atoms connected either to three Hg, two Hg, and one Cd; one Hg and two Cd; or three Cd, respectively. The energy differences separating these different microclusters has been shown[15] to be ~ 25 meV if the bond whose energy is being shifted is either a HgTe or CdTe bond. Therefore, energy separations of ~ 15 meV are the right order to be due to clustering and shifts caused by metallization terms. The lowest energy state should be the one on which there are three Hg atoms, since it has been demonstrated[16] that Cd always shifts adjacent bond energies more positively than Hg atoms do. Thus, according to this model, the lowest state is interpreted as a Te bond at the interface with three neighboring HgTe bonds. Since this state has the largest density, it implies that at the surface is HgTe rich as expected. However, the area ratio of the lowest and highest states is roughly 6.9/3.6 = 1.9 and, for a random $Hg_{1-x}Cd_xTe$ alloy, the probability of a cluster with m Cd atoms is

$$P_m = \frac{3!}{(3-m)!m!} (1-x)^{3-m} x^m .$$

Using this relation, one deduces a Cd concentration equivalent to that in an $x = 0.45$ alloy for the immediate surface. This composition corresponds to a considerably wider energy gap (~ 0.7 eV) than our direct observation indicates. However, since the surface region composition appears to be severely altered from an actual alloy stoichiometry, estimating a surface region energy gap from empirical expressions relating energy gap to x value derived for stoichiometric alloy may not be entirely accurate. In any event, it is clear from this study that the surface valence band maximum is offset to higher energies. The offset is consistent with either a vacancy or Cd-rich surface, since vacancy and Cd energies lie above those of Hg and the valence band edge is nearly all p-state (mostly Te, but with a significant cation admixture).

The higher lying states, labeled τ_2', have a total of ~ 7×10^{10} states/cm^2; although this is small compared to the surface atom density, it is more than a factor of 10 larger number than the τ_2 states. These states are densely packed and are mostly unresolved. Judging from their higher energy position and density, these states are likely to be associated with bonds between cations and the interface. The random substitutional species are then on the sites responsible for the states, and are affected by the composition of clusters including 9 to 12 second neighbors. Thus, the number of possibilities is large and the separation between the lines arising from different clusters will be small. Once again, a complete surface theory is needed to understand the implications of these data.

We have demonstrated that the HgCdTe/Photox SiO_2 interface exhibits a wealth of features. The particular oxide studied is an ionic conductor at low frequencies, and has a high frequency dispersion of its bulk dielectric constant with a characteristic frequency of ~ 2MHz. Both the surface valence and conduction band edges lie above their bulk counterparts. The surface is more p-type than the bulk. Continuum band tail states and discrete localized states are observed. Aspects of the data are consistent with the interface region being more Cd rich than the bulk.

Acknowledgments

We wish to acknowledge partial support for this work from DARPA Contract MDA 903-83-C-0108 and AFVAL/ML Contract F33615-82-C-5122). We especially wish to thank Professor C. W. Bates for helpful comments, and S. Gaisford for help with the software programming.

References

1. E.H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. 45, 1055 (1967).

2 A. Goetzberger, E. Klausman, and M.J. Schultz, CRC Crit. Rev. Solid State Sci. 6, 1(1967).

3. E.H. Nicollian and J. R. Brews, MOS Physics and Technology (Wiley, New York, 1982)

4. W.F. Leonard and M. Michael, J. Appl. Phys. 50, 1450 (1979).

5. Y. Nemirovsky and I. Kidron, Solid-State Electronics 22, 831 (1979).

6. N.E. Byer, T.S. Sun, and S.P. Buchner (unpublished results).

7. M.A. Kinch, Semiconductors and Semimetals, ed. R.K. Willardson and A. Beer, Vol. 18 (Academic Press, New York, 1981), p. 313.

8. A. Sher, H.J. Hoffman, P. Su, and Y.H. Tsuo, J. Appl. Phys. 54, 5183 (1983).

9. J.A. Wilson and V.A. Cotton, J. Vac. Sci. Technol. A 3, 199 (1985).

10. C.N. Berglund, IEEE Trans. Electron Devices ED-13, 701 (1966).

11. S.M. Sze, Physics of Semiconductor Devices (Wiley, New York, 1969).

12. G.L. Hansen, J.L. Schmit, and T.N. Casselman, J. Appl. Phys. 53, 7099 (1982).

13. G.L. Hansen and J.L. Schmit, J. Appl. Phys. 54, 1639 (1983).

14. A. Sher, Y.H. Tsuo, J. A. Moriarty, W.E. Miller, and B.K. Crouch, J. Appl. Phys. 45, 2137 (1980).

15. T.N. Casselman, A.Sher, J. Silberman, W.E. Spicer and A.-B. Chen, J. Vac. Sci. Technol. 1, 1735 (1983).

16. A. Sher, A.-B. Chen, W.E. Spicer, and C.-K. Shih, J. Vac. Sci. Technol. A3, 105 (1985).

Figure Captions

1. Equivalent circuits of the metal/insulator/semiconductor system. (a) The insulator capacitance in series with the parallel combination of the depletion/accumulation capacitance $C_{d/a}$, and the interface and inversion layer modeled by frequency-dependent capacitance C_p and conductance G_p. The insulator capacitance C_{ins} accounts for the ionic and bulk dielectric response of the insulator. (b) A more detailed model where G_p, C_p in (a) are decomposed into individual inversion layer and interface state responses.

2. The equivalent total series capacitance C_s vs applied voltage V plot at various frequencies.

3. $1/C_d^2$ vs $\psi_s(v)-\psi_s(v_o) + kT/e$ plot, where C_d is the high-frequency depletion layer capacitance and $\psi_s(v)$ is the surface potential at the applied voltage V. The intercept at the horizontal axis is the surface potential additive constant $\psi_s(v_o)$.

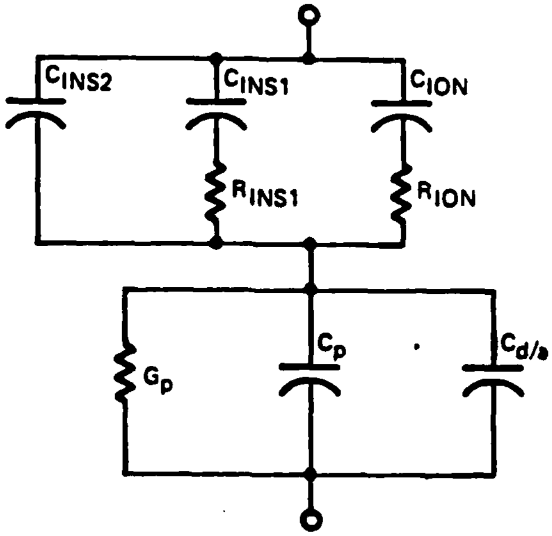
4. The surface potential $\psi_s(v)$ as a function of the applied dc bias V.

5. Parallel conductance divided by the angular frequency as a function of the applied ac signal frequency.

6. Interface state densities as a function of energy relative to the valence bandedge. The curves are :(i)total interface state density curve obtained from the high- and low-frequency measurements, and (ii)

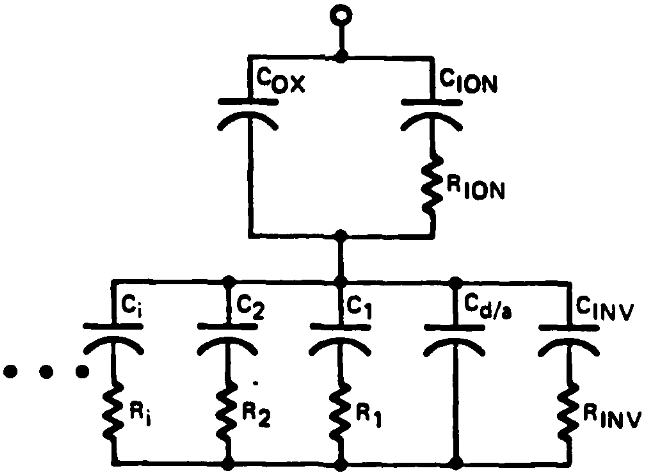
interface state densities curves determined from the conductance method.

7. The interface state response time constants as a function of energy in the silicon band gap.



C_INS2
C_INS1
C_ION
R_INS1
R_ION
G_p
C_p
C_d/a

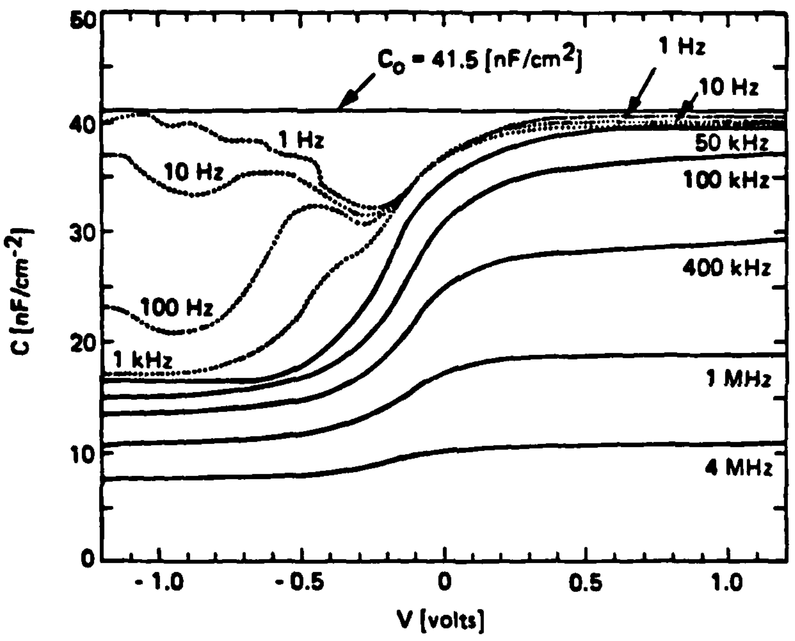
(a)



C_OX
C_ION
R_ION
C_i
C_2
C_1
C_d/a
C_INV
R_i
R_2
R_1
R_INV

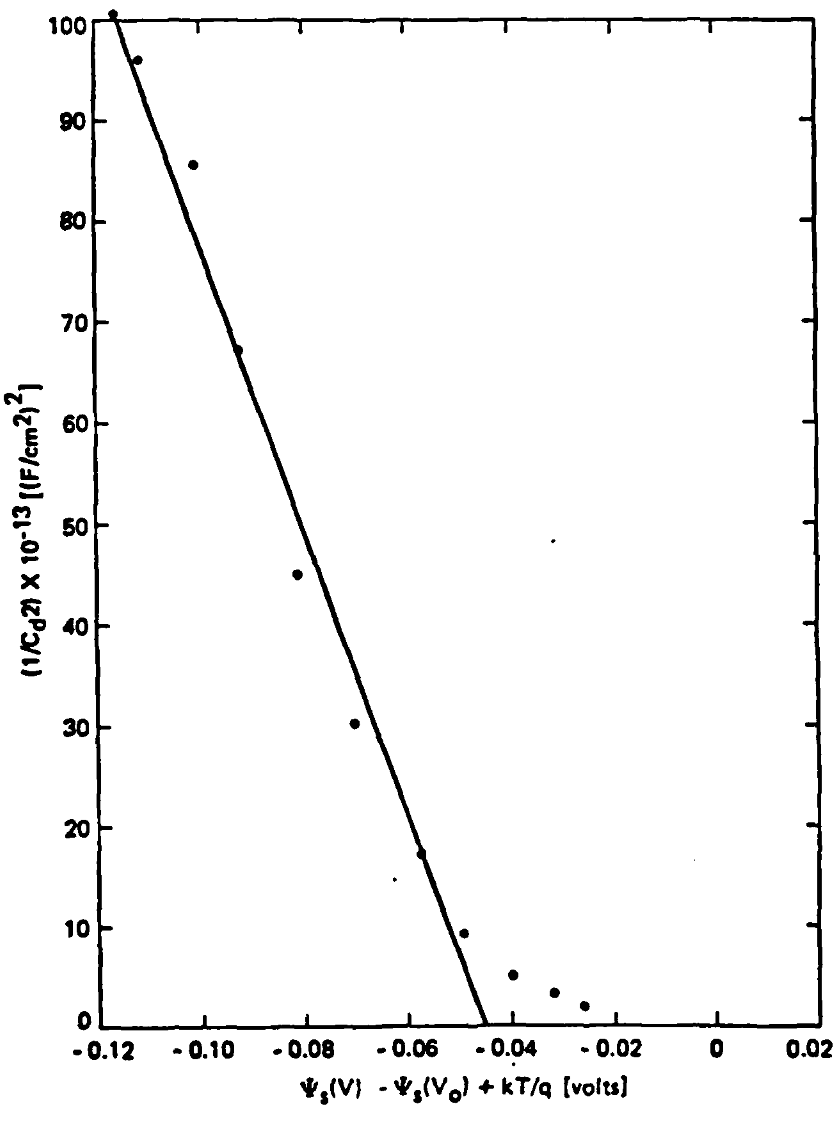
(b)

Figure 1



C_o = 41.5 [nF/cm^2]
1 Hz
10 Hz
1 Hz
10 Hz
50 kHz
100 kHz
100 Hz
1 kHz
400 kHz
1 MHz
4 MHz
C [nF/cm^{-2}]
V [volts]
0
10
20
30
40
50
- 1.0
- 0.5
0
0.5
1.0

Figure 2


100
90
80
70
60
50
40
30
20
10
0
$(1/C_d{}^2) \times 10^{-13}$ [(F/cm^2)^2]
-0.12
-0.10
-0.08
-0.06
-0.04
-0.02
0
0.02
$\Psi_s(V) - \Psi_s(V_o) + kT/q$ [volts]

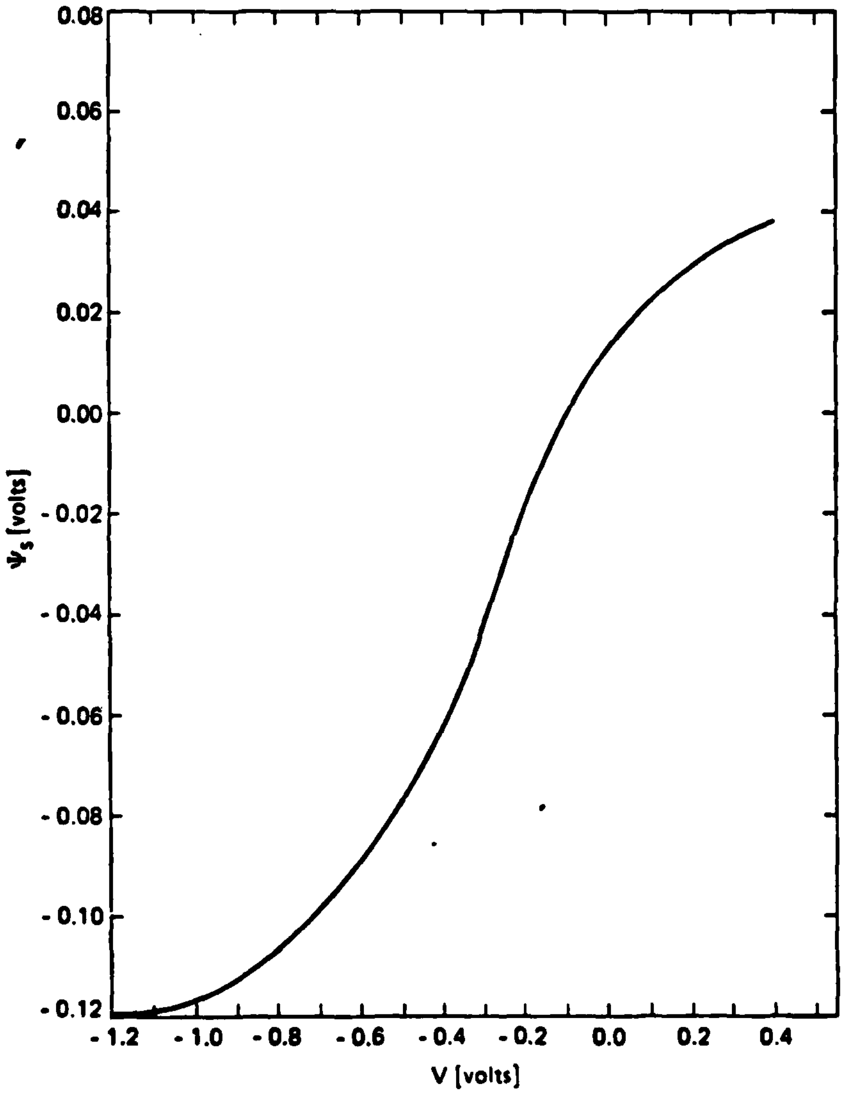
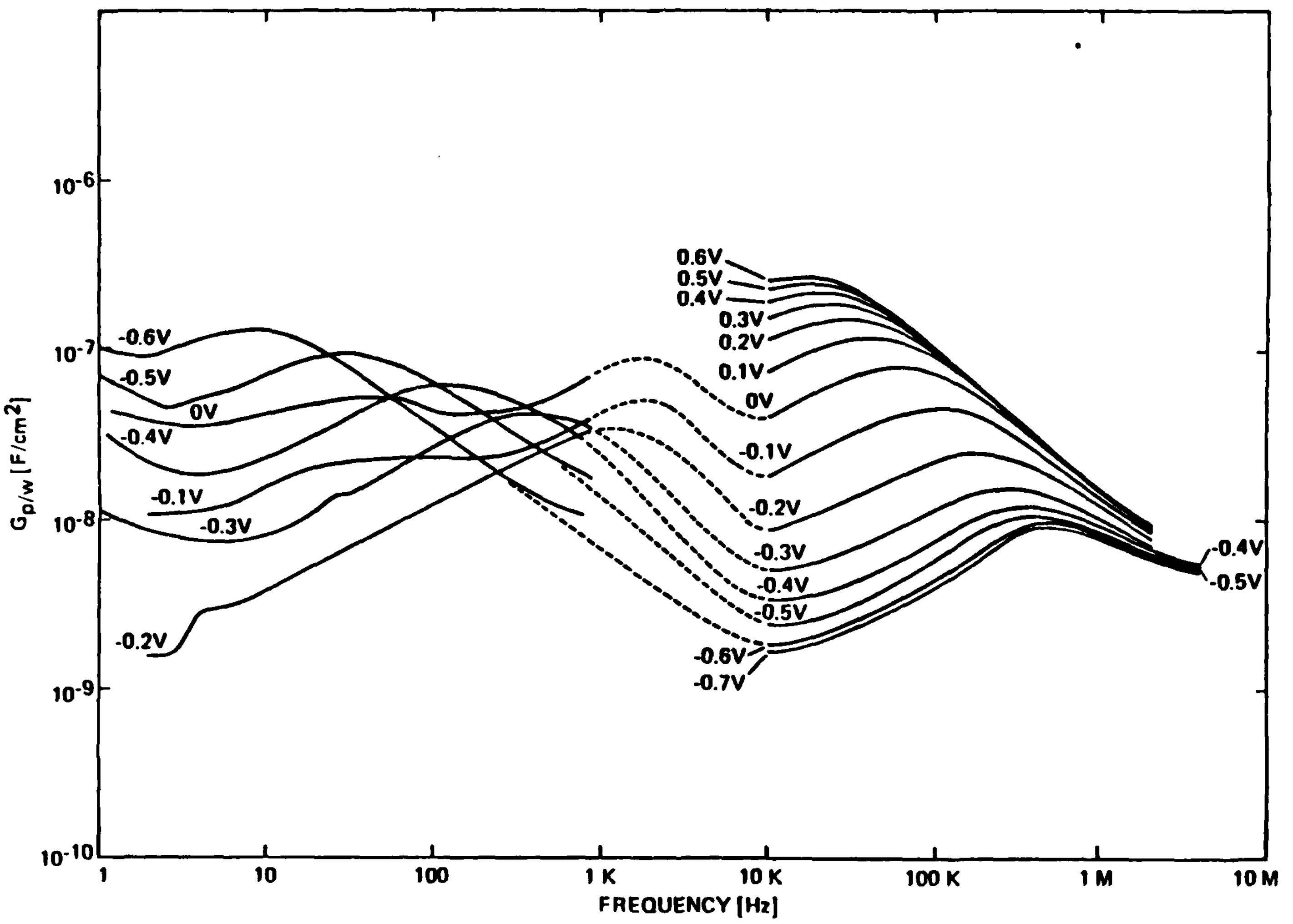
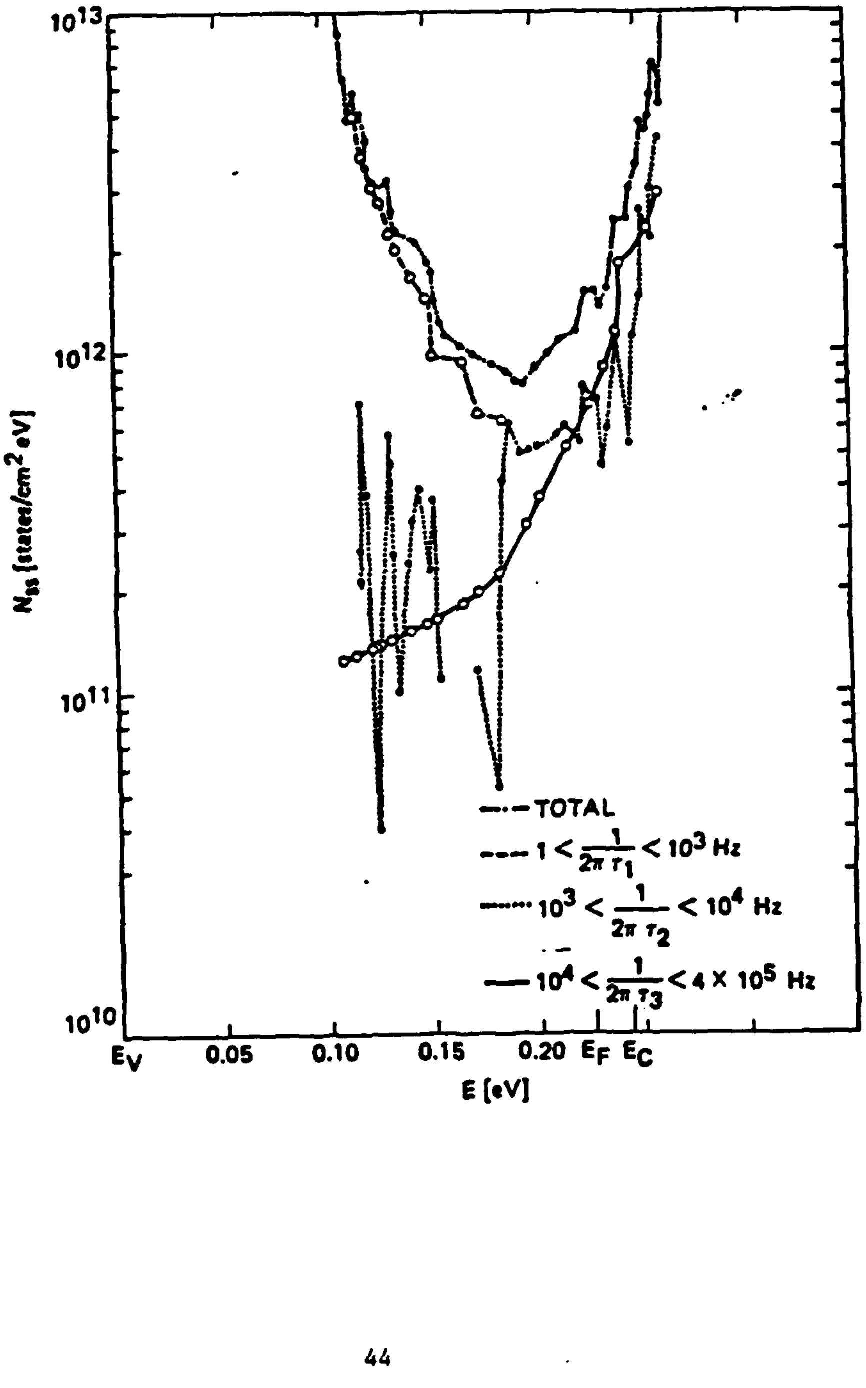

0.08
0.06
0.04
0.02
0.00
- 0.02
- 0.04
- 0.06
- 0.08
- 0.10
- 0.12
ψ_s [volts]
- 1.2
- 1.0
- 0.8
- 0.6
- 0.4
- 0.2
0.0
0.2
0.4
V [volts]

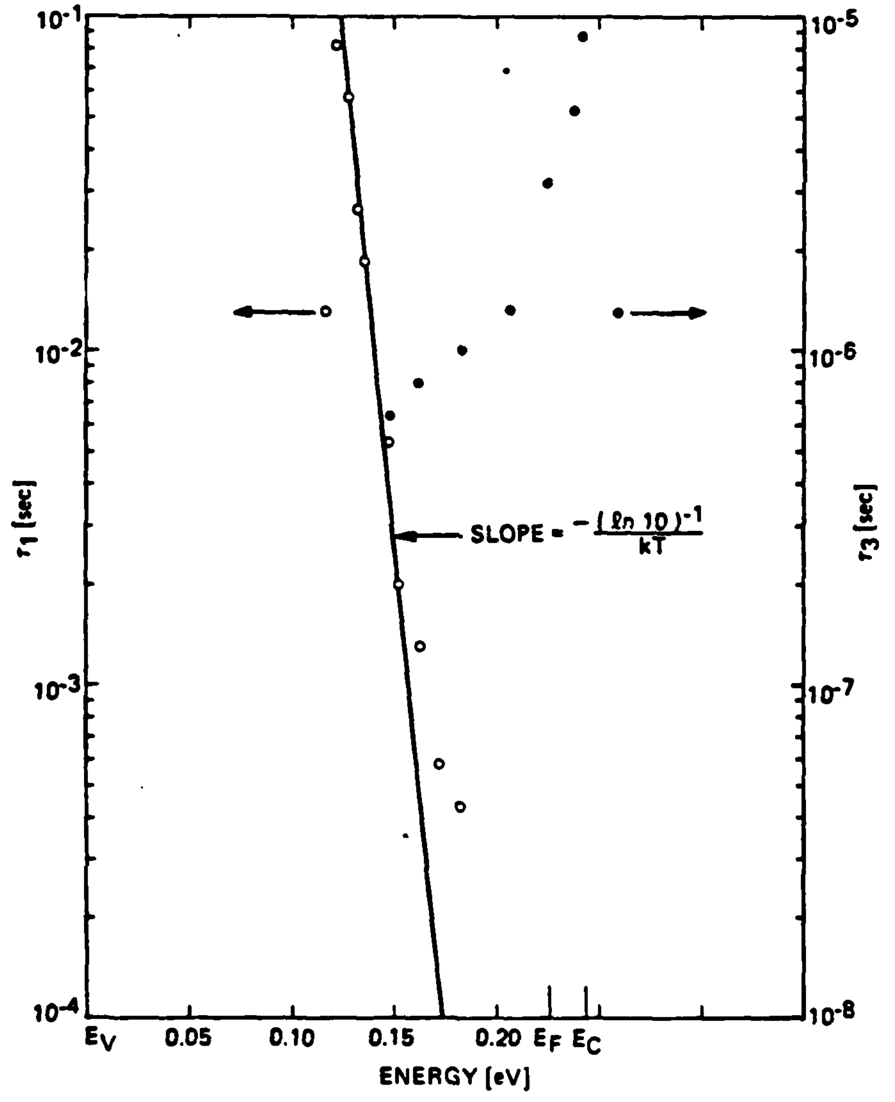
Figure 4



10-6
10-7
10-8
10-9
10-10
$G_{p/w}$ [F/cm^2]
1
10
100
1 K
10 K
100 K
1 M
10 M
FREQUENCY [Hz]
-0.6V
-0.5V
0V
-0.4V
-0.1V
-0.3V
-0.2V
0.6V
0.5V
0.4V
0.3V
0.2V
0.1V
0V
-0.1V
-0.2V
-0.3V
-0.4V
-0.5V
-0.6V
-0.7V
-0.4V
-0.5V

Figure 5


10^{13}
10^{12}
10^{11}
10^{10}
N_{ss} [states/cm^2 eV]
E_V
0.05
0.10
0.15
0.20
E_F
E_C
E [eV]
TOTAL
$1 < \frac{1}{2\pi \tau_1} < 10^3$ Hz
$10^3 < \frac{1}{2\pi \tau_2} < 10^4$ Hz
$10^4 < \frac{1}{2\pi \tau_3} < 4 \times 10^5$ Hz



10^{-1}
10^{-2}
10^{-3}
10^{-4}
τ_1 [sec]
10^{-5}
10^{-6}
10^{-7}
10^{-8}
τ_3 [sec]
SLOPE = $-\frac{(\ln 10)^{-1}}{kT}$
E_V
0.05
0.10
0.15
0.20
E_F
E_C
ENERGY [eV]

Figure 7

(THIS PAGE INTENTIONALLY LEFT BLANK)

Section II

CHANNELED IMPLANT STUDY

A. INTRODUCTION

While ion implantation has become the doping method of choice in the Si industry and is also used successfully in III-V semiconductor materials, it has not found a similar role in HgCdTe processing. This is of course due to the well-known effect that due to the ease of damaging the HgCdTe lattice, implantation of any ionic species always results in n-type activity due to these defects. In order to use implantation as a doping process, the ion must be expressed preferentially over the lattice damage. In order to do this, the damage must either be annealed out or not produced to begin with. Various annealing techniques have been extensively explored in ongoing efforts and are reported elsewhere by many groups. To date, however, no technique has been found to be completely satisfactory since all annealing processes examined produce more defects, principally Hg loss, than are removed. No conclusive demonstration of doping due to expression of the implanted species has yet been made.

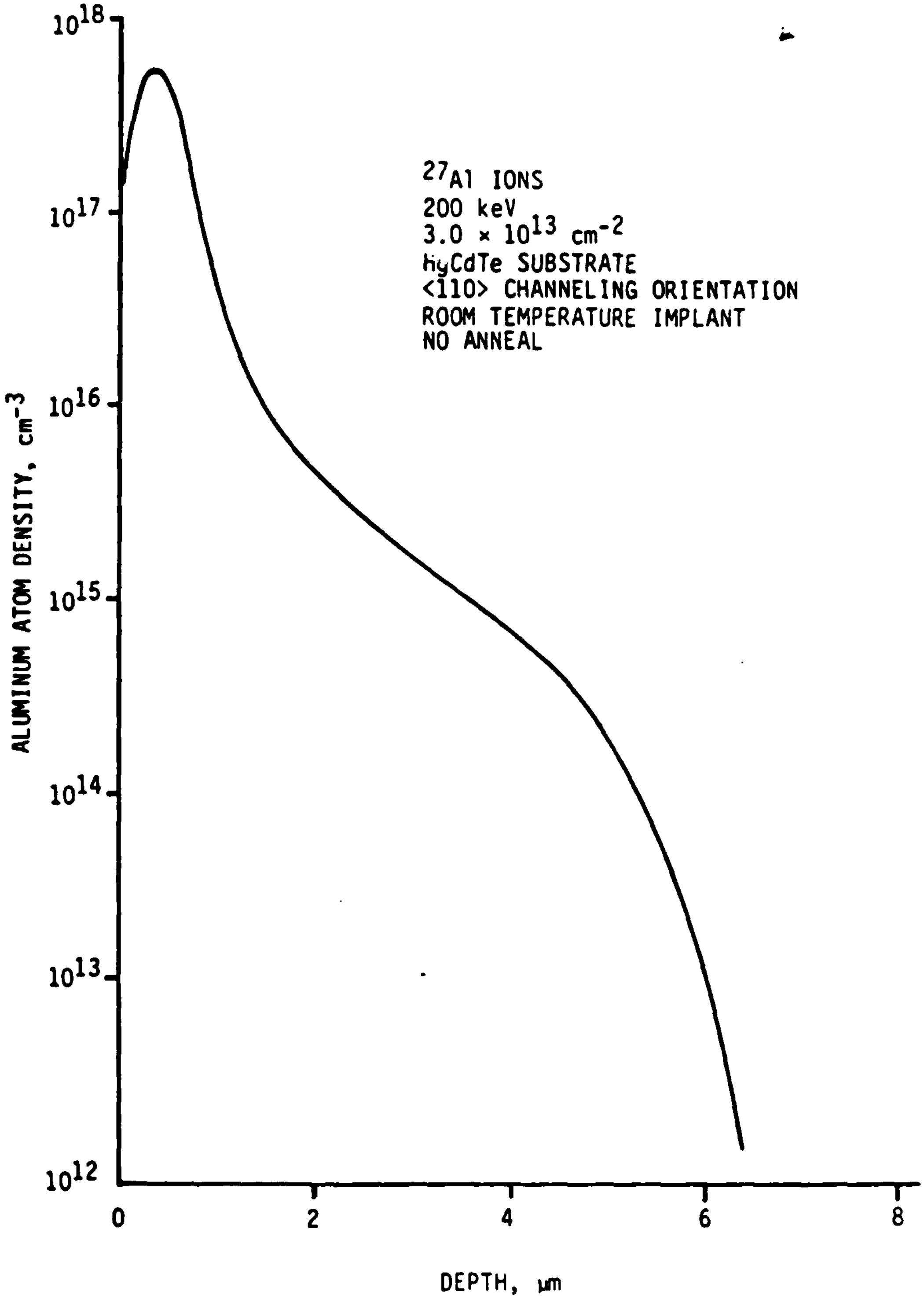
The alternate approach of channeling the implant into the (110) face of HgCdTe's zinc-blende lattice is the topic of this preliminary investigation. A number of channeled implants, as well as the more usual random implants, have been performed with a range of ions which are predicted to have n- or p-type activity on the cation or anion sites. A few implants have also been performed that are expected to be electrically inactive to isolated lattice damage effects. Ion profiles were obtained by SIMS analysis on all the implanted wafers, and because they contain useful and previously unpublished range statistics for various ions in HgCdTe, selected plots are presented in this section.

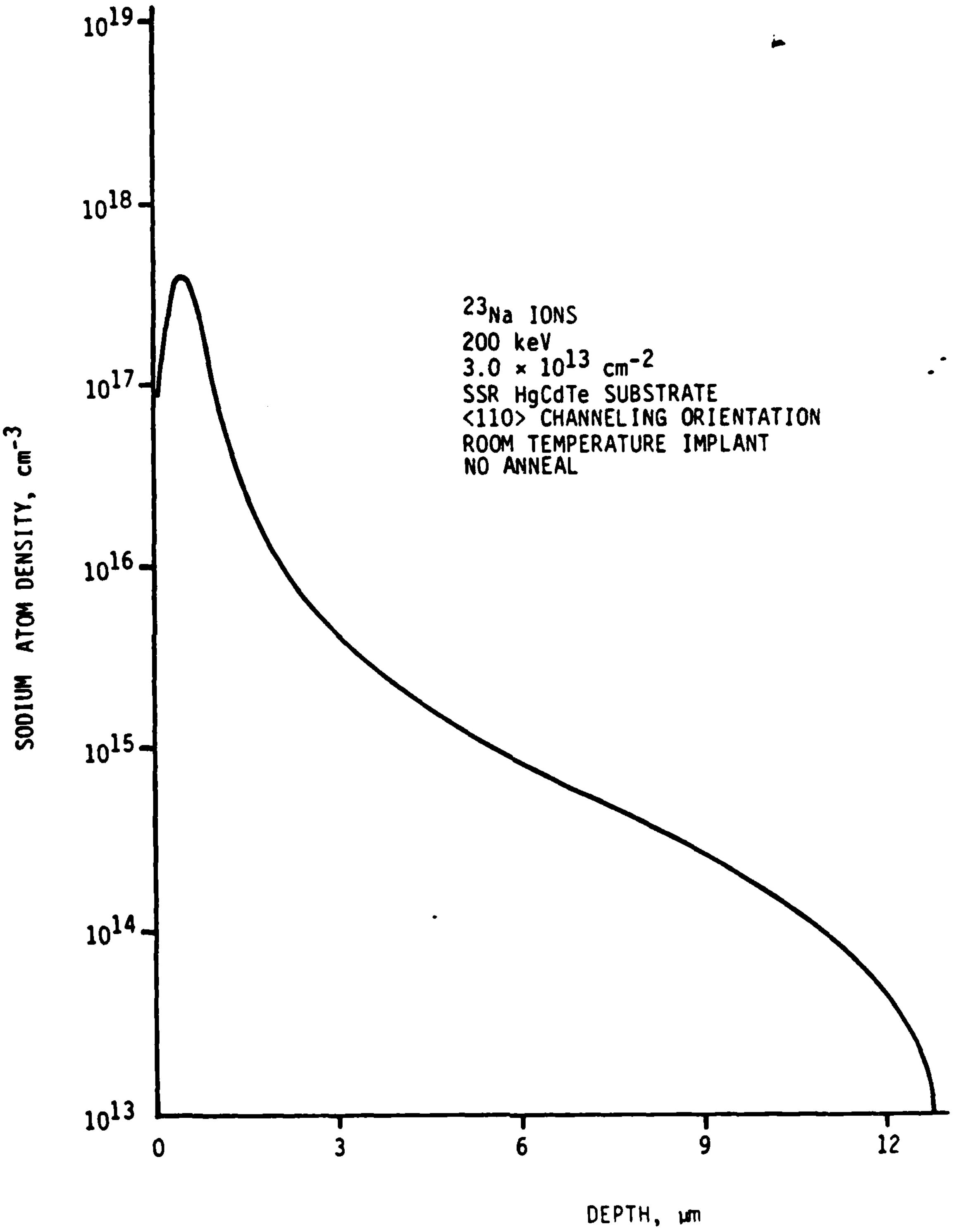
The wafers are made into MIS devices and examined for doping profiles and deep electron trap activity profiles using C-V analysis and deep level transient spectroscopy (DLTS). A trap due to the lattice damage has been identified near midgap. A report of this trap's characteristics is presented in the preprint contained in this section.

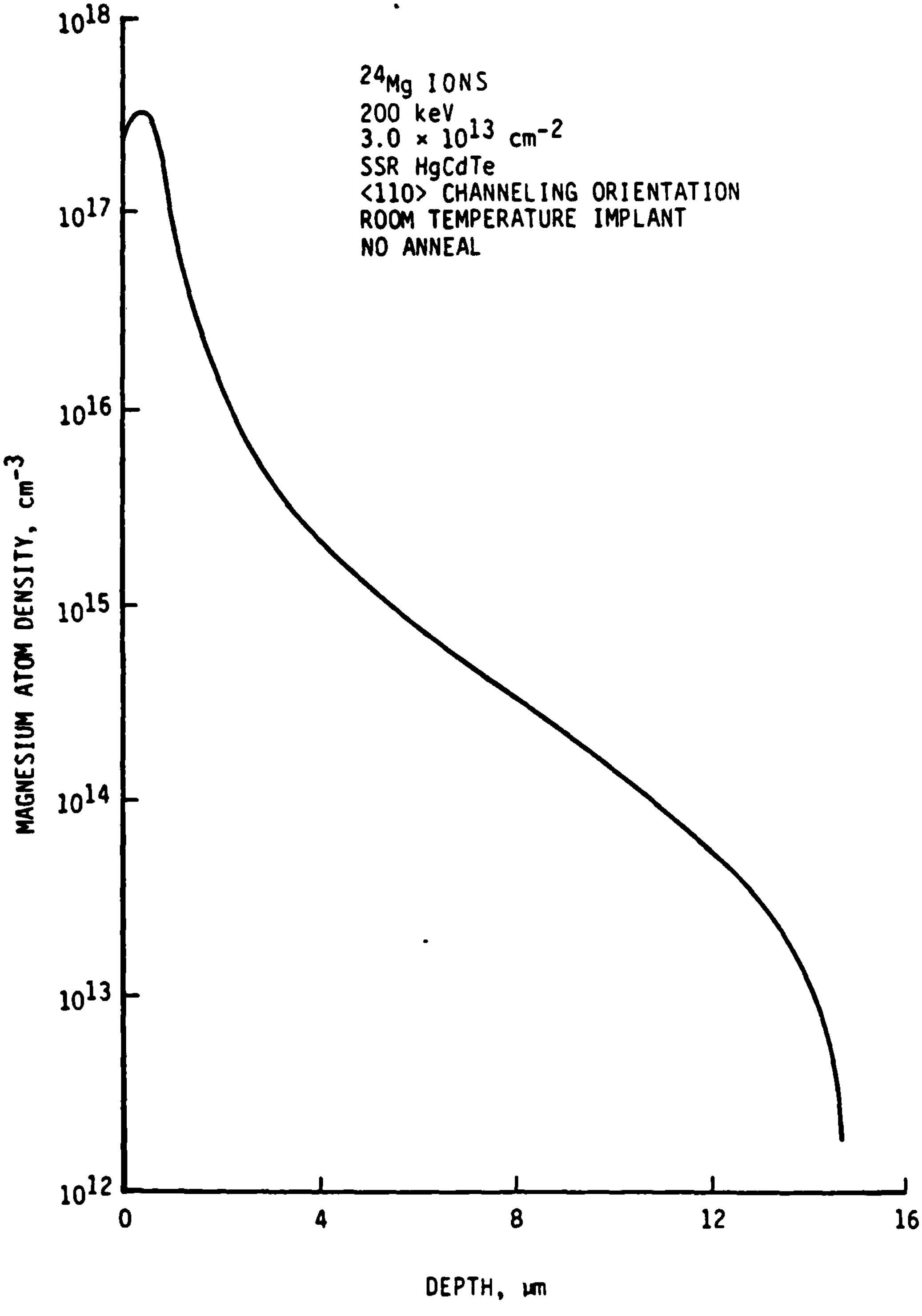
None of the samples prepared in this study can be claimed to demonstrate expression of the implanted ion. As seen in the SIMS profiles, the implants which were produced as channeled implants do demonstrate that a component of the incident ions actually channels; however, the overwhelmingly dominant proportion of the incident ions shows a final distribution that is characteristic of a random implant. In some cases no channeling component is seen at all even though the identical procedure produced some channeling in other samples. Considering the damage effects of surface processing discussed in the data presented in the other sections of this report, particularly the surface Cd depletion found after processes identical to those performed on these wafers prior to implant, it is most likely that the Te lattice near the surface has collapsed and blocked the channels normally associated with the (110) orientation of an ideal zinc-blende structure. New samples are being prepared that have been treated with a surface process believed to be much less prone to induce surface damage. These new samples are being implanted, and results will be reported in the final technical report.

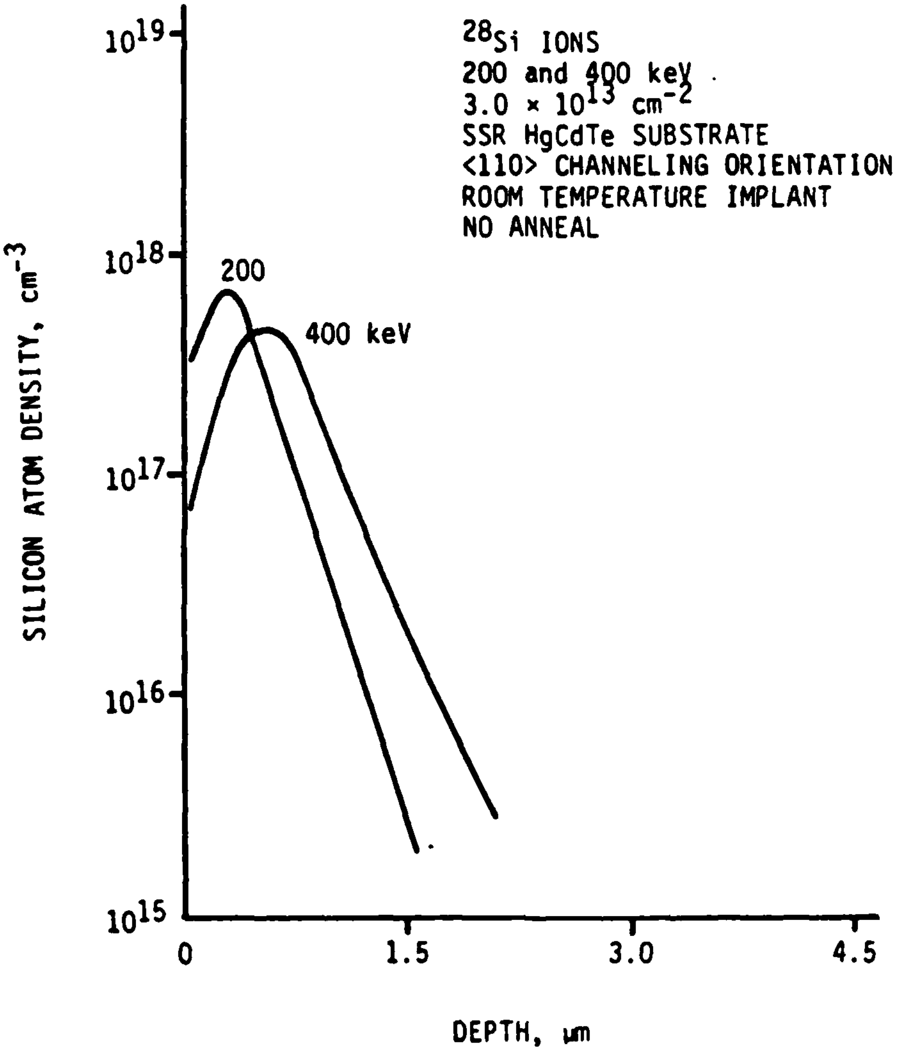
The following SIMS profiles of a selection of ions implanted into HgCdTe is presented because of the useful range information which they contain. This type of data for HgCdTe has not been previously available in the literature. The channeled implants were made into the (110) plane of wafers grown by solid-state recrystallization, while the random implants were made into horizontal zone melt wafers. None of the samples received any form of post-implant annealing, though all the wafers used were polished, Hg annealed, and etched prior to implantation. The ion implantations are done at Hughes Research Laboratory and the SIMS at Charles Evans, both under the direction of R.G. Wilson.

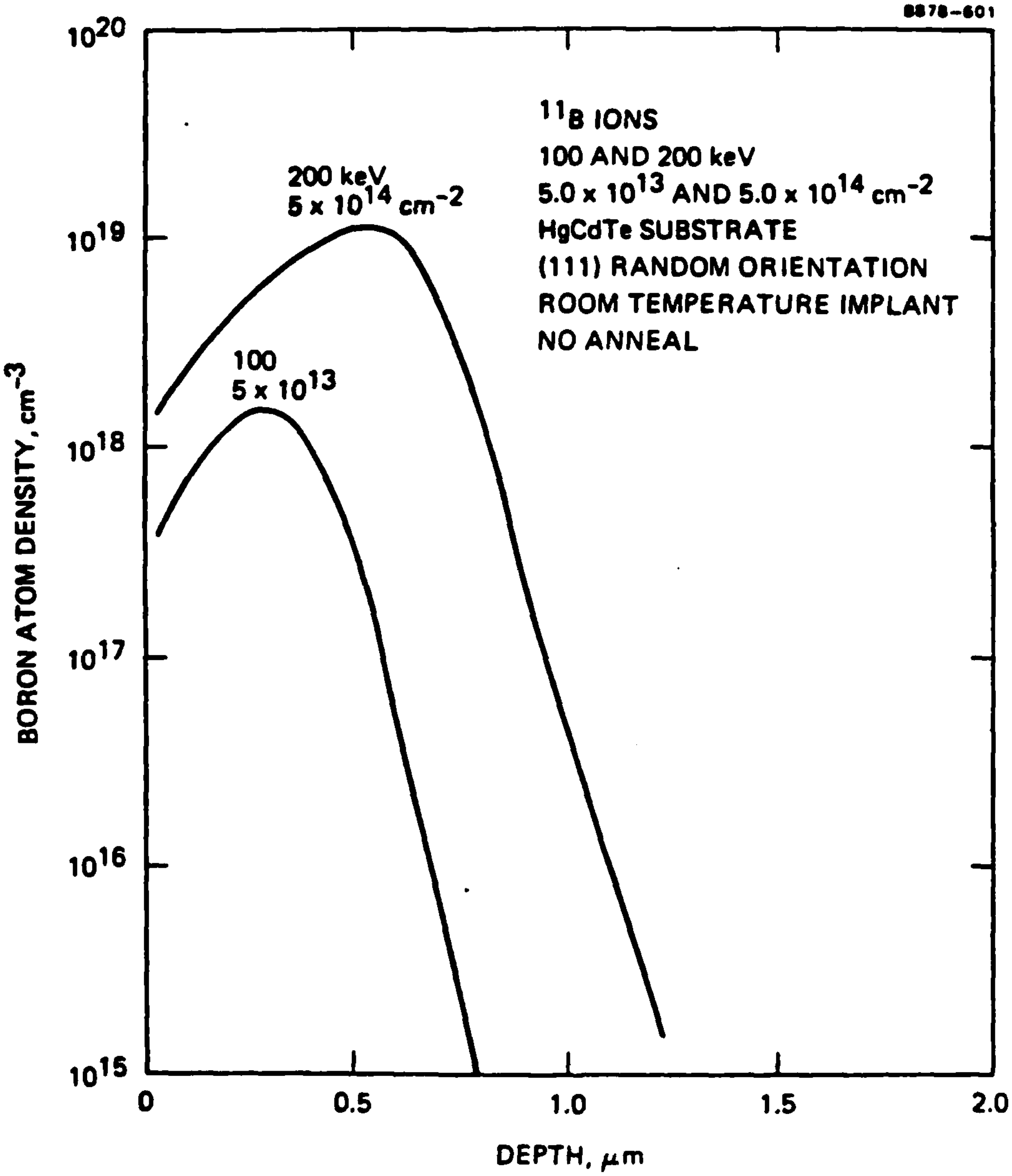
Ion	Energy (keV)	Dose (cm^{-2})	Type
Al	200	3×10^{13}	Channeled
Na	200	3×10^{13}	Channeled
Mg	200	3×10^{13}	Channeled
Si	200 and 400	3×10^{13}	Channeled
P	200	3×10^{13}	Channeled
As	200	3×10^{13}	Channeled
H	50, 100 and 200	3×10^{13}	Channeled
Br	200	3×10^{13}	Channeled
Cu	200	3×10^{13}	Channeled
B	100 and 200	5×10^{13} and 5×10^{14}	Random
Cu	200	5×10^{13}	Random

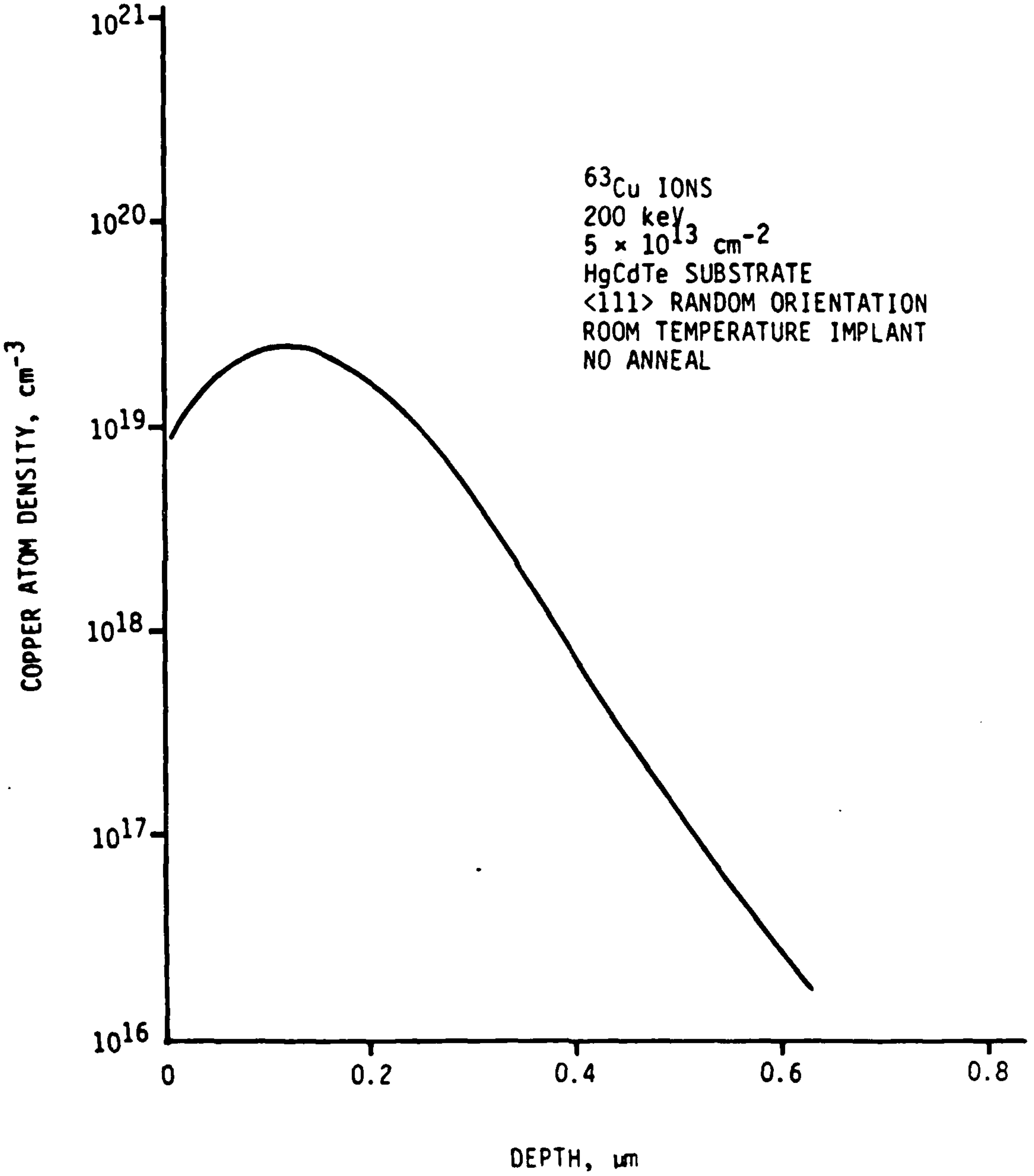

^{27}Al IONS
200 keV
3.0×10^{13} cm^{-2}
HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
ALUMINUM ATOM DENSITY, cm^{-3}
10^{18}
10^{17}
10^{16}
10^{15}
10^{14}
10^{13}
10^{12}
0
2
4
6
8
DEPTH, µm

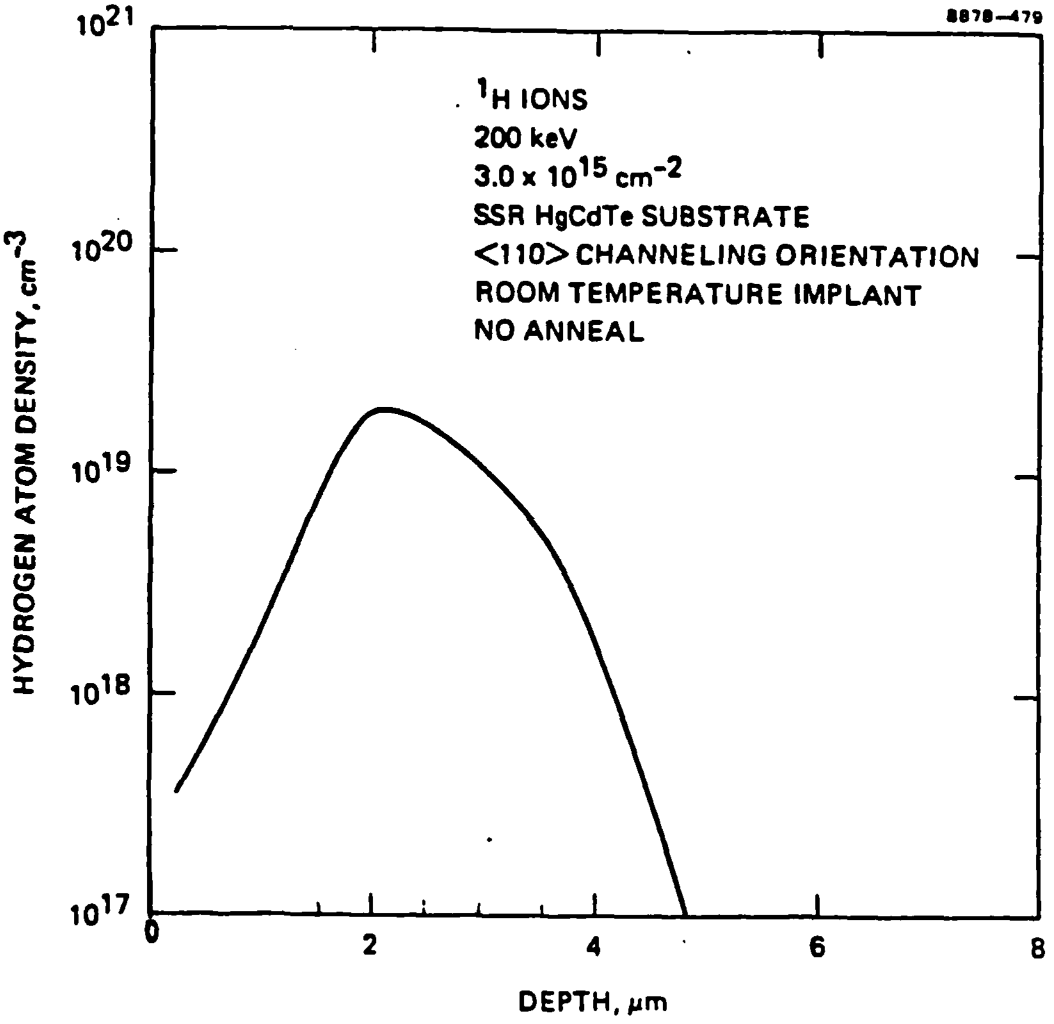

^{23}Na IONS
200 keV
3.0 × 10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
SODIUM ATOM DENSITY, cm^{-3}
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
10^{14}
10^{13}
0
3
6
9
12
DEPTH, µm

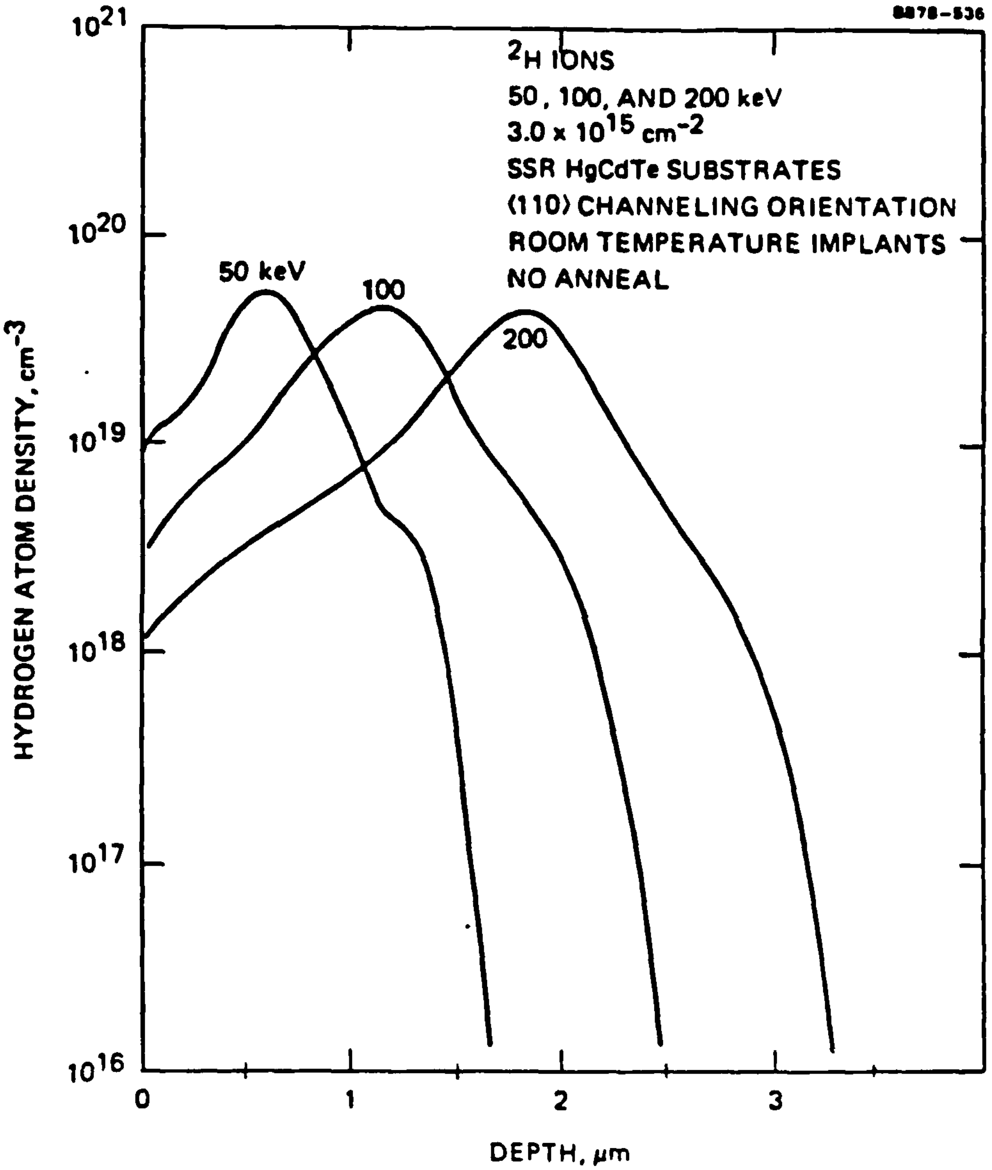

^{24}Mg IONS
200 keV
3.0 × 10^{13} cm^{-2}
SSR HgCdTe
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
MAGNESIUM ATOM DENSITY, cm^{-3}
10^{18}
10^{17}
10^{16}
10^{15}
10^{14}
10^{13}
10^{12}
0
4
8
12
16
DEPTH, µm

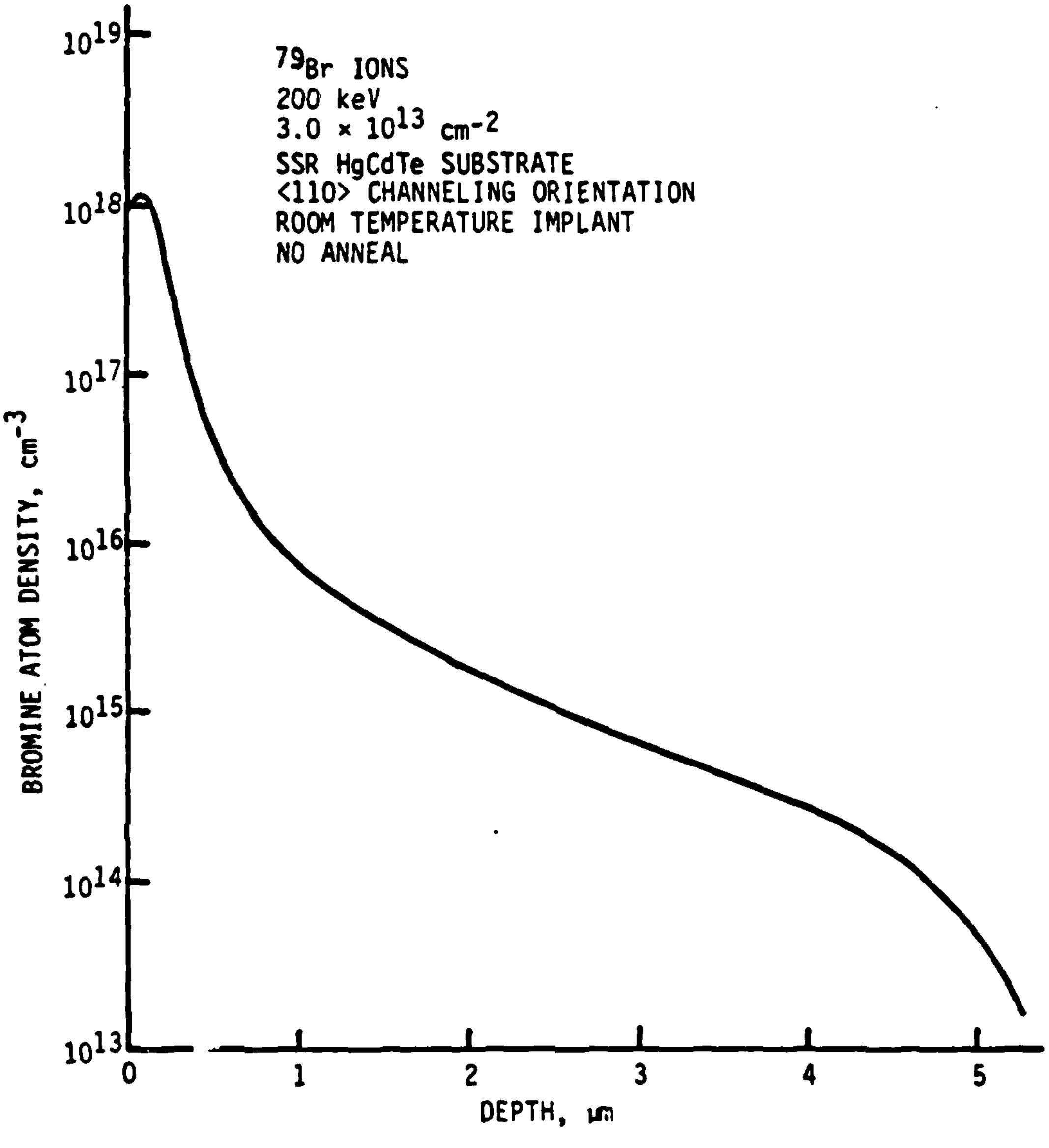

^{28}Si IONS
200 and 400 keV
3.0 × 10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
200
400 keV
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
SILICON ATOM DENSITY, cm^{-3}
0
1.5
3.0
4.5
DEPTH, μm


8878-601
^{11}B IONS
100 AND 200 keV
5.0 x 10^{13} AND 5.0 x 10^{14} cm^{-2}
HgCdTe SUBSTRATE
(111) RANDOM ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
200 keV
5 x 10^{14} cm^{-2}
100
5 x 10^{13}
BORON ATOM DENSITY, cm^{-3}
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
0
0.5
1.0
1.5
2.0
DEPTH, μm

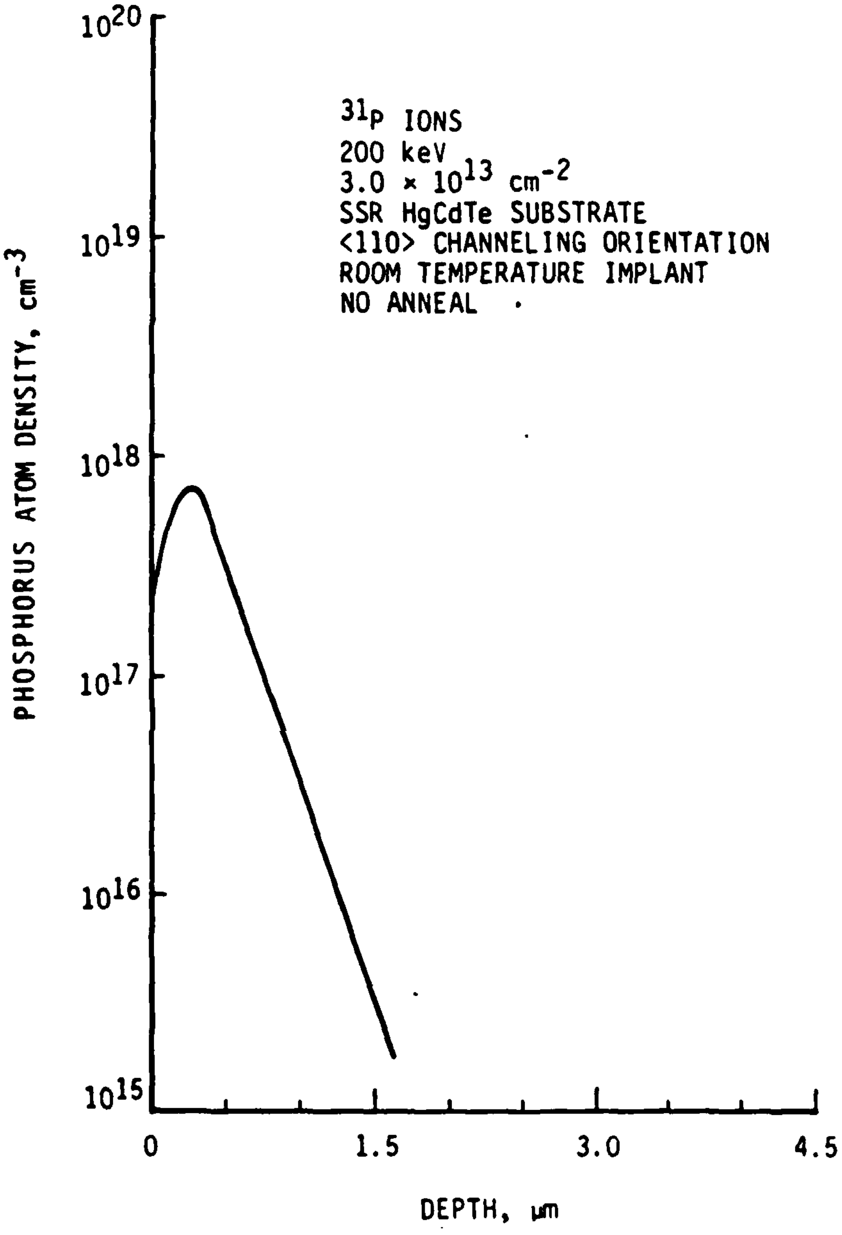

^{63}Cu IONS
200 keV
5 × 10^{13} cm^{-2}
HgCdTe SUBSTRATE
<111> RANDOM ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
COPPER ATOM DENSITY, cm^{-3}
10^{21}
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
0
0.2
0.4
0.6
0.8
DEPTH, μm

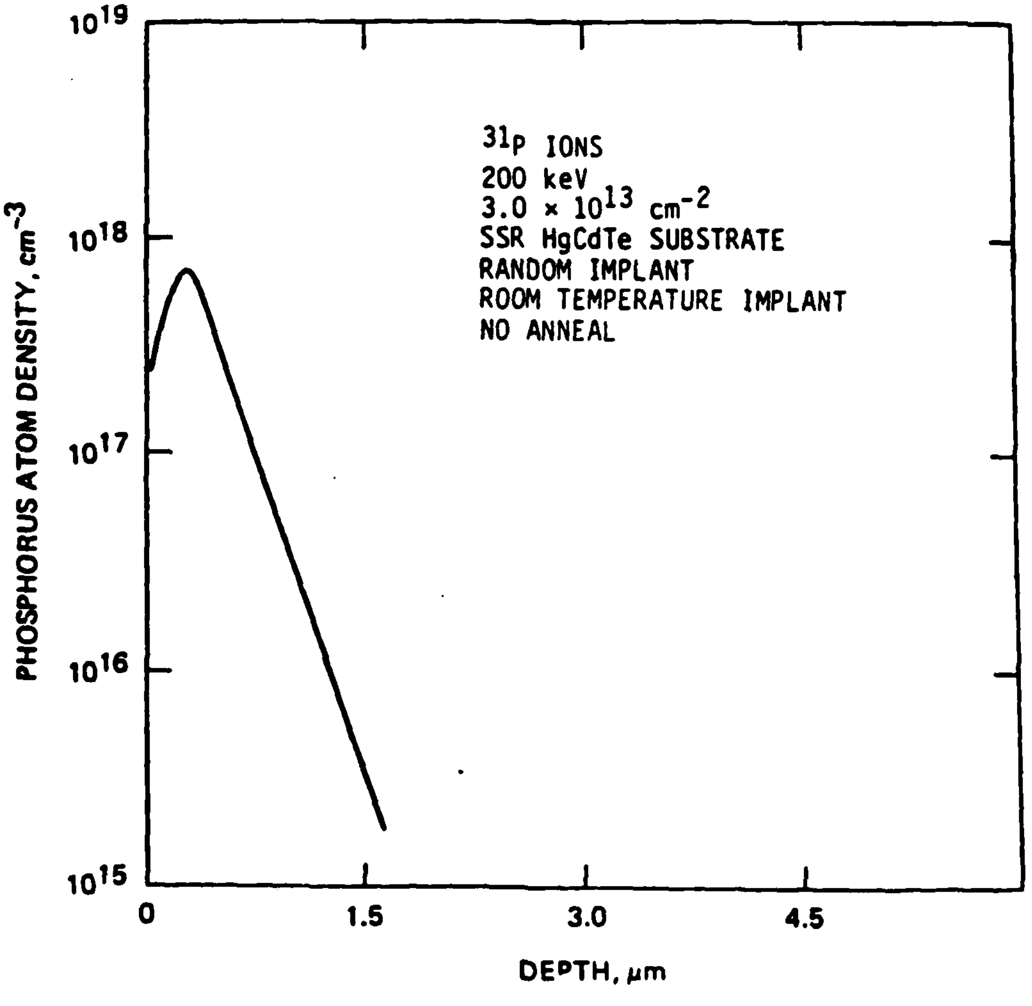

8878—479
^{1}H IONS
200 keV
3.0×10^{15} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
HYDROGEN ATOM DENSITY, cm^{-3}
10^{21}
10^{20}
10^{19}
10^{18}
10^{17}
0
2
4
6
8
DEPTH, μm

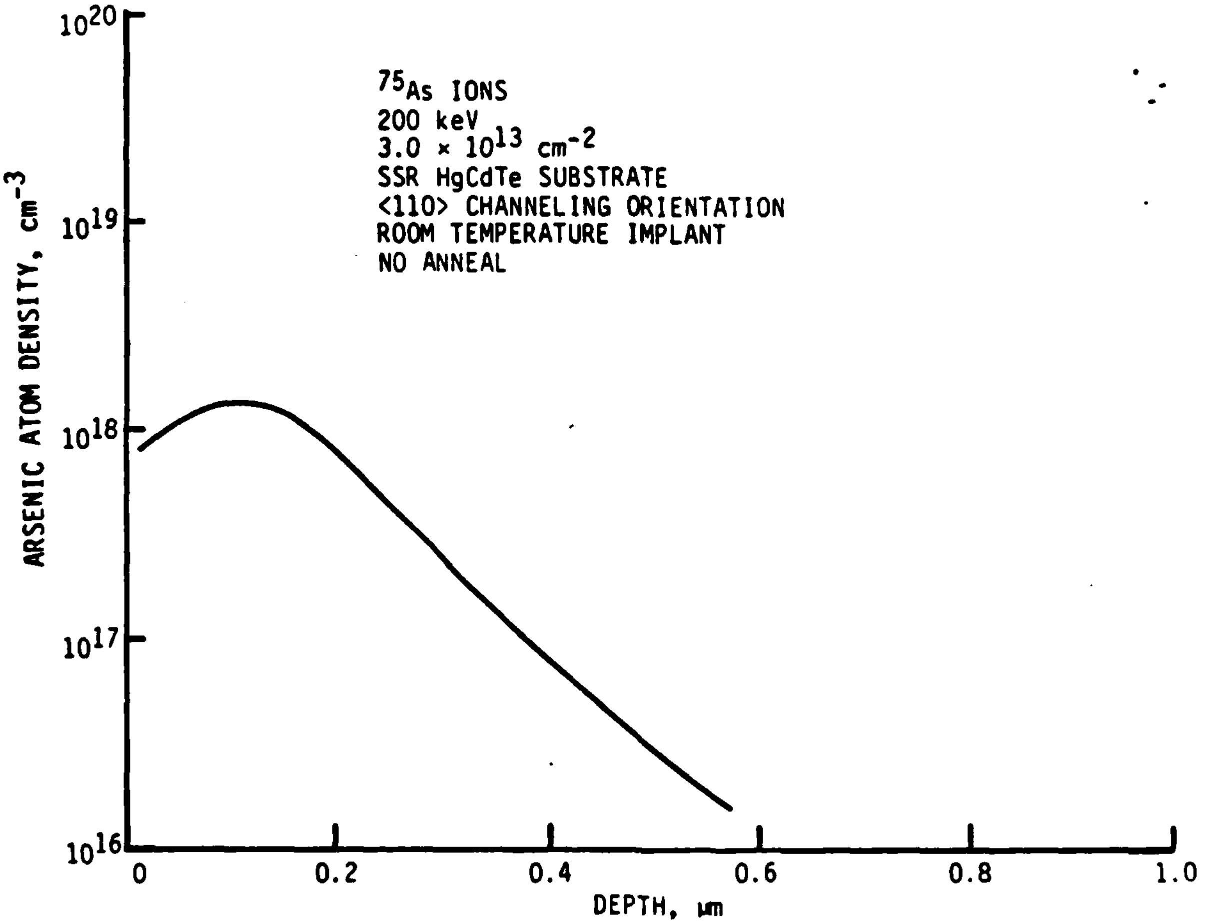

8878-536
^{2}H IONS
50, 100, AND 200 keV
3.0×10^{15} cm^{-2}
SSR HgCdTe SUBSTRATES
(110) CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANTS
NO ANNEAL
50 keV
100
200
HYDROGEN ATOM DENSITY, cm^{-3}
10^{21}
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
0
1
2
3
DEPTH, μm

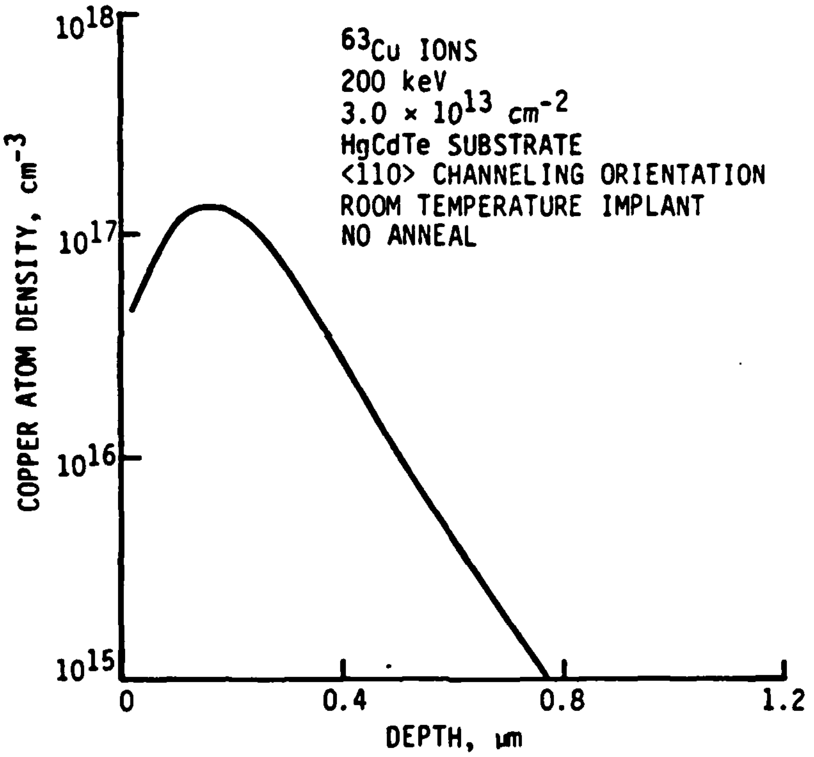


^{79}Br IONS
200 keV
3.0 × 10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL


10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
PHOSPHORUS ATOM DENSITY, cm^{-3}
^{31}P IONS
200 keV
3.0×10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
0
1.5
3.0
4.5
DEPTH, μm


^{31}P IONS
200 keV
3.0 × 10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
RANDOM IMPLANT
ROOM TEMPERATURE IMPLANT
NO ANNEAL
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
PHOSPHORUS ATOM DENSITY, cm^{-3}
0
1.5
3.0
4.5
DEPTH, μm


^{75}As IONS
200 keV
3.0×10^{13} cm^{-2}
SSR HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
ARSENIC ATOM DENSITY, cm^{-3}
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
0
0.2
0.4
0.6
0.8
1.0
DEPTH, µm


^{63}Cu IONS
200 keV
3.0 × 10^{13} cm^{-2}
HgCdTe SUBSTRATE
<110> CHANNELING ORIENTATION
ROOM TEMPERATURE IMPLANT
NO ANNEAL
COPPER ATOM DENSITY, cm^{-3}
10^{18}
10^{17}
10^{16}
10^{15}
0
0.4
0.8
1.2
DEPTH, µm

1. EFFECTS OF ION IMPLANTATION ON DEEP ELECTRON TRAPS IN $Hg_{0.7}Cd_{0.3}Te$

V.A. Cotton and J.A. Wilson
Santa Barbara Research Center
Goleta, CA 93117

ABSTRACT

Ion implantation is commonly used to create n^+ on p junctions for photovoltaic detectors in HgCdTe. This process is known to create a high level of donors due to lattice damage.[1] This work reports results of a deep level transient spectroscopy (DLTS) study of traps in n-type $Hg_{0.7}Cd_{0.3}Te$ also associated with the implant induced lattice damage.

A new trapping center is found at $E_c - 0.19$ eV which is distinctly due to implant damage. The electron trapping characteristics of this level are unusual, showing a small capture cross section of $\sigma_n = 5 \times 10^{-21}$ cm^2 and a large barrier potential to electron capture. This indicates a weak electron trap and implies a strong hole trapping nature, effective at limiting minority carrier lifetime in n-type HgCdTe.

INTRODUCTION

Ion implantation is a widely used process in many semiconductors for creation of n^+ on p junctions, ohmic contacts and barrier regions. In many materials (Si, InSb, and HgCdTe) this process creates n-type electrical activity due to lattice damage. HgCdTe presents unique problems in that a successful anneal to remove this damage, without severe Hg loss, has not been developed. The result is that many device properties are determined by the electrical activity created by implantation damage. In addition to high donor levels, there is the possibility of increased deep trapping centers in the bulk as well as increased interface traps. We investigate here deep electron traps associated with ion implant damage.

SAMPLE PREPARATION

Samples of $x = 0.3$ $Hg_{(1-x)}Cd_xTe$ were grown by both the horizontal zone melt and solid state recrystallization methods. Some samples were oriented <110>. This is the channeling direction for this crystal structure (zincblende) and hence should result in an implant dominated by electronic stopping since the nuclear stopping component is strongly reduced when channeling into an ordered lattice. Annealing at 240°C in Hg vapor for four weeks resulted in n-type material with $|N_D - N_A| \cong 3 \times 10^{15}\ cm^{-3}$. Chlorine, boron and hydrogen were implanted into both oriented and unoriented wafers with an accelerating voltage of 200 keV. The samples received no post-implant anneal. Each wafer was then masked and step-etched in sections to various depths along the implant direction.

Metal-Insulator-Semiconductor (MIS) devices were formed on each step and used as the measurement vehicle. These were fabricated by deposition of 1500 Å of PHOTOX" SiO_2 as the insulator followed by thermal evaporation of 400/4000 Å of Ti/Au through a shadow mask to form gates. Indium soldered directly to the wafer provided the ground.

MEASUREMENT TECHNIQUE

Deep Level Transient Spectroscopy (DLTS) was used to identify and characterize trapping centers. DLTS takes advantage of the temperature dependence of the capacitance transient resulting from traps emptying within the depletion region after a rapid pulse into deep depletion. Use of MIS devices on n-type material results in characterization of electron traps only, since they are the majority carriers, and yields capture cross section, activation energy and trap density.

Capacitance-voltage (C-V) measurements were used to measure average doping density within the depletion region, carrier type, and mobile oxide charge. These measurements are also used to select the appropriate bias voltages for DLTS scans. Mobile oxide charge must be minimal to ensure accurate knowledge of the depletion region width at a given bias. Secondary Ion Mass Spectroscopy provided the implant species depth profile.

EXPERIMENTAL RESULTS

SIMS profiles of unoriented wafers (unchanneled implant) show a profile closely approximated by LSS theory.[2] The channeled implants show a peak similar to that for unchanneled samples followed by a "tail" of channeled ions which penetrate more deeply into the material. A comparison of these two types of implants is shown in Figure 1 for B. Channeled samples will always show some unchanneled component due to the fact that some incoming ions do not "see" channels but scatter from the ends of the atomic planes at the lattice surface. This may be enhanced by loss of crystallinity near the surface due to sawing, polishing and other processing.

Donor densities obtained from C-V on unchanneled samples show a concentration higher than the implant species, by a factor of ~100 for Cl and ~25 for B, and following the same shape, as shown in Figure 2. This is consistent with findings that implanted junctions are located deeper than expected based on ion concentration profiles. The damage-induced donor level becomes low enough to just compensate the substrate acceptor concentration at two to five times the depth of the implant species. Hydrogen is the exception. The donor concentration in this case is below the implanted H concentration (Figure 3). This implies that the induced damage decreases with decreasing mass of the implanted species. Donor density profiles of channeled samples are very similar to those for unchanneled indicating that the nuclear stopping component is still dominant. A clear evaluation of damage due to electronic stopping is thus not yet possible.

Quasi-static C-V measurements, attempted in order to get interface trap densities (D_{it}), were not successful on implanted samples due to high drift of the measured current. This indicates a high density of interface traps and very high surface leakage.

Deep bulk traps, believed to be those commonly present in unimplanted material, were detected by DLTS in these samples. These have activation energies of E_c - 0.11 eV and E_c - 0.14 eV, which are very similar to those previously reported.[4,5] However, more accurate capture cross section measurements, limited in this case by bias drift, are needed to completely verify this. The concentration of these traps are independent of implanted ion concentration or damage-induced donor density and are typically an order of magnitude lower than the background donor density.

The major finding of this work is a new trap in the 200K region of the DLTS spectra which is not present in unimplanted samples. Figure 4 shows a DLTS spectra for an unimplanted bulk sample compared with one which received an unchanneled B implant. What appears to be a decrease in concentration of bulk trap levels is actually an artifact of the higher doping density and hence narrower depletion region in the implanted case. The concentration per unit volume is the same in both cases for levels seen in the 100K region. Figure 5 shows a comparison of channeled and unchanneled implanted boron. As the doping density decreases in the channeled sample, bulk traps become more visible. Also, a slight shift in position of the trap in the 200K region is seen.

This new trap has a DLTS activation energy of 0.1 eV to 2.3 eV depending on implant species and whether the implant was channeled or not. This measured activation energy is composed of two elements; the actual trap activation energy, E_t, referenced to the conduction band edge, and the capture cross section activation energy, E_σ, which is a measure of local conduction band barrier to capture such that:

$$E_{DLTS} = E_t + E_\sigma$$

After separation of these effects the new trap is found to have an activation energy, E_t, of E_c - 0.19 eV ± 0.02 eV and an usually large barrier potential which is measured at 0.06 eV and 0.25 eV for channeled and unchanneled boron, respectively, and 0.04 eV and 1.8 eV in unchanneled hydrogen and chlorine. There is a slight shift in activation energy also but this is less than 0.02 eV between the extremes of hydrogen and chlorine.

Captured cross section was measured to be 5×10^{-21} cm^2, again with a slight shift to 2×10^{-21} cm^2 for channeled hydrogen.

As shown for unchanneled Cl in Figure 6, the trap concentration is higher than the implant ion profile, in this case, by a factor of five. Even in the case of H, the trap concentration is comparable to the implanted H concentration. For both B and Cl the concentration is higher than the implant species but lower than the donor concentration.

CONCLUSION

A new trap has been found which is distinctly correlated with implant damage. Since the concentration of this trap is greater than the implanted ion concentration, direct host atom substitution or clustering is not likely as this would yield a one-to-one correlation. This indicates damage due to propagated energy loss as the source of the trap. The activation energy, E_t, and capture cross section, σ_n, are both relatively constant with respect to both mass and channeled versus unchanneled implant conditions. However, the barrier potential, E_σ, changes dramatically. It appears, then, that implanted ion mass, as well as both electronic stopping and nuclear stopping components play a role in determining the trap characteristics.

Existing deep trap concentrations are, to first order, unaffected by the implantation process. However, the slight shift in activation energies may be reflective of local potential changes due to resulting damage in the lattice.

Both the small capture cross section and the large barrier potential measured for the trap which results from implantation indicate that this is a weak electron trap. The concentration of this trap will be fairly low on the p-side of the junction where it would act as a minority carrier lifetime limiter. This trap is expected to be a very effective hole trap due the difference of ≈ 200 in effective mass. Thus it may be a lifetime limiter on the n^+ side of the junction, however this is the side with the narrow depletion region and will not dominate. Lifetime measurements on variable area diodes formed by ion implantation indicate that the bulk lifetime of the material is shortened by this process.[6] Surface leakage may be enhanced, but is not dominant. So far, traps seen by DLTS are not sufficient to account for this degradation. A closer look at alteration of existing deep traps may shed some light on this question.

Acknowledgments

We gratefully acknowledge many helpful discussions with C. E. Jones. We also acknowledge the technical assistance of R. G. Wilson in performing the implants, obtaining SIMS data as well as many fruitful discussions, and M. E. Boyd for assistance in recording capacitance data. This work was supported under DARPA contract number MDA-903-83-C-0108. The program sponsor is R.A. Reynolds and the technical monitor is J. Murphy both of DARPA/DSO.

REFERENCES

[1] R.K. Willardson and A.C. Beer, Semiconductors and Semimetals, Vol. 18, (Academic Press, N.Y., 1981).

[2] J. Lindhard and M. Scharff, Phys. Rev., 124, 1. 128 (1961).

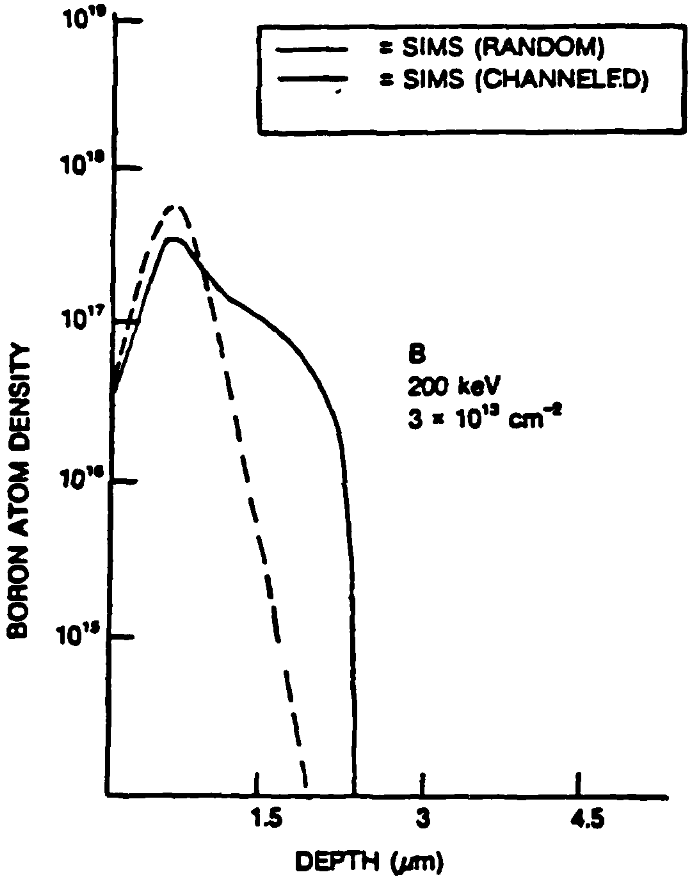
[3] J. Baars, A. Hurrie, W. Rothamund, C.R. Fritzche and T. Jakobus, J. Appl. Phys. 53(3), March 1982.

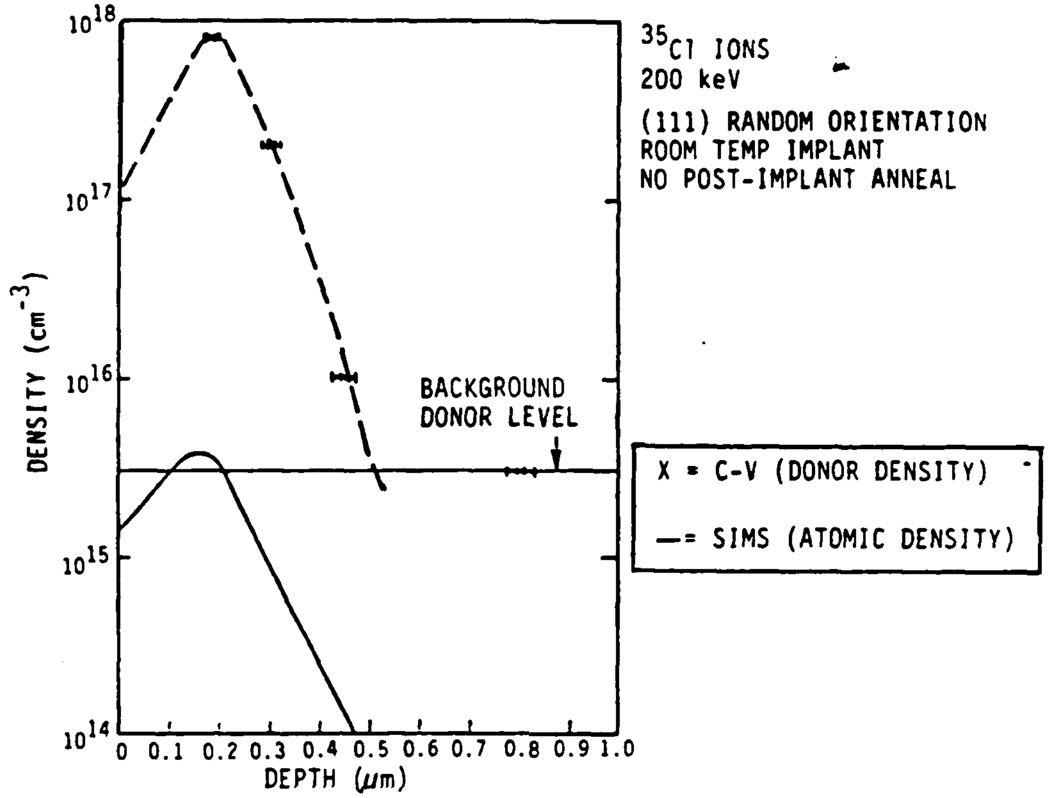
[4] C.E. Jones, V. Nair and D.L. Polla, Appl. Phys. Lett. 39, 248 (1981).

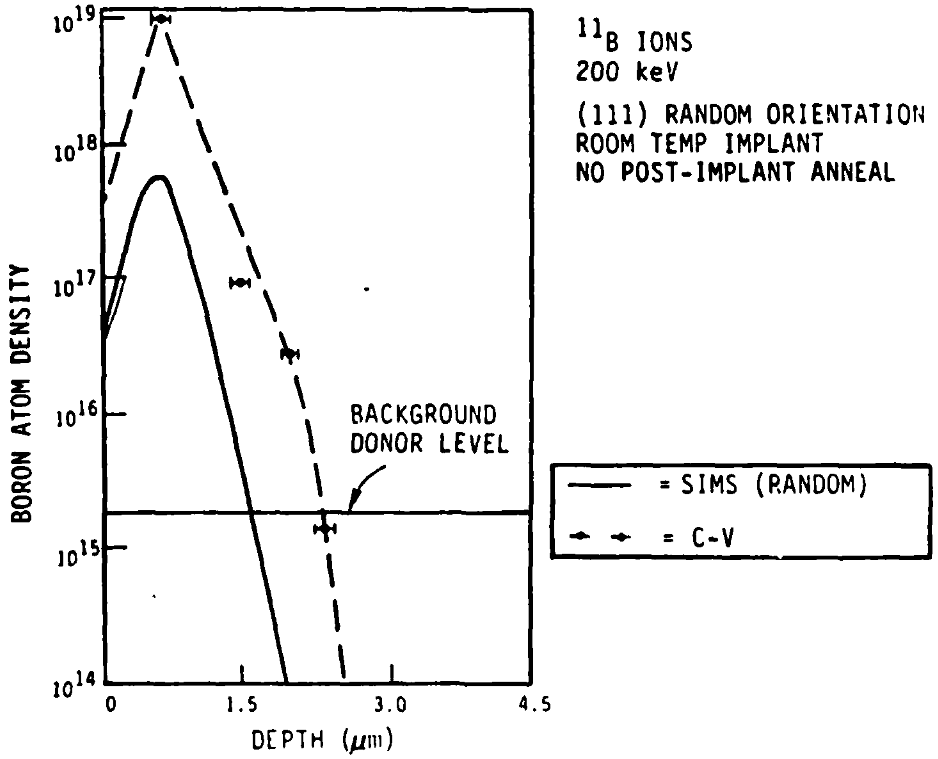
[5] V.A. Cotton, J.A. Wilson and C.E. Jones, J. Appl. Phys. 58(6), 15 Sept. 1985.

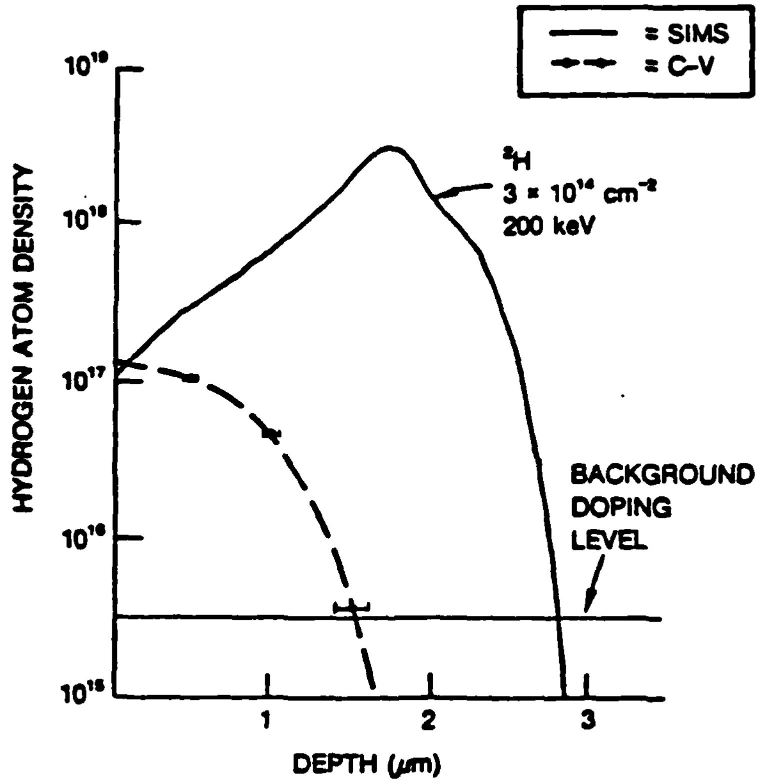
[6] W.A. Radford and C.E. Jones, Second Interim Report, 'Exploratory Development of Advanced HgCdTe Detectors', Contract No. F33615-82-C-5122, May 1985.

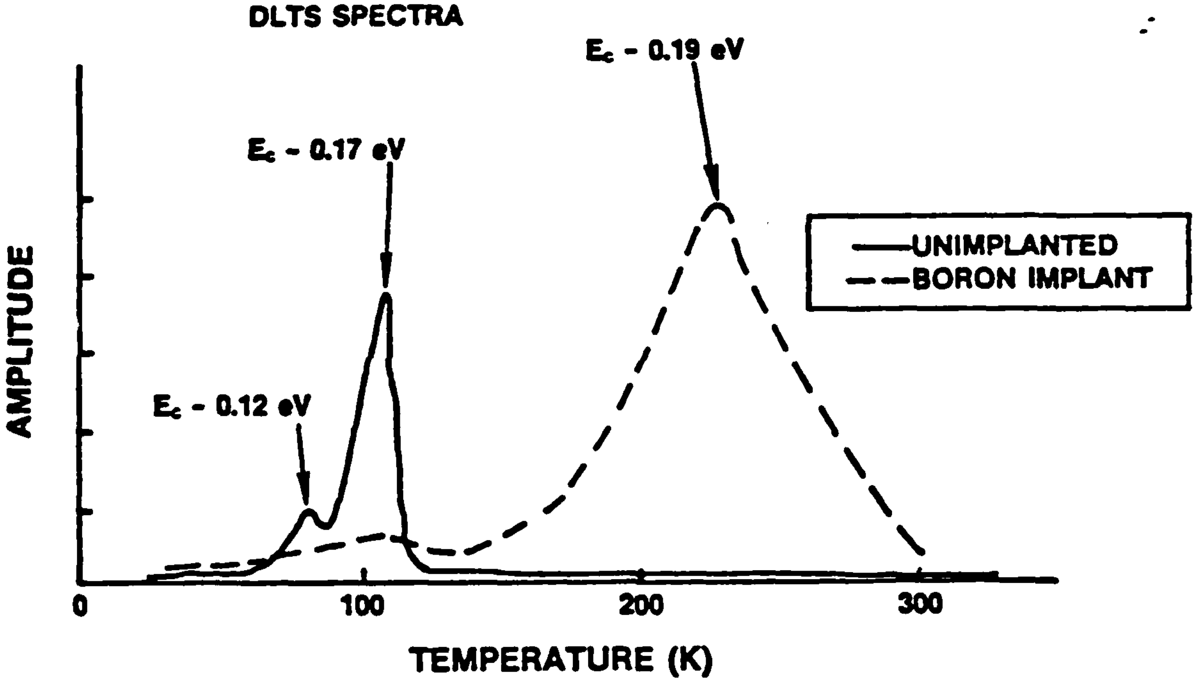
FIGURES

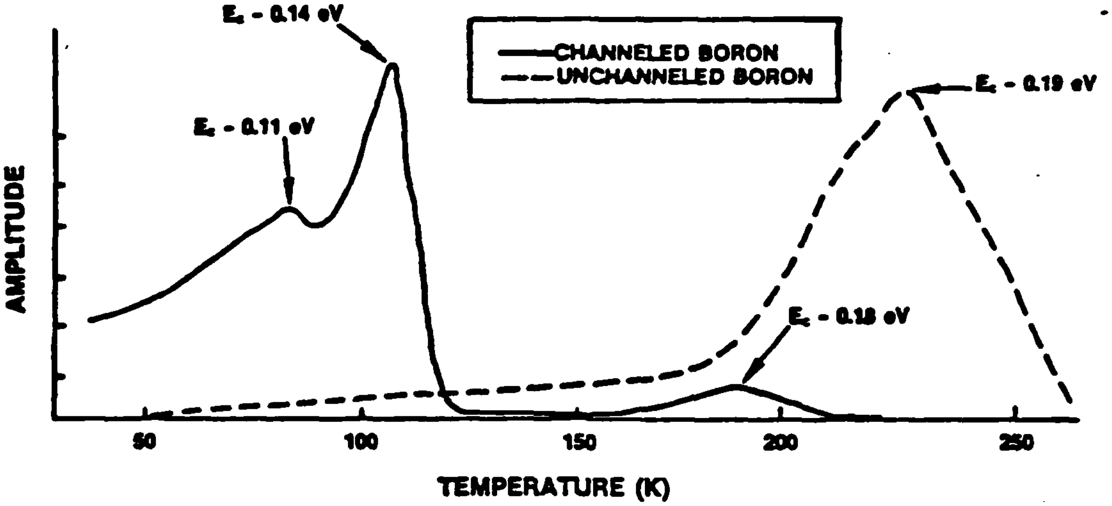

= SIMS (RANDOM)
= SIMS (CHANNELED)
10¹⁹
10¹⁸
10¹⁷
10¹⁶
10¹⁵
BORON ATOM DENSITY
B
200 keV
3 × 10¹³ cm⁻²
1.5
3
4.5
DEPTH (μm)

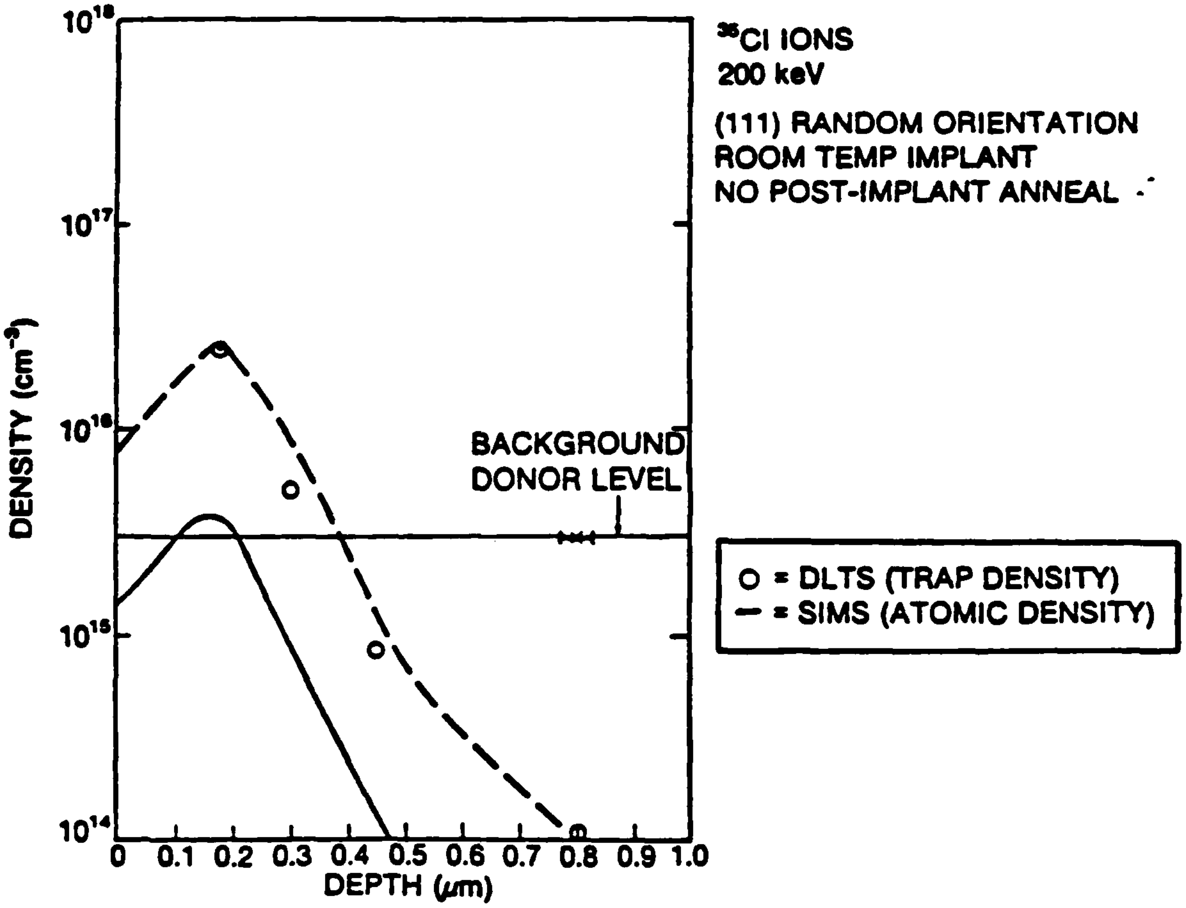

^{35}Cl IONS
200 keV
(111) RANDOM ORIENTATION
ROOM TEMP IMPLANT
NO POST-IMPLANT ANNEAL
BACKGROUND DONOR LEVEL
X = C-V (DONOR DENSITY)
— = SIMS (ATOMIC DENSITY)
DENSITY (cm^{-3})
10^{18}
10^{17}
10^{16}
10^{15}
10^{14}
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
DEPTH (μm)


^{11}B IONS
200 keV
(111) RANDOM ORIENTATION
ROOM TEMP IMPLANT
NO POST-IMPLANT ANNEAL
BACKGROUND DONOR LEVEL
= SIMS (RANDOM)
= C-V
BORON ATOM DENSITY
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
10^{14}
0
1.5
3.0
4.5
DEPTH (μm)


— = SIMS
- - = C-V
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
HYDROGEN ATOM DENSITY
^{2}H
3×10^{14} cm^{-2}
200 keV
BACKGROUND
DOPING
LEVEL
1
2
3
DEPTH (μm)


DLTS SPECTRA
AMPLITUDE
TEMPERATURE (K)
0
100
200
300
E_c - 0.12 eV
E_c - 0.17 eV
E_c - 0.19 eV
——UNIMPLANTED
– – –BORON IMPLANT


E_c - 0.14 eV
E_c - 0.11 eV
CHANNELED BORON
UNCHANNELED BORON
E_c - 0.19 eV
E_c - 0.18 eV
AMPLITUDE
50
100
150
200
250
TEMPERATURE (K)



^{35}Cl IONS
200 keV

(111) RANDOM ORIENTATION
ROOM TEMP IMPLANT
NO POST-IMPLANT ANNEAL

Section III
SURFACE STUDY

A. INTRODUCTION

In order that manufacturing of HgCdTe be optimized, it is necessary that fundamental understanding be obtained of critical parameters which may be affected by processing. Under this contract, attention is directed at surfaces and interfaces. In order to achieve a fundamental understanding of the effects of technologically important processes on HgCdTe surfaces and interfaces, it is essential to characterize the changes that occur in composition, atomic structure, and electronic structure in near-surface regions. The sensitivity of the HgCdTe surface to chemical processing procedures and the interaction of the surface with passivation layers and deposited metal films is of utmost concern to device stability and performance. Of particular concern is the effect of procedures that alter surface stoichiometry and their resulting influence on defect formation and band bending. Of extreme importance also is the effect of removal of cations (corresponding to a modified surface x-value) on the bandgap and other electronically sensitive parameters at the surface. These important concerns are the focus of investigations reported herein.

Characterization of the interaction between metal overlayers evaporated onto atomically clean HgCdTe interfaces is critical to an understanding of practical metal/HgCdTe interfaces, the electrical characteristics of metal contacts, and the time stability of such contacts. In the previous semiannual report, we described initial studies of the interface chemistry between HgCdTe and the metals Al and Ag, specifically chosen to represent extremes of thermodynamic reactivity with the alloy. As expected, the Al/HgCdTe interface exhibited significant chemical reaction, with a depletion of Hg at the surface. In Section 1 we present further studies of the very different character of the nonreactive Ag/HgCdTe interface, which indicates that Ag diffuses 10^2 to 10^3Å into the bulk, with a change in surface composition occurring due to Hg displacement.

The chemistry and electronic structure of the HgCdTe surface with intentionally altered stoichiometry has been studied by taking advantage of electron beam induced desorption of Hg from the lattice. We outlined the experiment in the past semiannual report and described early results concerning the controlled depletion of Hg from the surface. In Section a. we present new information which describes the loss of surface crystallinity and valence band dispersion that occurs with higher Hg loss. These results indicate that the significant alteration of the lattice and electronic structure of the alloy is taking place; however, at a much lower rate than might be expected (e.g., a surprising amount of Hg must be removed before its effects are seen).

Characterization of chemically prepared HgCdTe surfaces permits one to extend fundamental studies of atomically clean surfaces to surfaces of technological importance. In Section 1 we present the results of our continuing examination of the surface composition and electronic properties (position of the surface fermi level) of surfaces etched with bromine-methanol. The importance of compiling a detailed history of sample growth and processing procedures has been stressed in previous reports. We include here additional background information obtained for previously examined as well as new samples. Cd cation loss relative to the nominal bulk composition is found for all bromine-methanol treated surfaces (with known bulk x-value).

1. DIFFUSION OF Ag AND Hg AT THE Ag/(Hg,Cd)Te INTERFACE

D. J. Friedman, G. P. Carey, C. K. Shih, I. Lindau, and W. E. Spicer
Stanford Electronics Laboratories, Stanford Univ., Stanford, CA, 94305

J. A. Wilson
Santa Barbara Research Center, Goleta, CA 93017

Abstract

The room temperature Ag/(Hg,Cd)Te interface has been studied in ultra-high vacuum with X-ray and ultraviolet photoelectron spectroscopy. The Ag evaporated onto (Hg,Cd)Te diffuses 10^2 to 10^3 Å into the bulk of the semiconductor, displacing Hg as it does so. The bands bend 0.05-0.1 eV upward at low coverage. No significant chemical reaction is observed.

Given the fundamental and technological importance of the metal/(Hg,Cd)Te interface, very little work has been done on the subject compared, for example, to the research done on Si or GaAs. For the early studies of this interface, metals lying at the extremes of reactivity with (Hg,Cd)Te (hereinafter referred to as MCT) will be of special interest. Since Ag has an especially small heat of formation[1] with Te, $\Delta H_f(Ag_2Te)=-8.6$ kcal/mole, Ag should show little chemical reaction with MCT. Other relevant ΔH_f's are[1] $\Delta H_f(HgTe)=-7.6$ kcal/mole, as well as[2] $\Delta H_f(Ag_2Cd)=-1$ kcal/mole, $\Delta H_f(Ag_2Cd)=-0.7$ kcal/mole. This paper presents a photoemission study of the Ag/MCT interface formed at room temperature, which proves to be dominated by Ag diffusion into the semiconductor. This interface thus behaves quite differently from the nonreactive Au/MCT[3] and reactive metal/MCT interfaces[4-6] studied to date, as well as from e.g. the Ag/Si interface, for which the largest amount of significant intermixing observed involves only the first two monolayers or so.[7]

A 5×5×15 mm^3 bar of p-type $Hg_{.77}Cd_{.23}Te$ was transferred into a previously-baked vacuum chamber and then cleaved in ultra-high vacuum (base pressure 6×10^{-11} torr) to reveal a (110) face. Upon cleaving the surface converted to n-type, as is often observed.[8] Sequential depositions of Ag were evaporated from a tungsten filament onto the surface, which was studied at each stage by X-ray and ultraviolet photoelectron spectroscopy (XPS and UPS) using respectively a Mg Kα X-ray source ($h\nu$ = 1253.6 eV) and a helium lamp (He I and He II, $h\nu$ = 21.2 and 40.8 eV). The photoelectrons were analyzed with a double-pass cylindrical-mirror analyzer (CMA). The amount of Ag deposited is given in units of a monolayer, ML, which we define to be the surface density of atoms on the MCT (110) face: 1 ML $\equiv 6.76\times10^{14}$ atoms/cm^2 = 1.15 Å of metallic Ag.

The XPS data will be used for quantitative composition analysis as follows. The area I of a photoemission peak at kinetic energy E for a subshell of atom A with quantum numbers n,l is related to the concentration c(x) of atoms A at distance x from the surface by

$$I = T(E)\sigma_{nl}(E,h\nu)\, F \int_0^\infty e^{-x/\lambda(E)} c(x)\, dx. \tag{1}$$

Here $\lambda(E)$ is the electron escape length at kinetic energy E, T(E) is the CMA transmission, σ_{nl} is the subshell partial photoionization cross section (PPCS), $h\nu$ is the photon energy, and F includes all other factors. Thus for a fixed experimental geometry and photon flux the concentrations c_1 and c_2 of atoms 1 and 2 in a sample of uniform composition will be related to the XPS peak areas I_1 and I_2 by $c_2/c_1=(S_2/S_1)(I_2/I_1)$, where $S\equiv[\lambda(E)T(E)\sigma(E)]^{-1}$ is the atomic sensitivity. The escape lengths we use are interpolated from escape lengths for CdTe given by Ref. 9, the PPCS are taken from Ref. 10, and T(E)=const/E for our CMA at constant pass energy. We shall denote the concentration of atom A calculated from the peak area of an n,l subshell excited by radiation $h\nu$ as $c(A\,nl;h\nu)$, or c(A) when the meaning is clear.

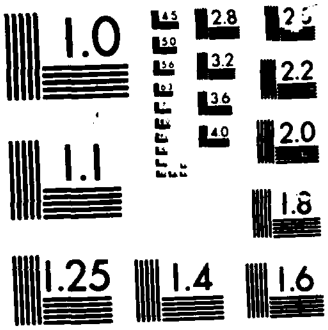
Fig. 1 shows as a function of Ag deposition the XPS peak areas of the semiconductor and Ag core levels, multiplied by the atomic sensitivity factors and normalized to the initial concentration of Te. While the Hg 4f peak area attenuates with increasing Ag coverage, neither the Te 3d nor the Cd 3d peak area attenuates significantly. These trends are reflected in the UPS spectra, which show the Hg 5d peak area going to zero at the highest coverages while the Cd 4d peak is virtually unchanged. This lack of attenuation, even at greater than 100 ML Ag deposition, suggests that either the Ag has a very low sticking coefficient c_S, defined as the fraction of atoms evaporated onto the substrate which do not reevaporate, or it diffuses fairly far into the bulk of the semiconductor (or a combination of the two effects). We have also measured in the same experimental configuration the XPS Ag 3d peak area I(Ag3d) from 6.4 ML Ag evaporated onto a stainless substrate and find it to be approximately equal to I(Ag3d) from 150 ML Ag deposited on the semiconductor, and thus also ten times greater than I(Ag3d) for 6.4 ML Ag/MCT. This is consistent with either in-diffusion of Ag or low c_S for Ag on MCT. Both types of behavior would be suggestive of very weak chemical reaction between the Ag and the semiconductor. No significant core level shifts could be resolved in the XPS data, confirming that minimal chemical reaction takes place. This inertness, also seen in the Au/MCT interface[3], is in sharp contrast to the behavior of the reactive Al/MCT interface, for which a reacted-phase Al 2p peak can be seen along with a metallic Al 2p peak at about 1 eV lower binding energy.[4,6] Strong chemical reaction is also seen at the Cr/MCT and In/MCT interfaces.[4,5] For these reactive metals, deposition is also accompanied by a marked Hg depletion at from the surface, as well as Te movement outward through the overlayer metal.[4-6]

If low sticking coefficient but not diffusion played a role for the Ag/MCT interface, the ten times greater I(Ag3d) for Ag/stainless compared to an equal amount of Ag/MCT would show that Ag sticks to itself at least ten times as effectively as to MCT. From eq. 1 we calculate that I(Ag3d) from the Ag/stainless corresponds to about 2.5 ML Ag actually stuck to the stainless. Since I(Ag/stainless)≈1.2×I(30ML Ag/MCT), at least 2 ML Ag must have stuck to the MCT at 30 ML deposition giving a lower bound of 1/15 for c_S. If the deposited Ag stayed on the surface, then after 1 ML stuck, Ag deposited subsequently would see an Ag-covered surface and would start to stick an order of magnitude more effectively. Thus I(Ag3d) should start to rise faster with Ag deposition, as indicated qualitatively by the dashed line in Fig. 1. Instead, the opposite is observed: I(Ag3d) starts to level off. Therefore, while low sticking coefficient c_S may play some role, the Ag must be diffusing at the very least on the order of 5-10 lattice constants (25-50 Å) into the bulk at 30 ML deposition. Since c_S should not decrease between 30 and 150 ML deposition, the amount of Ag actually in the near-surface should increase by at least a factor of five from 30 to 150 ML deposition. However, the Ag peak area rises by only about 50%. Therefore the additional Ag must diffuse to a depth of at least about 100-200 Å.

Since the Ag diffuses so far, the composition of the near-surface will be approximately uniform within an XPS escape depth. Therefore the intensity-sensitivity products in Fig. 1 represent actual atomic concentrations relative to the bulk Te concentration. Note that (with no adjusted parameters) the calculated composition for the clean surface agrees closely with the nominal composition of $Hg_{.77}Cd_{.23}Te$, lending support to the procedure. From the figure we see that at 150 ML deposition, c(Ag)/c(Te) has started to level off at about 1.0 while c(Hg)/c(Ag) has fallen to about 0.1, suggesting that Ag is replacing Hg in the lattice. The fact that c(Ag)/c(Te) actually rises somewhat above the original Hg/Te ratio of 0.77 at the highest coverage indicates that Ag is starting to pile up at the surface. For this reason, some emission up to the Fermi level E_f starts to be visible in the He I UPS spectra at the highest coverage. Even at this coverage, the diffusion makes the emission near E_f much weaker than for the 6.4 ML Ag/stainless reference.

AD-A166 794 HGCDTE SURFACE AND DEFECT STUDY PROGRAM(U) SANTA BARBARA RESEARCH CENTER GOLETA CALIF J A WILSON ET AL. JUL 85 SBRC-60070 MDA903-83-C-0108 2/3

UNCLASSIFIED F/G 20/12 NL



MICROCOP CHART

Because the escape lengths $\lambda(\text{Ag4d:HeII}) \approx {}^1/_3\lambda(\text{Ag3d:MgK}\alpha)$, the ratio of the Ag concentrations observed at these different surface sensitivities will indicate the gradient of the Ag concentration. This ratio may be estimated with a minor extension of the method described above for the XPS data. Except for the lowest coverages, the concentration ratio c(Hg5d:HeII)/c(Hg4f:MgKα) is close to one, implying a minimal gradient in the Ag concentration within an XPS escape length of order 25 Å and thus further supporting our contention that the Ag diffuses an order of magnitude farther than this. At the highest coverages c(Hg5d:HeII)/c(Hg4f:MgKα) starts to rise, due to the start of Ag pile-up at the surface.

The data of Fig. 1 are replotted in Fig. 2 to show the percent composition of the Ag-diffused semiconductor near-surface as a function of Ag deposition. The figure graphically illustrates the substitution of Ag for Hg in the lattice. It is of interest to calculate how far into the bulk a given amount of Ag can replace Hg to yield the Ag concentration seen in the near-surface. From Fig. 2, after 30 ML has been deposited about half the Hg is replaced by Ag at the XPS probing depth. Since 40% of the atoms in the MCT (110) plane are Hg atoms, 30 ML Ag of which a fraction c_s sticks can replace half the Hg atoms in the first $150 \times c_s$ layers, or about $700 \times c_s$ Å into the bulk (since the interlayer spacing is 4.57 Å). For the lower bound on c_s, $c_s = 1/15$, the Ag would replace the Hg 45 Å into the bulk, consistent with the previous conservative estimate of Ag diffusion at 30 ML deposition. For a more realistic but still conservative $c_s = 0.2$ the Hg replacement model would give an Ag diffusion length of 140 Å. This distance provides a good estimate of the order of magnitude of the diffusion length, and it is consistent with the estimates given above. To estimate the upper limit on the diffusion length note that at 150 ML deposition the Ag replaces essentially all the Hg in the near-surface. Even if $c_S=1$, there would only be enough Ag to replace the Hg less than 2000 Å into the bulk.

The Au/MCT interface might be expected to be very similar to the Ag/MCT case on the basis of heat-of-formation considerations. As with Ag/MCT, little chemical reaction is seen[3]; however, the Cd and Hg peak areas decrease with increasing Au coverage while the Te peak area remains essentially constant. This suggests that the Au covers up the Hg and Cd while letting the Te diffuse through, or that both Hg and Cd are depleted from the surface. For the case of CdTe both the Cd 3d and Te 3d core level peak areas attenuate quickly with Ag coverage[11], implying that there is little if any diffusion of Ag into CdTe, only an Ag overlayer being formed. Therefore the presences of Hg is critical to the diffusion process. Ag replacement of Hg may not be the only diffusion mechanism. As the inset in Fig. 1 shows, the slope of the c(Ag)-*vs*- deposition curve decreases for low depositions (< 5ML), suggesting that a different diffusion mechanism operates at very low Ag concentration. Interstitial diffusion could carry Ag deep into the bulk at low concentrations; the Ag replacement, with much shorter diffusion length but higher Ag solubility, would dominate for higher concentrations.

Finally, a brief mention is made of the band bending induced by the Ag deposition at low coverages (the high coverage behavior will be described elsewhere[6]). Between 0 and 1 ML coverage, all core level and valence band features shift 0.05-0.1 eV to lower binding energy (BE), indicating that the bands are bending upward by this amount relative to the cleaved surface. A shift of 0.2 eV to lower BE has been seen for Au/MCT[3], and was attributed to Au acting as an acceptor in MCT. A similar explanation probably applies for the Ag-induced band bending at low coverages: the low concentration of Ag acts as a p-type dopant.

In summary, while Ag does not react strongly with MCT upon deposition, it diffuses into the bulk on the order of 10^2 to 10^3 Å by replacing Hg. Only after on the order of 100 ML Ag is deposited does it start to pile up at the surface. The interface is very different in character from reactive metal/MCT interfaces, and also differs (in the nature of the diffusion) from the nonreactive Au/MCT interface. The in-diffused Ag alters the electrical properties of the surface.

This work was supported by DARPA under contract 916610-85.

References

[1] K. C. Mills, *Thermodynamic Data for Inorganic Sulphides, Selenides and Tellurides* (Butterworths, London, 1974).

[2] A. K. Niessen, F. R. de Boer, R. Boom, P. F. de Chatel, W. C. M. Mattens, and A. R. Miedema, Calphad 7, 51 (1985).

[3] G. D. Davis, W. A. Beck, N. E. Byer, R. R. Daniels, and G. Margaritondo, J. Vac. Sci. Technol. A 2, 546 (1984).

[4] G. D. Davis, N. E. Byer, R. A. Riedel, and G. Margaritondo, J. Appl. Phys. 57, 1915 (1985).

[5] D. J. Peterman and A. Franciosi, Appl. Phys. Lett. 45, 1305 (1984); P. Philip, A. Franciosi, and D. J. Peterman, J. Vac. Sci. Technol. A 3, 1007 (1985).

[6] D. J. Friedman, G. P. Carey, C. K. Shih, I. Lindau, W. E. Spicer, and J. A. Wilson, to be published.

[7] G. Rossi, I. Abbati, L. Braicovich, I. Lindau, and W. E. Spicer, Surf. Sci. Lett. 112, L765 (1981).

[8] J. A. Silberman, P. Morgen, I. Lindau, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. 21, 154 (1982).

[9] J. Szajman, R. C. G. Leckey, J. Liesegang, and J. G. Jenkin, J. Electron Spectrosc. Relat. Phenom. 20, 323 (1980).

[10] J. J. Yeh and I. Lindau, At. Data Nucl. Data Tables 32, 1 (1985).

[11] T. P. Humphreys, M. H. Patterson, and R. H. Williams, J. Vac. Sci. Technol. 17, 886 (1980).

Figure Captions

Figure 1.

The Hg 4f, Cd 3d, Te 3d, and Ag 3d core level peak areas, multiplied by their respective atomic sensitivities and normalized to the value for Te at zero deposition, as a function of Ag deposition. The peak areas so normalized are equal to the near-surface concentrations of their respective atoms relative to the bulk Te concentration. The dashed line is the shape expected for the Ag 3d curve if there were no diffusion.

Figure 2.

Fractional composition of the Ag-diffused semiconductor near-surface as a function of Ag coverage.

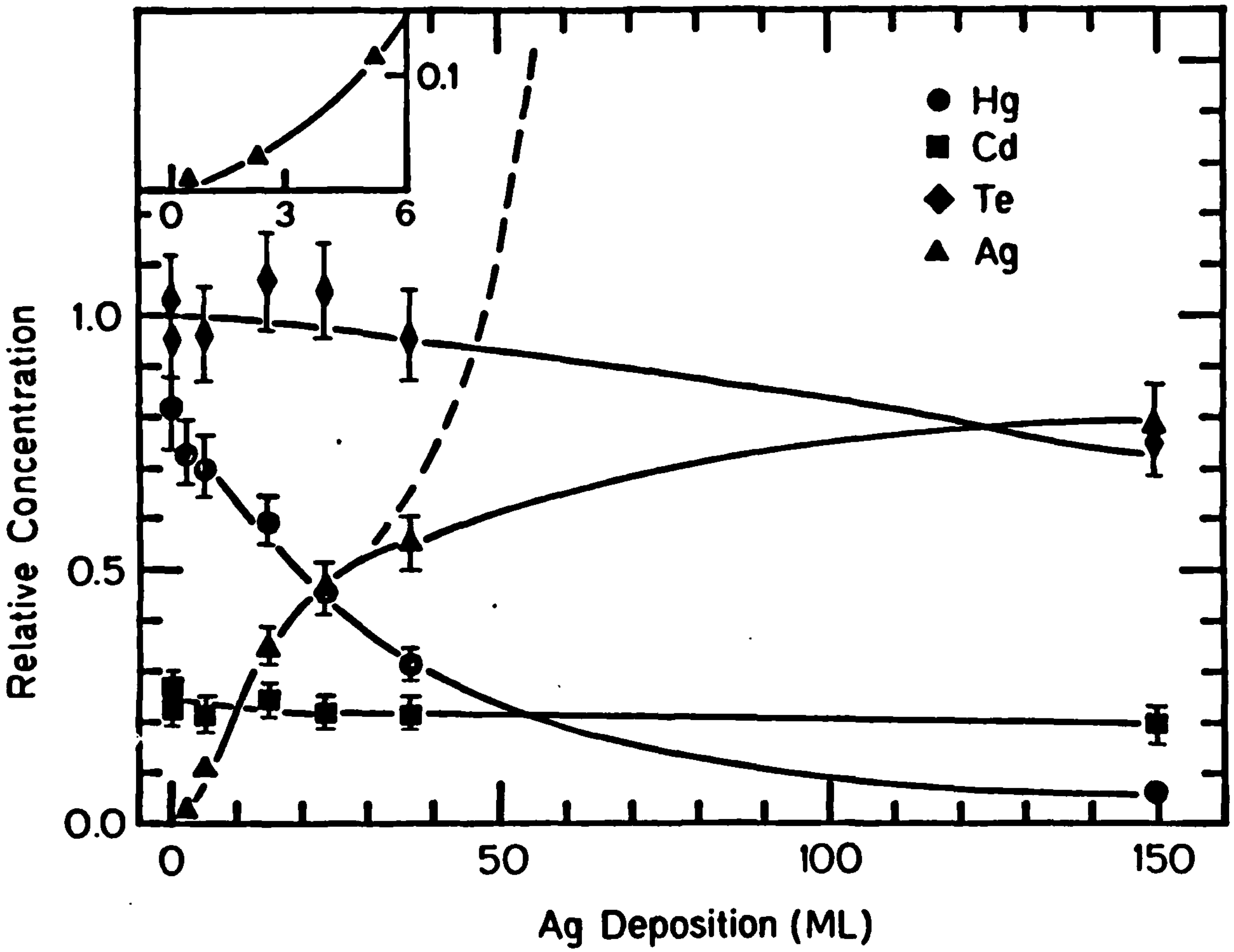

0
3
6
0.1
● Hg
■ Cd
◆ Te
▲ Ag
Relative Concentration
1.0
0.5
0.0
0
50
100
150
Ag Deposition (ML)

Figure 1

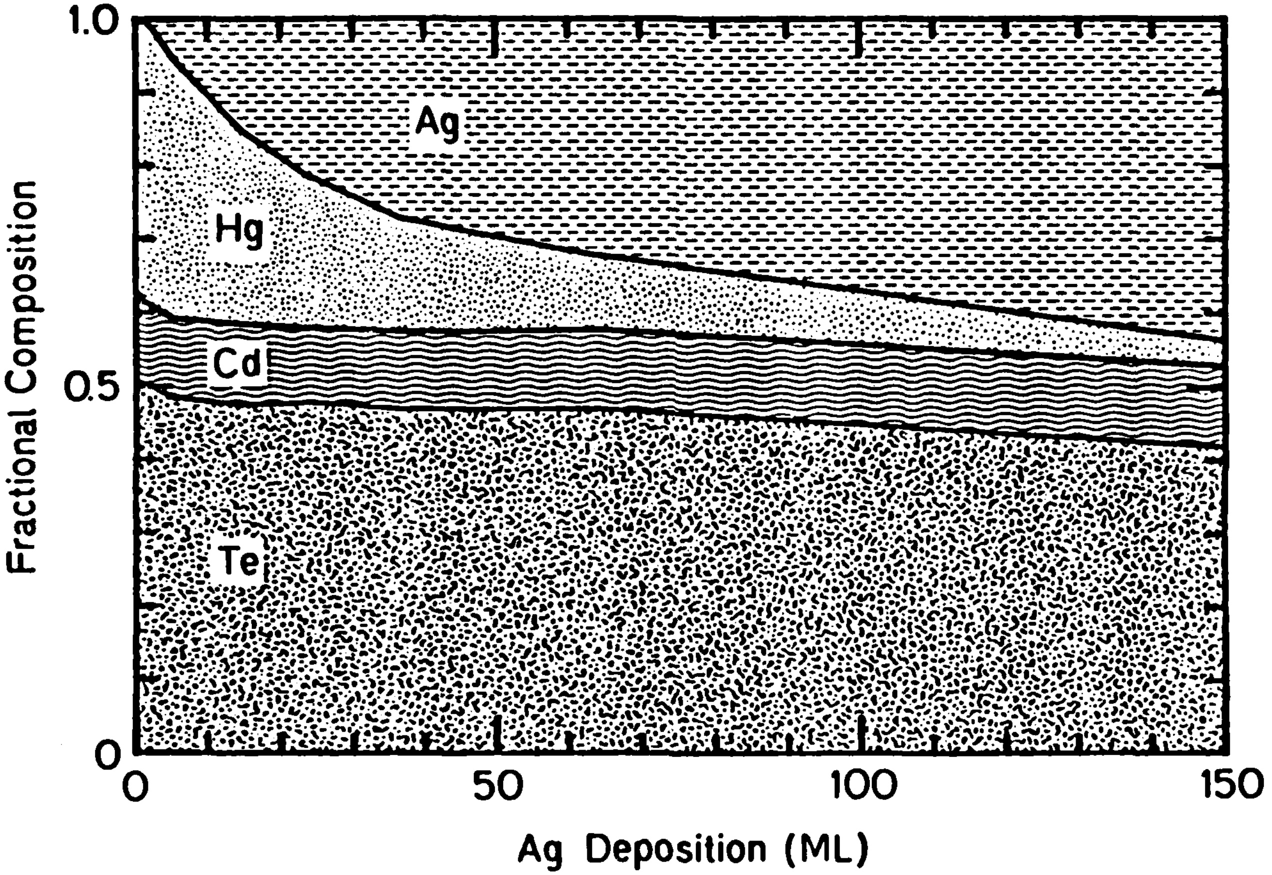

Ag
Hg
Cd
Te
Fractional Composition
1.0
0.5
0
0
50
100
150
Ag Deposition (ML)

Figure 2

3

a. Electronbeam Induced Hg Desorption and the Electronic Structure of the Hg Depleted Surface of $Hg_{1-x}Cd_xTe$

C. K. Shih, D. J. Friedman, K. A. Bertness, I. Lindau, and W. E. Spicer
Stanford University, Stanford, California 94305

J. A. Wilson
Santa Barbara Research Center, Goleta, California 93017

Abstract

Auger electron spectroscopy (AES), X-ray photoemission spectroscopy (XPS), Low energy electron diffraction (LEED), and Angle-resolved ultraviolet photoemission spectroscopy (ARPES) were used to study the electron beam induced Hg desorption from a cleaved (110) $Hg_{1-x}Cd_xTe$ surface and the electronic structure of the Hg depleted surface. Solid state recrystalized $Hg_{1-x}Cd_xTe$ single crystals were used. It was found that the electron beam heating dominated the electron beam induced Hg desorption on $Hg_{1-x}Cd_xTe$. At the electron beam energy used, the electron beam heating extended several thousand angstrons deep. However, the Hg depletion saturated after a few monolayers were depleted of Hg atoms. At the initial stage of Hg loss (only 3%), the surface band bend upward (more p-type). The ARPES spectrum showed the loss of some E vs. k dispersion after 22% Hg atoms were removed from the surface region, and no dispersion was observed after 43% Hg atoms were removed. These results have important implication on the electronic structure of the surfaces and interfaces of which the stoichiometry is altered.

I. Introduction

$Hg_{1-x}Cd_xTe$ (MCT) appears to be unique in that cations are easily removed from the surface region by various surface treatments. In low pressure oxidation[1] and metals on clean MCT surface[2,3], Hg loss is uniquely strong. In chemical preparation of surface with methonal etching[4] appears to cause depletion of Cd. Also Nitz et al[5] reported Hg loss due to use of surface analysis tool using electron beam. As a result it is important for device work to know how this removal of cations effects the electronic and lattice structure. Conversely since Hg can be removed so easily an opportunity is provided to obtain fundamental understanding of the effects of such removal on the electronic and lattice structures. The first need is to provide a sensitive and reproducible means of removing Hg atoms unifomly from the surface. We have achieved this goal by studying the electron beam induced Hg desoption and calibrating the Hg loss as a function of beam parameters. Here we report preliminary studies of the elctron beam induced Hg desorption and the electronic and lattice structure of the Hg depleted surface of cleaved $Hg_{1-x}Cd_xTe$ (110) surface.

II. Experimentals

$Hg_{1-x}Cd_xTe$ crystals with $x = 0.23$ were introduced through a transfer system into a baked UHV system which was equipped with sample cooling and heating capabilities. The sample was cleaved to produce an atomically clean (110) surface. Electron beam was rastered over the sample to produce Hg desorption. By rastering the electron beam over a fine wire mesh with known wire spacing and watching the electron absorbtion image of the wire mesh on the TV monitor, the rastering area was determined. By moving the electron beam across a sharp knife edge, the current pass through the knife was measured as a function of the moving distance, and the electron beam diameter was determined as the distance between the positions of 10% and 90% current rise.

Electron beam induced Hg desoption was first studied and calibrated by rastering the electron beam over a restricted area and varying the substrate temperature. Because Aguer electrons are emitted from the sample during irradiation by the rastering electron beam, Auger electron spectroscopy (AES) was used to monitor the surface composition during electron beam exposure. In order to do LEED and photoemission measurements the electron beam was rastered over the whole surface to produce uniform damage. LEED was used to study the surface crystal structure, angle-resolved photoemmision spectroscopy (ARPES) was used to study the valence band dispersion and the band bending, and x-ray photoemission spectroscopy (XPS) with various detection angles was used to study the composition with different surface sensitivities.

III. Results and discussion

A. Electron beam induced Hg desorption

Figure 1 shows the Auger signals of Hg and Te as a function of electron beam irradiation time at different substrate temperatures. The beam parameters are: $E_0 = 3$ KV, $I_0 = 40$ μA, area = 1 mm x 0.7 mm, and $d_0 \approx 100$ μm, where E_0 is the electron beam energy, I_0 is the beam current and d_0 is the electron beam diameter. With these beam parameters, a calculation which models the power deposited on the sample (appendix A) shows the dynamical heating effect of a thermal spike at a given location with 13 μsec duration and amplitude, $T_{max} - T_{sub}$, estimated between 150 °C - 300 °C. The uncertainty is due to the uncertainty of the thermal conductivity[6]. In Figure 1 one sees that Hg depletion is suppressed with cooling of the substrate. Combining this result with the calculation of the heating effect suggests that the electron beam heating plays the dominant role in the electron beam induced Hg desorption. One also finds that the Hg signal decreases to 60% depletion then saturates, and the Te signal increases by 20%. The surface sensitivities of Hg Auger signal (76 eV) and Te Auger signal (483 eV) are about 5 Å and 15 Å respectively[7,8]. If we assume that for each Hg atom desorpted an excess Te atoms is produced, and that only Hg is removed from the top few monolayers, we can explain that the Te signal increased by 20% while the Hg signal decreased by 60%. One would also expect that the Cd/Te signal remain unchanged if the above assumption is true. By measuring the XPS signal at different photoelectron take-off angles one can change the effective probing depth of XPS. Let us say the off-normal emission angle is θ_i, and the escape depth of the XPS signal is ℓ_0, then the effective probing depth, ℓ_i, equals to $\ell_0(\cos\theta_i)$. In figure 2 we have plotted the cations (Hg or Cd) to Te ratio of an uniformly Hg-depleted surface as a function of $\ell_0(\cos\theta_i)$. The Hg/Te and Cd/Te signals are normalized to that of the cleaved surface, ie. taking both of the Hg/Te and Cd/Te ratio as unity. The abscissa of figure 2 is plotted from zero to ℓ_0 corresponding to $\cos\theta_i$ from zero to unity. We observed that the Cd/Te ratio stays constant while the Hg/Te ratio drops. If one approximates the XPS signal of Hg/Te (normalized to the signal of the cleaved surface) at probing depth ℓ_i which is denoted as $S(\ell_i)$ is the average from the suface to ℓ_i, then the concentration ratio at different depths, $C(\ell)$ can be deconvoluted using the following equation:

$$C[(\ell_i + \ell_{i+1})/2] = [S(\ell_{i+1}) * \ell_{i+1} - S(\ell_i) * \ell_i] / (\ell_{i+1} - \ell_i) \qquad (1)$$

The deconvolution result can not be consider as an accurate representation of the depth profile since the approximation is very crude. However, qualitatively it does give the trend of the Hg/Te ratio as a function of depth. The XPS result is consistent with the above interpretation of AES result -- that for each Hg atom desorpted an excess Te atoms is produced, and that only Hg is removed from the top few monolayers.

Is the electron beam induced Hg desoption on $Hg_{1-x}Cd_xTe$ caused by electronic transition, the so-called Electron stimulated desoption (ESD)[9]? For the ESD phenomena there is a threshold dose of electrons which produce the onset of the desorption, and the desorption is linearly dependent on doses. In our study, for the beam parameters of: E_0 = 2kV, I_0= 10 μA , we observed *no Hg depletion* and the surface LEED pattern was sharp after 3 hours; whereas for the same beam parameters except with I_0 = 30 μA, we observed 22 % Hg depletion in 60 minutes. The electron doses for the two cases are identical which indicates the ESD mechanism can not explain the electron beam induced Hg desoroption in $Hg_{1-x}Cd_xTe$. On the other hand, the later case yields a larger heating effect which indicates the electron beam heating is the dominant effect to cause the release of the surface Hg atoms. However, we can not exclude the possibility of incoporating the electronic transition to break the surface Hg-Te bonds. Although with 3 kV beam energy the electron beam heating extended as deep as a few thousands angstroms, the Hg depletion is limited at the top few monolayers. We think that the surface layer with excess Te film might act as a barrier for the Hg outdiffusion to the surface for the subsquent desorption. The following mechanism is suggested: The electron beam heating causes the release of the Hg atoms from the surface, after the top few monolayers are depleted of Hg the process is limited by the outdiffusion of the subsurface Hg atoms to the surface, and the crystal collapses with excess Te on the surface which acts as diffusion barrier.

B. Electronic and lattice structure of the Hg depleted surface

As Hg atoms were removed from the $Hg_{1-x}Cd_xTe$ surface, LEED was used to measure the suface lattice periodicity, and the valence band-dispersion was measured by ARPES. At the same time, the valence band maximum and the Cd 4d and the Hg 5d core levels were used to detemined the band bending. These measurements were done as a function of Hg atoms removed. The fermi level of the fresh cleaved surface was at the VBM, which indicated a p-type surface. At only 3% Hg loss (determined by XPS) we first observed that both the Hg 5d and Cd 4d levels (measured at 21 eV photon energy) shifted upward by about 0.1 eV (figure 3(a)). This was attributed to the band bending upward and the surface becoming degenerate p-type. However, the LEED pattern and the valence band dispersion were identical to that of the cleaved suface. The previous result can be understood by assuming that Hg vacancies have formed and act like acceptors. The Hg5d and Cd 4d core level shifts as a function of the surface Hg loss are shown in figure3(b). The kinetic energies of these levels in UPS measurements are less than 10 eV thich corresonds to the escape depth of about 20Å[7]. If one assumes that the Hg 5d emission comes from the underlying stoichiometric region of $Hg_{1-x}Cd_xTe$ where the UPS is still sensitive to, the movement of the 5d core level at higher Hg loss indicates that the underlying $Hg_{1-x}Cd_xTe$ becomes more p-type because a large amount of Hg vacancies have formed. At higher Hg depletion the ARPES spectrum showed the valence band E vs. k dispersion was lost. In figure 4(a) and 4(b) the ARPES normal emission spectra with hυ from 17 eV to 21 eV of both the cleaved surface and the suface with 43% Hg loss are shown respectively. For the cleaved surface the strong dependence of photoemission peak positions on the incident photon energies represents the valence band E vs k dispersion[10]. This dispersion relation was not observed on the Hg depleted surface. Instead, two broad peaks with stationary binding energy at 4.8 eV and 1.4 eV were observed. No LEED pattern was observed for this Hg-depleted surface indicating that the lattice periodicity also disappeared. In table 1 we summarize the results of the LEED patterns and the band dispersions as a function of Hg loss. An interesting result is that one can remove at least 7% of the cations (corresponding to 9 % Hg loss) within approximately 20 Å of the surface without loss of the band dispersion. It should be noted that our LEED measurment which is very suface sensitive (about 5Å) also showed a very sharp LEED pattern. The loss of the E vs k dispersion relationship tells us that the elctrons in this region are no longer Bloch wave like, and one can say that there is no "Bloch band". However, one should notice that in amorphous Si, although there is also no "Bloch band", one can still talk in terms of a valence band and conduction band. The difference is that in a "Bloch band" electrons travel as Bloch waves, and in amorphous material electrons travel by a other mechanism.

Is there a change in the x value and thus a change in the band gap corresponding to the change in the Cd/Hg ratio? It is hard to imagine that after a large amount of Hg loss, there is still a band structure of $Hg_{1-x}Cd_xTe$ with larger x value, unless most of the missing Hg atoms are replaced by the Cd atoms, ie. the (Hg+Cd)/Te remains unity, and the crystal structure are still preserved. In our study we found that the Cd/Te ratio remained constant as Hg was decreased, and the valence band lost its E vs k dispersion. However, one can not say there is no band structure even though the E vs k dispersion is lost (as discussed above for the example of amorphous Si). What possibly happens in our case is that the mixtures of the clusters of Te, CdTe and $Hg_{1-x}Cd_xTe$ with various x values are formed, and the band structures associated with these clusters are very complex depending on the size and the crystallinity of the clusters. Morphologically how these clusters are arranged is very impotant to the device work. For example, if clusters of amorphous Te (which is metal like) spread out beneath a passivation layer, it will cause the leakage of the device.

IV. Conclusion

We have studied the electron beam induced Hg desorption and the effect of the loss of Hg on the electronic and lattice structure of the cleaved $Hg_{1-x}Cd_xTe$. The major results are summarized as follows:

(1) The ESD[8] mechanism can not explain the electron beam induced Hg desorption on the cleaved $Hg_{1-x}Cd_xTe$ where the electron beam heating is the dominant effect in causing the release of the surface Hg atoms. After the top few monolayers are depleted of Hg atoms, the crystal collapses leaving the surface with excess Te atoms which act as a diffusion barrier for Hg outdiffusion.

(2) At the initial stage of the Hg loss (only 3%), the surface crystal structure and the band structure are preserved and the surface becomes degenerate p-type. Suprisingly, good lattice structure and band dispersion are still preserved up to 9% Hg loss. At higher Hg loss, the surface cyrtal structure and band dispersion are gradually lost.

As mentioned earlier, the $Hg_{1-x}Cd_xTe$ surface stoichiometry is easily altered[1,2,3,4,5]. The result of this on the electronic structure is very important for the device application. For example, if type conversion happens at the interface, it will create an extra junction which affects the device performance, and if band gap narrowing happens at the interface, it will cause the leakage, and etc. The information needed are so crucial and yet little is known to date. What we presented here represents the first step in an effort to understand the electronic structure of the stoichiometry altered surfaces and interfaces.

Acknowledgement

This work was supported by DARPA contract No. 916610-B5.

APPENDIX A

ELECTRON BEAM HEATING EFFECT

The power deposited in a solid sample by an electron beam can cause heating of the sample. The differential equation of the heat conduction is

$$\partial T/\partial t = Q/C_p - \nabla(\alpha \nabla T) \qquad \text{(A1)}$$

where T - temperature

t - time

Q - power deposited per unit volume

C_p - heat capacity per unit volume

α - thermal diffusivity.

For a given position, the first term on the right hand side of the equation is the energy deposited in from the electron beam and the second term is the energy flows out resulting from the gradient of heat flow.

Steady State: According to Pittway[11], the steady state temperature rise is directly proportional to the beam power ($I_p \times E_p$) and inversely proportional to both the thermal conductivity and the beam diameter. At the center of the beam the surface temperature rise is given by[12] :

$$T_m - T_s = P/[(8\pi)^{1/2} K_s r_b] \qquad \text{(A2)}$$

where T_m - temperature at the center of the beam which is maximum

T_s - substrate temperature

P - power of the electron beam

K_s - thermal conductivity

r_b - beam diamete.

For materials like $Hg_{1-x}Cd_xTe$, where the thermal conductivity is a function of temperature, K_s is the effective thermal conductivity which lies between the maximum value and the minimum blaue. For our purpose, let's assume K_s is a constant.

Transient Behavior: To know the transient heating effect at the center of the beam one needs to solve equation (A1) $\partial T/\partial t = Q/C_p - \nabla(\alpha \nabla T)$. To a good approximation, on the surface the term $\nabla(\alpha \nabla T)$ is proportional to the difference between the surface temperature and the substrate temperature, $\nabla(\alpha \nabla T) = \beta(T-T_s)$. Then equation (A1) becomes

$$\partial T/\partial t = Q/C_p - \beta(T-T_s) \qquad (A3)$$

At steady state $\partial T/\partial t = 0$, we obtain

$$Q/C_p = \beta(T_m-T_s) \qquad (A4)$$

Thus the equation becomes

$$\partial T/\partial t = (Q/C_p)[1-(T-T_s)/(T_m-T_s)] \qquad (A5)$$

The solution is

$$T(t)-T_s = (T_m-T_s)/[1-\exp(-t/\tau)] \qquad (A6)$$

Where $\tau = C_p(T_m-T_s)/Q$. Q can be approximated as $Q = P/\pi r^2{}_b R_g$, where R_g is called the gruen range which represents the range of the electron beam power deposited into the solid. T_m can be obtained from (A2).

Dynamical Heating effect of a rastering electron beam: We have obtained the heating effect at the center of an electron beam for the steady state and transient situations. We can then apply this knowledge to the heating effect of a rastering electron beam.

For a given position when the electron beam on a pass through, the temperature rises up and then falls down. The rising and falling time is characterized by τ and the duration time t_0 is determined by the spot size.

d_0, the length of the rastering line, l_0, the TV frame rate, f, and the number of rastering lines per frame, n,

$$t_0 = (d_0/l_0)/nf \qquad (A7)$$

The amplitude of the thermal spike then can be obtained from (A6) and equals to $(T_m-T_s)/[1-\exp(-t_0/\tau)]$. When the scanning line spacing is smaller than the spot size, similar thermal spike occurs when the next scanning line passes through, and the amplitude is smaller.

References:

1. J. A. Silberman, D. Laser, I. Lindau, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. **A 3**, 222 (1985).

2. D. J. Friedman, G. P. Carey, C. K. Shih, I. Lindau, W. E. Spicer, and J. A. Wilson, this conference proceeding.

3. G. D. Davis, N. E. Byer, R. R. Daniels and G.Margaritondo, J. Vac. Sci. Technol. **A 1**, 1726 (1983)..

4. R. T. Talasek, and A. J. Syllaion, J. Electrochem. Soc. **132** (3), 656 (1985).

5 H. M. Nitz, O. Ganschow, U. Kaiser, L. Wiedmann, and A. Benninghoven, Surf. Sci. **104**, 365 (1981).

6. Thermo Physical Prop. of Matter **1**, 1267 and 1321, IFI/Plenum. Newyork-Washington (1972).

7. H. Ibach, Topics in Current Physics, **4**, 4, Springer-Verlag, Berlin Heidelberg New York (1977).

8. J. Szajmaan, R. C. G. Leckey, J. Liesegang, and J. G. Jenkin, J. of Electron Spectroscopy and Related Phnomena, **20**, 323 (1980).

9. T. E. Madey and J. T. Yates, Jr., J. Vac. Sci. Technol. **8**, 525 (1971).

10. J. A. Silberman, D. Laser, C. K. Shih, D. J. Friedman, I. Lindau, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. **A 3**, 233 (1985).

11. L. G. Pittaway, Br. J. Appl. Phys. **15**, 967 (1964).

12. Ali A. Iranmanesh and R. F. W. Pease, J. Vac. Sci. Technol. **B 1**, 91 (1983)

Table 1. The ARPES and LEED results at various surface Hg depletions (as determined by XPS) at room temperature.

% Hg LOSS (XPS)	ARPES RESULTS	LEED RESULTS
0	dispersion in valence band	sharp 1x1 pattern
3	"	"
9	"	"
19	loss of some dispersion	weak 1x1 pattern
32	loss of more dispersion	no pattern
43	very little dispersion	"
51	no dispersion	"

Figure Captions

Figure 1 Auger signals (peak-to-peak height) of Hg (75 eV) and Te (487 eV) as a function of electron beam irradiation time at different substrate temperatures. The escape depths of the Hg and Te Auger signals are about 5Å and 15Å respectively.

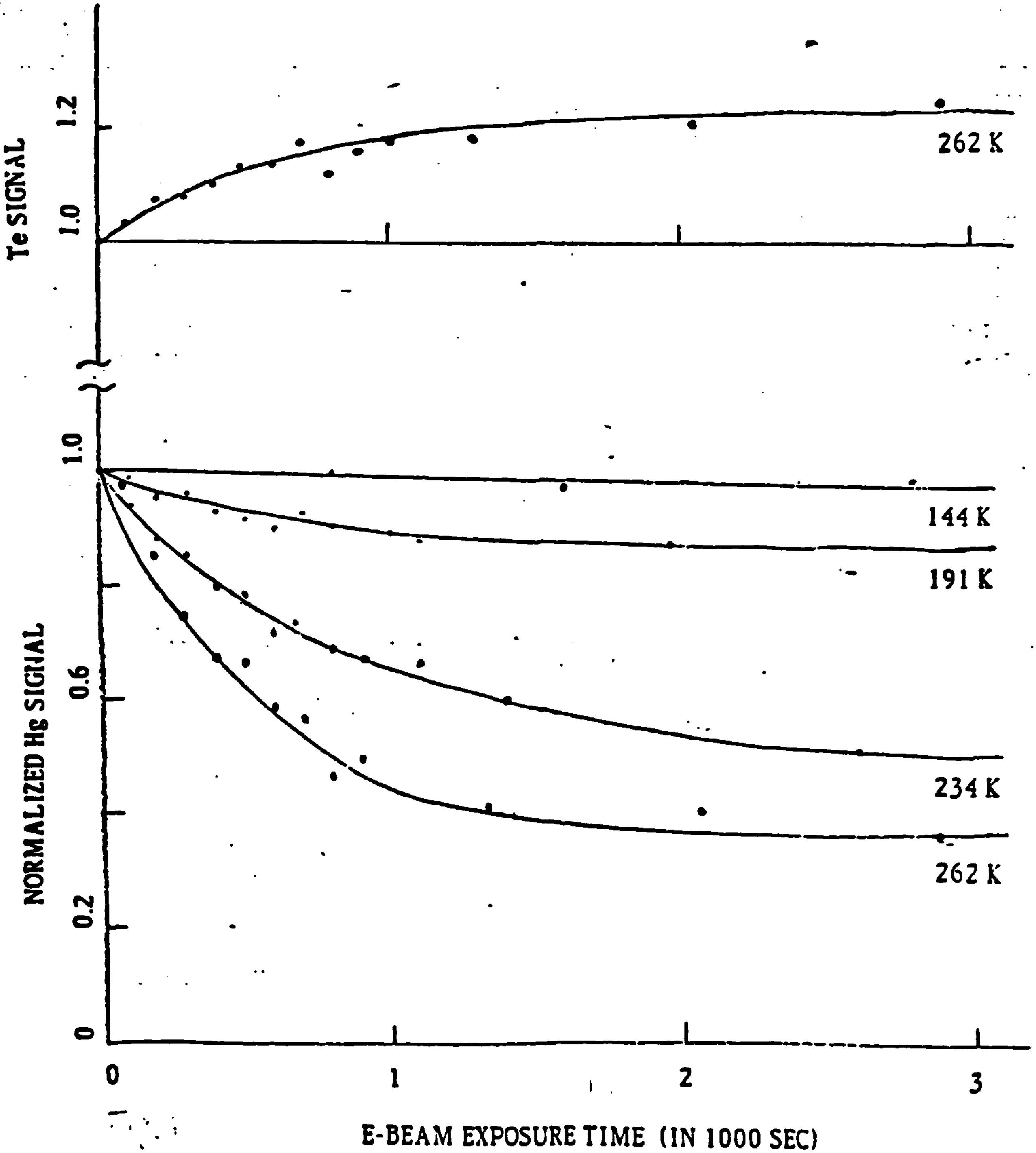
Figure 2 XPS signals of cation (Hg or Cd) to Te ratios as a function of the effective probing depth, $\ell_0 \cos\theta_1$. θ_1 is the off-normal emission angle. ℓ_0 is the escape depth of XPS, which is about 20Å. The deconvolution result is obtained using equation 1. (see text for details).

Figure 3a The binding energies of Hg 5d and Cd 4d core levels for the clean cleaved surface and the surface with 3% Hg loss. The shift upward of both Hg 5d and Cd 4d after 3% Hg loss is attributed to the band bending. The escape depths for these core level electrons are about 20Å. The deconvolution result is obtained by using equation1. (see text for details).

Figure 3b The binding energy shifts of Hg 5d and Cd 4d as a function of the percent Hg loss. The energy resolution of the relative movement of the core levels is better than 0.05 eV.

Figure 4a ARPES normal emission spectra of the cleaved surface ($x = 0.233$) valence band at different photon energies. The strong dependence of the peaks on the photon energy respresents the E vs k dispersion. For detailed discussion about ARPES normal emission spectra of $Hg_{1-x}Cd_xTe$, see reference 10.

Figure 4b ARPES normal emission spectra of the uniformly Hg-depleted surface (43% Hg loss as determined by XPS). The strong dependence of peaks on the photon energy has disappeared, instead two peaks with stationary binding energies at 4.8 eV and 1.4 eV are observed. The detailed discussion of this E vs k disappearance is in the text.


Te SIGNAL
1.0
1.2
262 K
NORMALIZED Hg SIGNAL
0
0.2
0.6
1.0
144 K
191 K
234 K
262 K
0
1
2
3
E-BEAM EXPOSURE TIME (IN 1000 SEC)

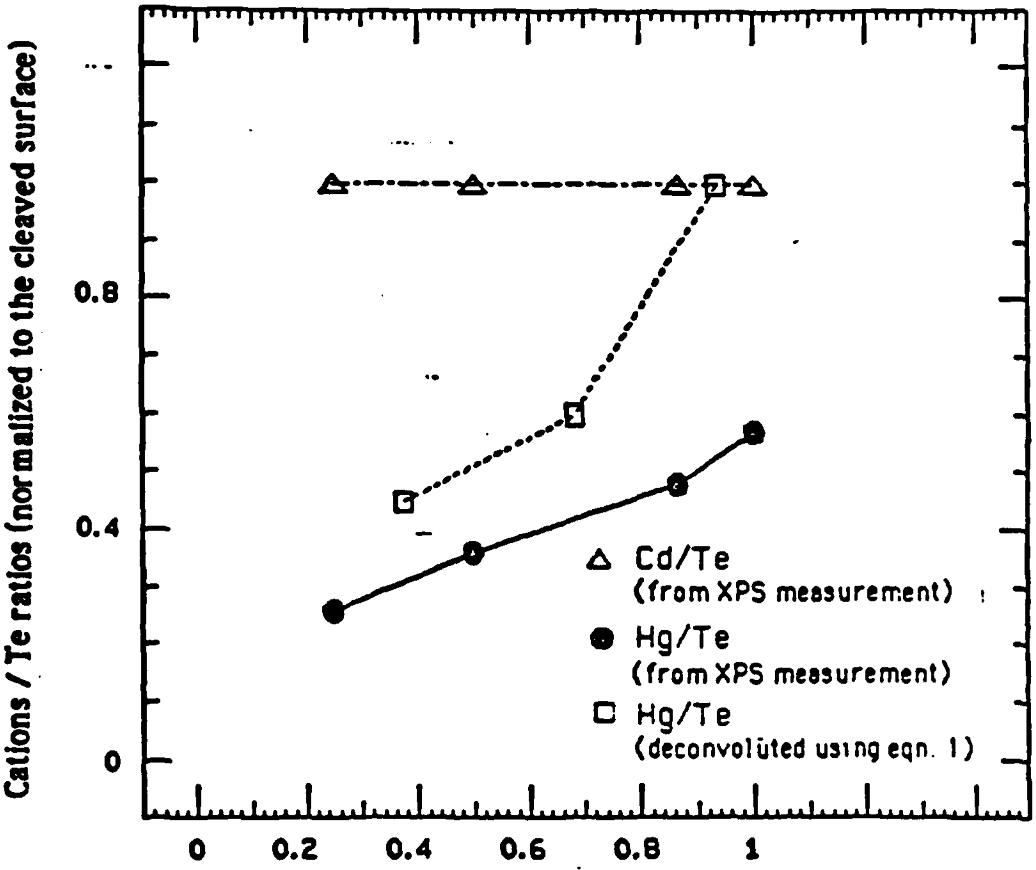
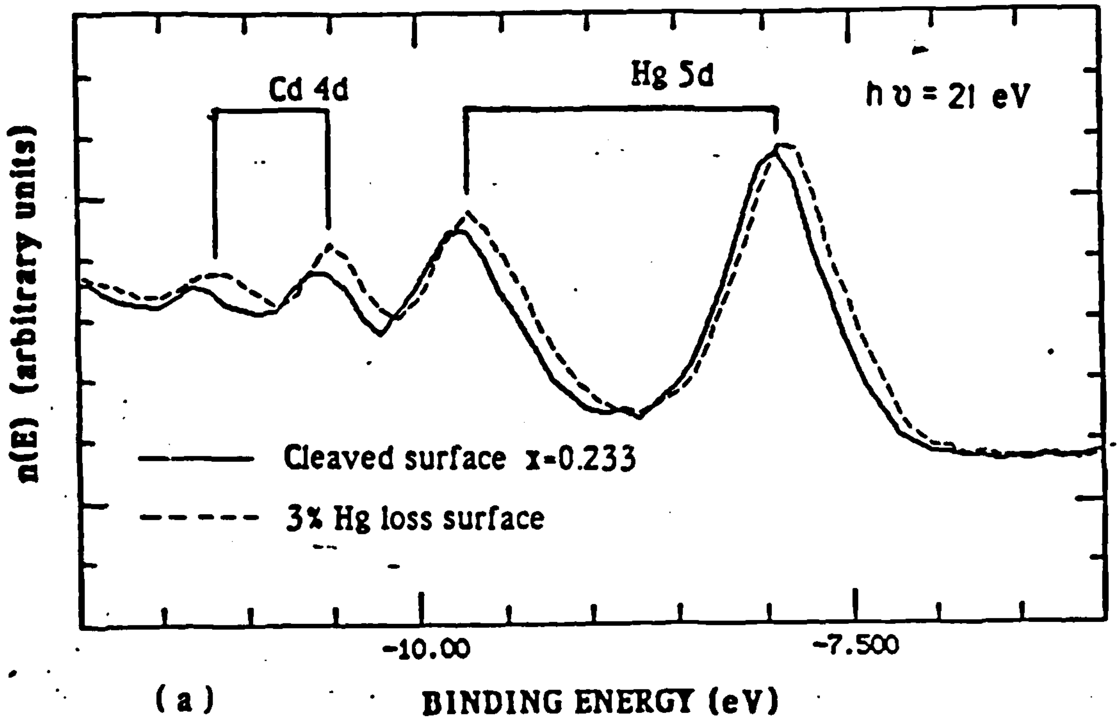
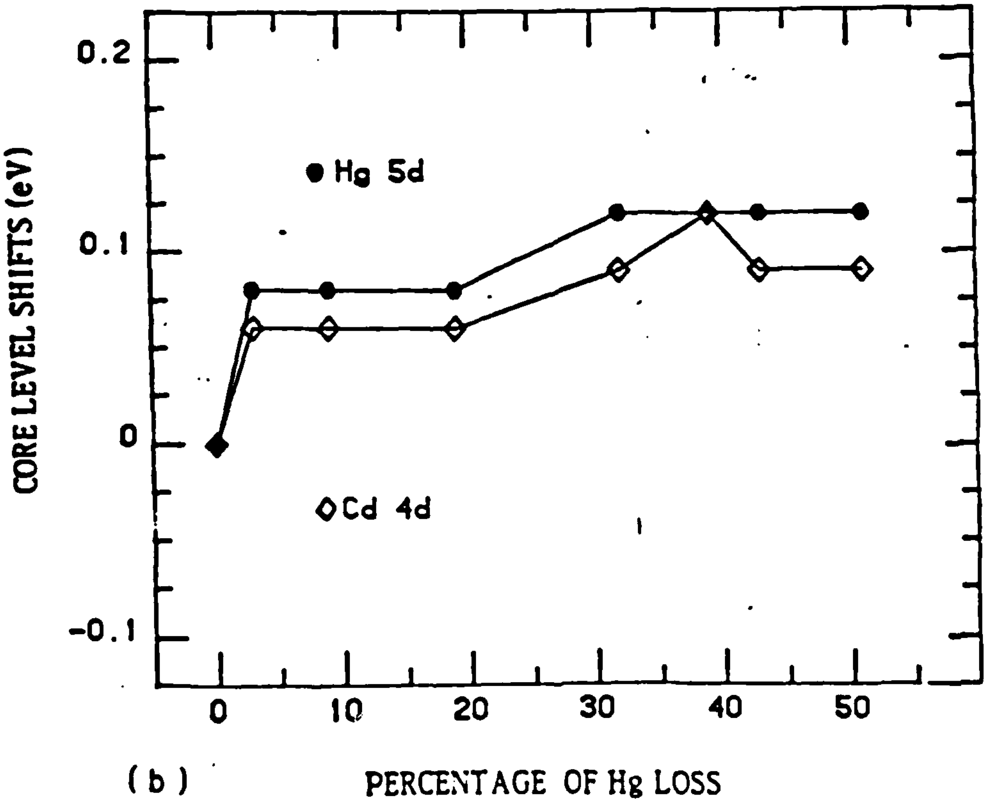


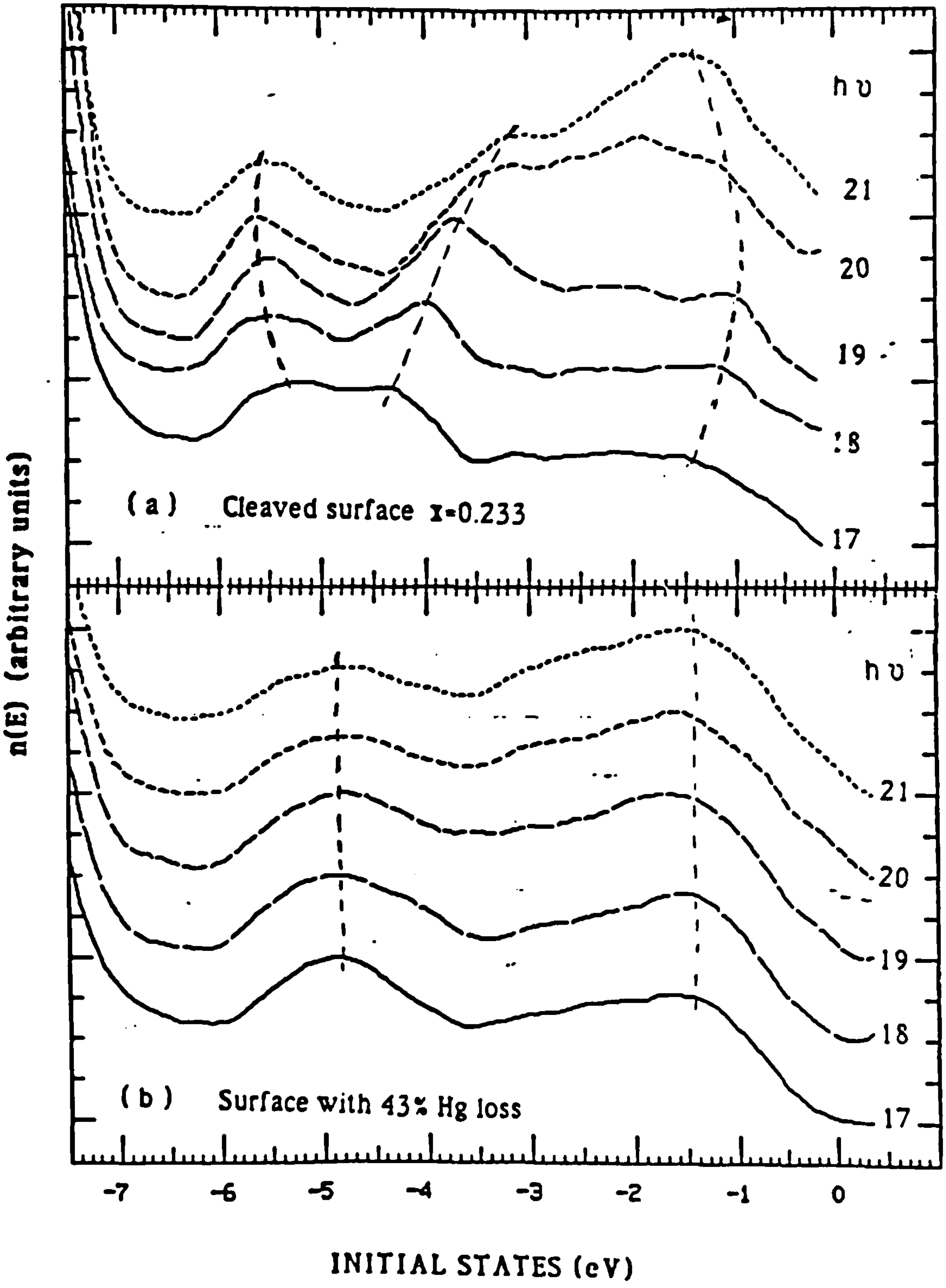
Fig. 2 Probing depth (in terms of the escape depth $\ell_o \approx 20$ Å)


Cd 4d
Hg 5d
h υ = 21 eV
n(E) (arbitrary units)
Cleaved surface x=0.233
3% Hg loss surface
-10.00
-7.500
BINDING ENERGY (eV)

(a)


0.2
0.1
0
-0.1
CORE LEVEL SHIFTS (eV)
● Hg 5d
◇ Cd 4d
0
10
20
30
40
50
PERCENTAGE OF Hg LOSS

(b)


n(E) (arbitrary units)
h υ
21
20
19
18
17
(a) Cleaved surface x=0.233
h υ
21
20
19
18
17
(b) Surface with 43% Hg loss
-7
-6
-5
-4
-3
-2
-1
0
INITIAL STATES (eV)

F:

(1) Photoemission Study of Etched HgCdTe Surfaces Revealing Cd Loss

G. P. Carey, A. Wahi, J. A. Silberman, C. M. Stahle and W. E. Spicer

I. Introduction

The complex chemistry of the ternary alloy system $Hg_{1-x}Cd_xTe$ is manifested in many aspects of the growth, processing, and characterization of this II-VI semiconductor. The large differences in the activities of the constituents and the case of diffusion lead to difficulties in controlling crystal growth and can lead to off-stoichiometry on the cation sublattice and a non-ideal cation to anion ratio. The weak Hg bonding in the lattice [1,2] which contributes to the disparities in the activity coefficients also increases the density of point defects (e.g. Hg vacancies) which can be electrically active and hence markedly affect the electrical properties of the semiconductor, as reported in section III of this report. In that section it is shown that one must be careful when using an electron beam as a probing source during characterization or processing since it may preferentially remove the cation Hg from the lattice and hence alter the electronic band structure in the surface of the alloy. But, because of the nature of the bonding in this lattice, very gentle handling of this system (relative to Si, e.g.) during both processing and characterization can markedly affect the composition of the alloy and hence impede potential device performance. It is therefore of great importance to study the effect on the surface and bulk electronic structure of $Hg_{1-x}Cd_xTe$ of the mechanical, chemical, and thermal interactions that will occur during growth, processing and characterization of the alloy.

Many well-controlled studies of the surface electronic structure and chemistry have been conducted in this program on atomically clean surfaces that have been produced by cleaving HgCdTe crystals in UHV environments. The initial stages of oxygen uptake on the surface have been studied in a contamination-free atmosphere, metal-HgCdTe interfaces [4], and the band structure [5] have all been studied on cleaved HgCdTe to give insights into the bonding nature of the alloy and its interaction with oxygen and metal overlayers.

Using these fundamental studies as a baseline, the extension of our investigations to those technologically important surfaces of HgCdTe that have been cleaned by chemical etching is important in enabling us to understand the complexities of these "device" surfaces. Composition changes might occur upon etching of the material with bromine methanol (BrM) since this etchant has shown anisotropic reaction etch rates on HgCdTe, with Cd removal being the fastest [6]. The goal of this research is to determine the composition, degree of contamination, and the production of oxides on these technologically important surfaces and to determine the pinning of the fermi level at the surface ($E_{f,s}$). Our ultimate goal is to fundametally understand the HgCdTe surface in which devices are made by combining our knowledge of the "theoretical" cleaved surfaces with the studies of these technologically important device surfaces.

II. Experimental

Substrates of HgCdTe that had undergone surface preparation similar to that of production line samples were prepared at SBRC and shipped to Stanford in an inert N_2 ambient. The method of shipping was worked out to minimize surface changes. The samples were shipped in batches of six to facilitate their transfer into the VG ESCALAB via a nitrogen glove bag. Care was taken to minimize exposure to air during this transfer. Once inside the fast entry lock of the chamber the samples were brought down to UHV within a few minutes and then transformed into the analytical chamber. UPS (He I, HeII) and XPS (MgK_α) were performed on each sample. The details of this procedure are listed in section III of our previous *semi-annual* report [4].

To provide for calibration of the sensitivity factors for the hemispherical energy analyzer (HEA) of the VG ESCALAB and to determine the extent of the contamination of the samples during their transfer to the chamber via the nitrogen glove bag, a HgCdTe (x = 0.23) crystal was cleaved in the glove bag and then transferred to the chamber in the same manner as the etched surfaces. (It is not possible to cleave until UHV within the ESCALAB.) An oxidized piece of CdTe was also introduced and was gently sputtered (1 keV Ar, 3 min.) to facilitate calculation of the sensitivety factors, UPS and XPS were performed on each sample. The HgCdTe sample was analyzed twice: 2 hours and 72 hours after being introduced into the chamber.

III. Results

The PES results obtained thus far in this ongoing study are located in Table I. The samples were provided by SBRC and chosen to have x-values near 0.30. Table I lists the surface fermi level and composition for each-sample studied thus far as determined by PES. Whether or not excess Te and/or Te oxide exists at the surface is also noted. Other sample parameters such as growth type, bulk doping, bulk x-value, and surface preparation procedure are listed for comparison. The most salient feature in all this data is a consistent Cd loss revealed by all samples of known bulk x-value. The samples that have 50% transmission cut off values are currently being correlated to x-value.

In most cases, the surface fermi level ($E_{f,s}$) is pinned 0.15 eV ± 0.05 above the valence band maximum (VBM) for both bulk n and p-type samples.

There are notable exceptions to the pinning at 0.15 eV ± 0.05; in particular the vapor transport sample ($E_{f,s}$-VBM = 0.30 eV) and the LPE grown HgCdTe on CdZnTe substrates ($E_{f,s}$-VBM = 0.0 and 0.10 eV) show aberrations that merit a closer look into more similarly prepared samples but with known x-values. In general, the statistics (i.e. more samples studied) must be improved in order to draw more qualitative conclusions from these values.

In this continuing study it is important to refine some of the analysis depicted in this table. Not only must more samples be studied, but the accuracy of our quantitative analysis must be refined. The surface x-values were calculated using the relation:

$$x = \frac{C_{Cd}}{C_{Cd} + C_{Hg}}$$

where $$\%Cd = \frac{C_{Cd}}{C_{cd} + C_{Hg} + C_{Te}}$$

and $$C_n = \frac{A_n/S_n \times 100}{\sum_i A_i/S_i}$$

where A_n is the XPS peak area of element n and S_n is the atomic sensitivity factor of this peak and i is summed over Hg,Cd, and Te. The atomic sensitivity factors used in generating this table were obtained from the Phi handbook [7]: 2.1 for the Hg $4f_{7/2}$, 2.55 for the Cd $3d_{5/2}$ and 4.0 for the Te $3d_{5/2}$. As outlined in section II of this report, the atomic sensitivity factor is defined as

$$S = [\lambda(E)T(E)\sigma(E)]^{-1}$$

where λ(E) is the escape length of the photoexcited electron at kinetic energy E, and T(E) is the transmission of the analyzer and σ is the partial photoionization cross section (PPCS). The λ(E) and σ(E) are fixed for our MgK_α X ray source and the analyzed core peaks, therefore the T(E) may differ from the one that we have assumed and hence lead to a systematic error in our calculation since the analyzer used in the ESCALAB is different from that used to generate the data in the Phi handbook. The sensitivity factors of our analyzed peaks for the HEA, i.e. the ESCALAB, [8] are: 2.4 for the Hg $4f_{7/2}$, 3.8 for the Cd $3d_{5/2}$, and 5.2 for the $Te3d_{5/2}$. As can be readily seen, qualitatively the Cd loss would be calculated to be approximately 10% more at the surface if these S_n values are used. Therefore, lower surface x values are expected than those listed in the table. Analysis on the cleaved HgCdTe and sputtered CdTe aimed at calculating more accurate S_n's is currently being performed in order to give us better confidence in our values. These more accurate values will be listed in the final report.

One of the difficulties in assigning much significant significance to this investigation thus far is that it has been diffuclt in obtaining some of the relevant materials parameters (e.g. x value, doping) that are essential to complete a meaningful analysis. Our powerful surface science techniques can only look at a resultant surface; we need to have the pertinent information as to sample preparation history and relevant parameters. SBRC is working closely to facilitate the generation of this information. More samples are planned with this in mind.

IV. Discussion

The major finding in our PES surface analysis of all the samples studied thus far is the 10 to 25% Cd loss at the surface of the etched HgCdTe. Since we are only looking at the top ~ 20 Å of the surface, the depth at which this depletion occurs is not known. On investigation of future samples we will obtain some qualitative depth profile information by taking XPS spectra of elemental peaks of differnt binding energy in addition to the peaks used in this analysis. In a similar study of BrM etching on CdTe and CdMnTe, Feldman et al. [9] estimated the Cd depletion to be 10 Å and 20 Å , respectively, using Auger sputter profiling. They did not report the degree of Cd depletion, however. Talasek et al. [6] reported enhanced Cd reaction rates relative to Hg and Te with BrM etching, thus our findings agree qualitatively with both of those studies. None of these electron spectroscopies are able to obtain information into the uniformity of the composition across the surface, therefor these studies do not rule out the possibility of a non-uniform surface composition.

The loss of Cd at the surface elicits questions as to the crystalline integrity of the lattice after etching the surface. If the HgCdTe loses crystallinity at the surface, then the physics that give rise to the electronic band structure will no longer hold. As discussed in the previous section of this report, surgical removal of 10-25% of the cation Hg with an electron beam results in some loss of dispersion in the valence band of HgCdTe. This may play a role in the surfaces of devices made of etched material, giving rise to leakage currents, recombination, and possible noise sources. Performing LEED on these etched surfaces should give some insight into the crystallinity of these etched surfaces. The studies performed by Shih (Section III of this report) with the preferential removal of the cation Hg from HgCdTe and its effect on the electronic band structure should give tremendous insight into these studies.

The initial analysis of the data on the cleaved HgCdTe reveals that the cleave produced in the nitrogen glove bag is much cleaner than the etched samples, with very little of the ubiguitous oxygen and carbon contaminating the surface. This suggest that our glove bag transfer procedure does not introduce appreciable contamination and that the contamination present on etched samples is introduced during the etching process or during the ~ 24 hours the samples reside in the N_2 envelope. The cleaved sample also acts as a calibration standard to determine the sensitivity factors for the HEA. Preliminary analysis indicates that the HEA sensitivity factors give more accurate quantitative agreement than the Phi sensitivity factors. Therefore as our data analysis continues and the quantitative accuracy refined, it now appears that the Cd loss appears to be much larger than we first estimated.

V. Summary

Detailed analysis of the surface of etched HgCdTe using XPS and UPS reveals 10 to 25% Cd loss in the top ~ 20 Å of the lattice which is consistant with the anistropic nature of the BrM etchant with HgCdTe. The depth profile is not yet determined, and the lateral uniformity of this loss is not obtainable with our analytical PES tools. The etched surfaces contain much more contamination than the atomically clean cleave, and most also exhibit oxidized Te at the surface. Most surfaces have $E_{f,s}$ pinned at 0.15 eV ± 0.05 above VBM, with few exceptions. LPE layers grown on CdZnTe and large x values at the surface ($x>0.27$) give a larger spread in $E_{f,s}$. Most of the histories of each sample have been obtained, and future analysis will be enhanced by completing this knowledge. The statistical fluctuation among the data taken from the surface can be reduced by further studies. Thus future work is planned to acquire better statistics.

References

1. W. A. Harrison, J. Vac. Sci. Technol., A 1, 1672 (1983).
2. W. E. Spicer, J. A. Silberman, I. Lindau, A-B. Chen, A. Sher, and J. A. Wilson, J. Vac. Sci. Technol. A 1, 1795 (1983).
3. J. A. Silberman, D. Laser, I. Lindau, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. A 3(1), 222 (1985).
4. HgCdTe Surface Study Program Semiannual Technical Report, Section V, Jan. 1, 1985.
5. J. A. Silberman, D. Laser, C. K. Shih, D. J. Friedman, I. Lindau, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. A 3 (1) 233 (1985).
6. R. T. Talasek and A. J. Syllaious, J. Electrochem.Soc., 132 (3) 656 (1985).
7. Phi Handbook of X-ray PHotoelectron Spectroscopy (1979).
8. VG ESCALAB Users handbook
9. R. D. Feldman, R. L. Opila, and P. M. Bridenbaugh, J. Vac. Sci. Technol. A 3, 1988 (1985).

TABLE I

Sample	$E_{f,s}$ - VBM (eV)	Bulk doping	Growth type	x-Value Bulk	x-Value Surface	Surface Prep A	Surface Prep B	Surface Prep C	Notes
LPE 148(A)	0.15	n	LPE	3.8μm*	0.20	1st	2nd		1,2,3
LPE 148(B)	0.15	n	LPE	3.8μm*	0.21		1st		3
ZH247(A)	0.15	p	HZM	0.30	0.23	1st	2nd		1,2
ZH 247(B)	0.15	p	HZM	0.30	0.23		1st		1,2
RCS-241(A)	0.15	n	SSR	0.30	0.24	1st	2nd		1,2
RCA-241(B)	0.15	n	SSR	0.30	0.26		1st		1,2
	room T low T								
LPE-148(A)	0.15 0.20	n	LPE	3.8μm*	0.22		1st		2,3
LPE-148(B)	0.15 0.15	n	LPE	3.8μm*	0.22		1st	2nd	3
RSC-241(A)	0.20 0.20	n	SSR	0.30	0.24	1st	2nd		2
RCS-241(B)	0.20 0.25	n	SSR	0.30	0.24		1st		2
ZH247-3C	0.15 0.20	p	HZM	0.30	0.24	1st	2nd		1,2
ZH247-3D	0.15 0.20	p	HZM	0.30	0.23		1st		2
RCS-209-1-13(A)	0.20	p	SSR	0.35	0.33	1st	2nd		1,2
RCS-209-1-20(A)	0.20	p	SSR	0.35	0.30	1st	2nd		1
F6067(B)	0.15	-	LPE	4.4μm*	0.24		1st	2nd	1,2
VTIE-28C(B)	0.30	p	VT	4.9μm*	0.27		1st	2nd	1,2,3
ZH258-2D(B)	0.20	n	HZM	0.30	0.25		1st	2nd	
ZH286-2J(A)	0.15	n	HZM	0.30	0.27		1st		2
ZH286-2J(B)	0.25	n	HZM	0.30	0.28	lactic/nitric etch			1,2
ZH286-1E(A)	0.10	n	HZM	0.30	0.29	1st	2nd		2
ZH286-1E(B)	0.10	n	HZM	0.30	0.26		1st	2nd	2
VII 544.1	0.20	p	LPE	3.9μm*	0.34		1st	2nd	2,3
VII 544.2	0.0	p	LPE	3.9μm*	0.35		1st	2nd	2,4
IV 489.1	0.15	p	LPE	3.7μm*	0.25		1st	2nd	1,2,3
IV 489.2	0.10	p	LPE	3.7μm*	0.23		1st	2nd	1,2,4
RCS-184-2-8(A)	0.35	p	SSR	0.39	0.33		1st	2nd	A vs. B face 1
RCS-184-2-8(B)	0.30	p	SSR	0.39	0.28		1st	2nd	1,2
RCS-163-3-13(A)	0.20	p	SSR	0.22	0.17		1st	2nd	2
RCS-163-3-13(B)	0.20	p	SSR	0.22	0.15		1st	2nd	2

surface preparation

A - POLISH: 2% bromine in Dimethylformamide (DMF)

B - SPRAY: 1/16% Bromine in Ethylene Glycol spray etch

C - UV - O_3: ozone oxidation ash etch

* transmission cutoff data, will correlate to bandgap values

notes

1 - Excess Te in surface region

2 - Oxidized Te at surface

3 - LPE on CdTe substrate

4 - LPE on CdZnTe substrate

growth types

HZM: Horizontal Zone Melt

LPE: Liquid Phase Epitaxy

SSR: Solid State Recrystallized

VT: Vapor Transport

(THIS PAGE INTENTIONALLY LEFT BLANK)

Section IV
ALLOY BONDING

A. INTRODUCTION

In this period, seven papers appeared in print as a result of this task (some were submitted in earlier periods). In addition, two papers were accepted for publication, and three others have been submitted but not yet refereed. Reprints and preprints of these papers are attached. The topics treated in these papers and in new work not yet written for publication are:

I Isoelectronic Defects (Low Concentration Limit)
    A. Bond energy changes
    B. Lattice Relaxation
    C. Effect on simple model of mixing enthalpies
    D. Strain coefficient modifications

II Concentrated Alloys
    A. Compositional clusters
    B. New theory of order-disorder transitions
    C. New theory of liquid-solid phase transitions

III Band Structure, Transport, and Optical Properties of Alloys

IV Deep States in Semiconductors

V Structural and Mechanical Properties of Semiconductors
    A. Dislocation energies
    B. Hardness
    C. Strategy for dislocation reduction in semiconductors
    D. Vacancy formation energies.

There is a broad range of applications of the ideas contained in these papers. However, our motivation for doing the work under this contract is to appreciate the underlying phenomena that control devices fabricated in narrow bandgap materials, especially near surface related properties of HgCdTe. We are at this point trying to understand the phenomena that cause the wide

(<1 μm) damage (or transition layer) between bulk HgCdTe and any interface: surface to vacuum, metal Shottkey barrier, insulator, or heterojunction. The feature we have uncovered which was not anticipated at the outset, is the presence of clusters. The remaining time under this contract will be spent trying to uncover the connection, if any, between properties of clusters and the unusual surface region in HgCdTe

1. Semiconductor Alloys: Local Bond Lengths, Mixing Enthalpies, and Microclusters

A.B. CHEN* AND A. SHER†
*Physics Department, Auburn University, AL 36849
†SRI International, Menlo Park, CA 94025

ABSTRACT

Several recent theoretical studies of the local structure of semiconductor alloys are summarized. First, dilute limit calculations of local bond lengths and mixing enthalpies are discussed. These calculations include effects due to both bond length and bond-angle distortions, as well as local chemical rearrangements. Then, a new statistical theory of concentrated alloys is described. Deviations from random alloy distributions (microclusters) are predicted.

INTRODUCTION

This paper summarizes our recent theoretical studies directed toward understanding the microscopic structures of pseudo-binary semiconductor alloys $A_xB_{1-x}C$. We first present a detailed calculation of the local bond length relaxation in the dilute limit $x \rightarrow 0$, i.e. the case where an A atom is substituted for a B atom in a BC compound. The mixing enthalpy parameter Ω is found to be related to small excess substitution energies. These excess energies are calculated directly through a minimization procedure. Thus, the accuracy of the predicted Ω is not limited by trying to find small differences between large numbers. The theory is then generalized to concentrated alloys using statistics based on combinations of tetrahedral clusters of five atoms. Our results predict that microclustering occurs in a majority of alloys. We conclude by identifying systematic correlations between the theory and several experiments.

Before discussing the calculation, it is useful to provide some background about the structure of these alloys. It was customary to assume that these alloys have two sublattices in which the C atoms occupy one sublattice, and A and B atoms are randomly distributed on the other. This picture, referred to as the virtual crystal approximation (VCA), implies that the nearest-neighbor (nn) bond lengths in the alloy are the concentration weighed average values, i.e. $d_{AC} = d_{BC} = \bar{d} = x\, d_{AC}^{(0)} + (1-x)\, d_{BC}^{(0)}$ where the values with a superscript (0) denote the pure-crystal values. On the other hand, according to Pauling's covalent radii approximation (CRA), the local bond lengths retain their respective pure-crystal values, i.e. $d_{AC} = d_{AC}^{(0)}$ and $d_{BC} = d_{BC}^{(0)}$

If we define $\delta_0 = (\bar{d} - d_{AC}^{(0)}) / \bar{d}$ and $\delta = (\bar{d} - d_{AC}) / \bar{d}$, then the ratio δ/δ_0 in VCA is zero, but in CRA it is 1. However, Mikkelsen and Boyce(1) found from their EXAFS experiment on $Ga_xIn_{1-x}As$ that the nn bond lengths do not fit either VCA or CRA. Instead, they found the value of δ/δ_0 to be close to 3/4. Since then, similar experiments have been done for a number of zinc-blende pseudo-binary alloys,(1) and the 3/4 rule appears to be quite general.

DILUTE LIMIT

The dilute limit is the easiest case but is still not trivial. Its solution provides both end-point results ($x = 0$ and 1), as well as insight into the extension to the concentrated alloy case. A complete description of this case is being reported elsewhere; here we summarize the essential results. The substitution energy Δ_s for an A atom replacing a B atom in a BC compound is calculated and minimized to find the relaxed configuration. Δ_s can be written as $\Delta_s = 4\,(\Delta E_b + \Delta E_s + \Delta E_{ch})$, where ΔE_b is the binding energy difference between the AC and BC compounds, ΔE_s is the strain energy, and ΔE_{ch} is a chemical energy shift. All $\Delta E'$ s are energies per bond. Then, $\Delta E = \Delta E_s + \Delta E_{ch}$ is the excess energy per bond for the impurity substitution. ΔE_s is calculated by dividing the crystal into two regions. Outside R (which is the distance of the second-shell atoms to the impurity), the distorted crystal is treated as an elastic continuum with a radial displacement field which is inversely proportional to the square of the radial distance, so $\Delta E_s^{(out)} = 1/4RCu^2$, where C is an effective shear coefficient, - -

$$C = \pi\,(1.6\,(C_{11} - C_{12}) + 4.8\,C_{44}), \tag{1}$$

and u is the magnitude of the displacement at R. Inside R, the strain energy $\Delta_s^{(in)}$ is treated with a valence force field (VFF).(2) Finally, the chemical energy shift ΔE_{ch} is calculated from Harrison's model and arises from changes in the metalization energies caused by different bond lengths $\Delta d \equiv d_{BC} - d_{AC}$ and covalent energies $\Delta V_3 \equiv V_3(AC) - V_3(BC)$. Note that $\delta_0 = (d_{BC}^{(0)} - d_{AC}^{(0)}) / d_{BC}^{(0)}$ and $\delta = (d_{BC}^{(0)} - d_{AC}) / d_{BC}^{(0)}$ in this dilute limit, so the excess energy ΔE can be expanded up to second order in δ, u, and ΔV_3 For a given pair A and B, ΔE is an explicit function of δ and u. Minimization of ΔE with respect to δ and U leads to the equilibrium local bond length d_{AC} and energy ΔE. Then, ΔE is used to estimate the mixing enthalpy parameter Ω in the mixing enthalpy $\Delta H = x(1-x)\Omega$ by

$$\Omega = 2\,(\Delta E\,(\text{A in BC}) + \Delta E\,(\text{B in AC})). \tag{2}$$

A systematic comparison with other models based on strain energy alone shows that an increase of the range of the fixed boundary R increases the relaxation of d_{AC}, i.e. it causes δ/δ_0 to increase. The inclusion of the bond angle restoring force, on the other hand, reduces the relaxation. It turns out that a delicate cancellation of these two effects causes a simple spring model pointed out by Shih et al. (SSHS)(3) to yield accurate results. In this model $\delta = \delta_0 / (1 + 1/3\,\alpha/\alpha_I)$, where α and α_I are the bond-stretching force constants for the host (BC) and the impurity (AC) crystals. With $\alpha \approx \alpha_I$, this model predicts $\delta = 3/4\delta_0$ for a zinc-blende alloy. Although our full perturbation theory (FPT) and the VFF model of Martin and Zunger (MZ)(4) predict d_{AC} with an average absolute deviation comparable to the experimental uncertainty of 0.01A, the simple spring model is even better.

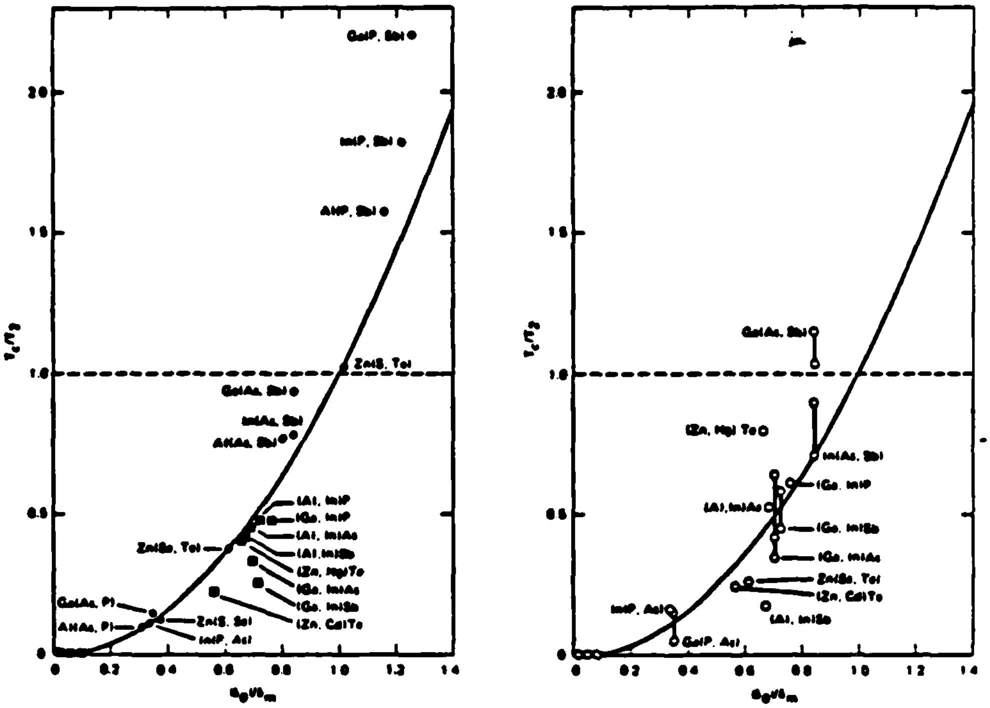
We note that while in our theory, MZ and SSHS, the Ω values are directly calculated without any adjustable parameters, our theory and SSHS agree with the experiment as well or even slightly better than the one-parameter theories.(5,6) Although our theory predicts a negative Ω value for all three (Ga, Al) alloys, the magnitude ($\Omega = -0.17$ kcal/mole) is too small to account for the ordering of $Ga_xAl_{1-x}As$ grown at 600 to 700°C found recently.(7) The calculated Ω values also provide guidance in separating miscible from immiscible alloys. In a random alloy, the criterion for alloy mixing for all x is $T > T_c$, where the critical temperature T is given by $T_c = \Omega/2R_g$ with R_g being the universal gas constant. Figure 1 is a plot of T_c/T_2 against $|\delta_0|/|\delta_m|$, where T_2 is the lower of the two constituent's melting temperatures, and $\delta_m = 1.83\chi_m$ with χ_m being the ratio of rms bond length amplitude fluctuation to the bond length at T_2. The simple spring model gives $T_c/T_2 = (\delta_0/\delta_m)^2$ as indicated by the solid curves.

There is an empirical rule stating that a miscibility gap will occur if $|\delta_0|$ between two alloy components exceeds 7.5%. However, if T_c/T_2 is plotted against $|\delta_0|$, the simple spring model would not exhibit a smooth simple quadratic curve, and our theoretical points would be much more scattered. This suggests that $|\delta_0|/|\delta_m| > 1$ is a better criterion than $|\delta_0| > 0.075$. Figure 2 also clearly shows the chemical effects, namely negative and positive chemical energies ΔE_{ch} for cation and anion substitutions respectively. The full theory and the experiments correlate within the experimental uncertainties. The simple SSHS model clearly is an excellent universal representation. However, T_c/T_2 varies faster than quadratically for larger $|\delta_0/\delta_m|$ values, as born out from both the experimental data and the full theory.

CONCENTRATED ALLOYS

Turn now to the concentrated alloy case. First, an improved statistical model is required. We have extended regular solution theory based on pair energies to one for five-atom clusters. For an $A_xB_{1-x}C$ alloy, the building blocks are clusters of A(m)B(4-m)C, where m ranges from 0 to 4. For a given alloy concentration x and for a given set of energies ϵ_m associated with these clusters, we have derived expressions for the cluster population distribution $x_m \equiv \bar{N}_m / N$, where N is the total number of unit cells and $\bar{N}_m$ is the averaged number of cells with A(m)B(4-m)C clusters. The partition function Z is obtained using a steepest descents argument which then yields the mixing Helmhotz free energy ΔF. The result reduces to Guggenheim's tetrahedron case(7) if pair potentials (for the second-neighbors) are assumed. Another major difference is that we only need to solve a single quartic equation, while Guggenheim needed to solve four simultaneous quartic equations.

The key to the problem, however, lies in the calculation of the energies ϵ_m. If one assumes that the size of the tetrahedra for all m-clusters at a given alloy concentration takes on the corresponding VCA values but allows the central C atom to relax, then the energies as functions of x behave like those shown in Figure 2(a). There are at least two major flaws in this result. First, the energies are too large and would correspond to Ω values many times the experimental values. Second, at $x = 0.75$, 0.5, and 0.2, these energies imply compound formation for $A_3B_1C_4$, $A_2B_2C_4$, and $A_1B_3C_4$, respectively, which is opposite to the known tendency for spinodal decomposition of $Ga_xIn_{1-x}As$ at low T. However, if the local cell volume of each cluster is allowed to be in mechanical equilibrium with a continuous medium with an effective shear coefficient $\bar{C} = xC_{AC}^{(0)} + (1-x)\,C_{AB}^{(0)}$, where the C value for the pure material is given by Eq. (1), then, the corresponding energies ϵ_m as a function of x are given in Figure 2(b), which now yields a reasonable value of mixing enthalpy and correctly predicts the tendency toward spinodal decomposition at low temperature. With this set of energies, one can then calculate the cluster distribution x_m, and compare them with the corresponding values for a random alloy, i.e. $x_m^{(0)} = \binom{4}{m} x^m (1-x)^{4-m}$. Figure 3 shows the deviation from randomness $\Delta x_m = x_m - x_m^{(0)}$ as a function of x for four arbitrarily chosen growth



(a) THEORY

(b) EXPERIMENT

FIGURE 1 PLOT OF T_c/T_2 AS A FUNCTION $|\delta_0|/\delta_m$ FOR T_c OBTAINED FROM (a) PRESENT THEORY AND (b) THE EXPERIMENTAL Ω. IN (a), THE DOTS ARE FOR ANION ALLOYS AND CROSSES FOR CATION SUBSTITUTION. THE SOLID LINES IN BOTH PARTS CORRESPOND TO THE Ω BASED ON THE SIMPLE THEORY OF REF. 3. THE DASHED LINES AT $T_c/T_2 = 0$ SEPARATE THE MISCIBLE FROM IMMISCIBLE GROUPS

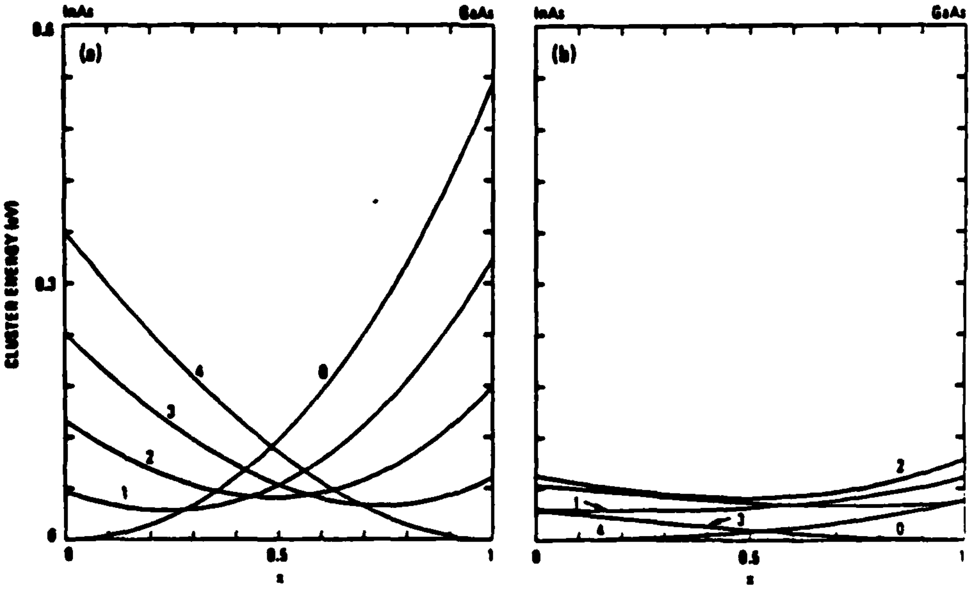


FIGURE 2 THE CLUSTER ENERGIES ϵ_m AS A FUNCTION OF x CALCULATED IN TWO DIFFERENT WAYS, AS DISCUSSED IN THE TEXT

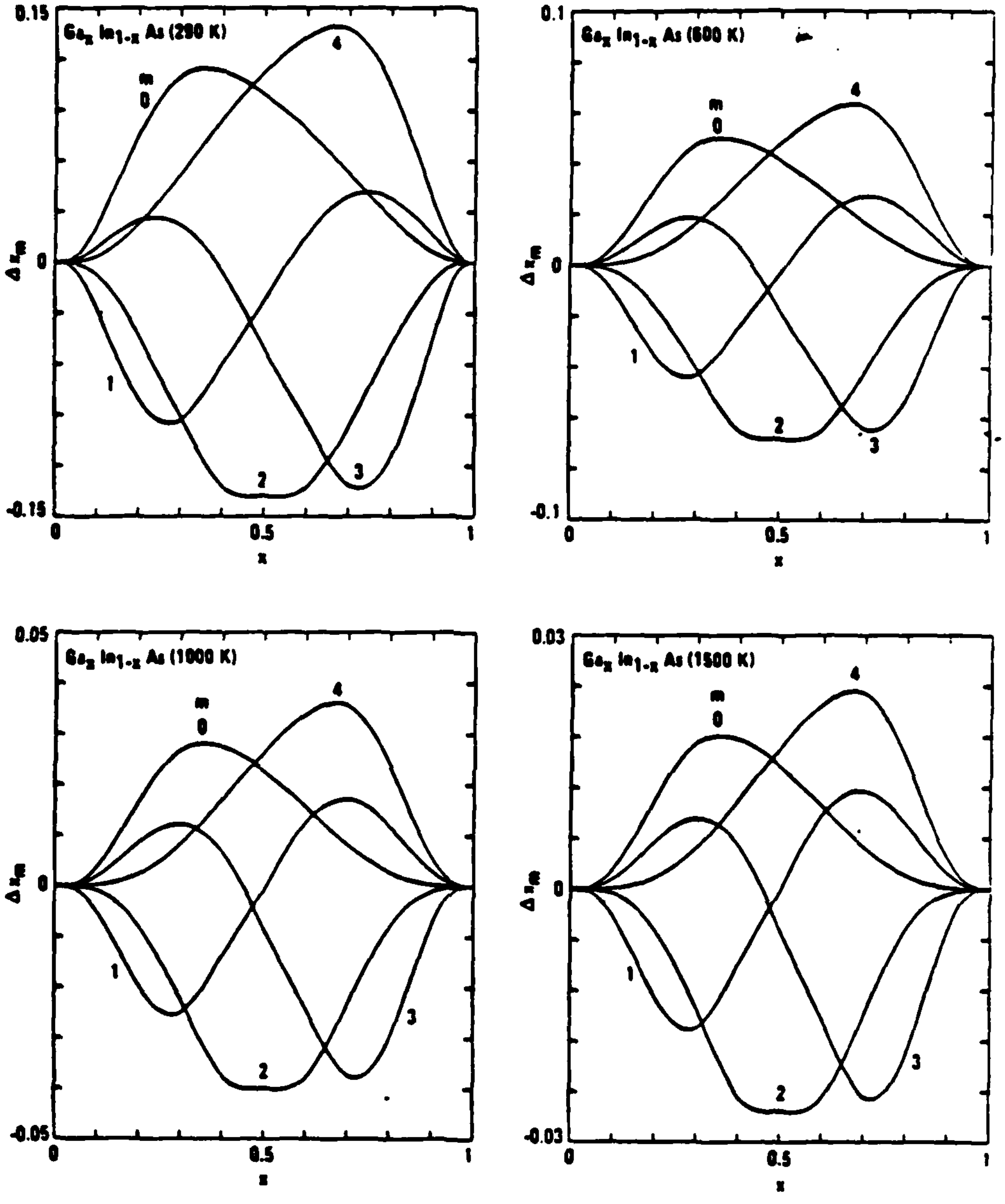


FIGURE 3 DEVIATIONS FROM RANDOM DISTRIBUTIONS FOR THE FIVE CLUSTERS AT SEVERAL GROWTH TEMPERATURES

temperatures. It shows that Δx_m can be several percent at ordinary growth temperatures and there are deviations from the symmetric behavior about $x = 1/2$ predicted by the pair energy model.

It is also interesting to examine the sets of bond lengths $d_{AC}^{(m)}$ and $d_{BC}^{(m)}$ corresponding to the ϵ_m values of Figure 2(b). The spread in the d values among different clusters is found to lie within the width of EXAFS lines.

The existence of microclusters, whose populations deviate from those of a random alloy, will impact on many experimental results. These include phase diagrams, EXAFS line positions and widths, magnetic quadruple splittings, modulation spectroscopy widths, Raman frequencies and intensities for different modes, etc. We have demonstrated that the theory agrees with EXAFS and measured mixing enthalpy parameters. It should be tested against more experiments to further our understanding of the structural properties of semiconductor alloys.

Acknowledgment

The work is supported in part by AFOSR-84-0284 and DARPA Contract MDA 903-83-C-0108. A.-B. Chen would like to thank Professor W.E. Spicer for his hospitality at Stanford University.

REFERENCES

1. J.C. Mikkelsen and J.B. Boyce, Phys. Rev. Lett. *49*, 1412 (1982); Phys. Rev. *28*, 7130 (1983); and private communications.

2. R.M. Martin, Phys. Rev. *B1*, 4005 (1970).

3. C.K. Shih, W.E. Spicer, W.A. Harrison and A.Sher, Phys Rev. B*31*, 1139 (1985).

4. J.L. Martins and A. Junger, Phys. Rev. B *30*, 6217 (1984).

5. G.B. Stringfellow, J. Cryst. Growth *27*, 21 (1974); *58*, 194 (1982).

6. P.A. Fedders and M.W. Muller, J. Phys. Chem. Solids *45*, 685 (1984).

7. T.S. Kuan, T.F. Kuech, W.I. Wang, and E.L. Wilkie, Phys. Rev. Lett. *54*, 201 (1985).

8. E.A. Guggenheim, *Mixtures*, (Oxford at the Clarendon Press, 1952).

a. Semiconductor pseudobinary alloys: Bond-length relaxation and mixing enthalpies

A.-B. Chen*
Department of Physics, Auburn University, Alabama 36849

A. Sher
SRI International, Menlo Park, California 94025

(Received 1 March 1985)

Harrison's bonding theory, the valence force field (VFF), and an elastic continuum are combined in a study of the substitution energies Δ_s and local (first-shell) bond lengths d_1 of isoelectronic impurities in semiconductors. Explicit expressions for Δ_s and d_1 are derived, which enable us to absorb measured elastic constants into the calculation and to study the chemical effects arising from differences in the covalent radii and polarities. Several models based on VFF alone are also derived for comparison. The full theory and at least five VFF models are found to produce impurity bond lengths in excellent agreement with experiment. The substitution energies are shown to provide good estimates of the mixing enthalpies Ω of pseudobinary alloys and to predict miscibility gaps properly. The chemical shifts in Ω are found to be negative for most cation alloys but positive for anion substitutions.

MS code no. BC3005 1985 PACS numbers: 61.55.Hg, 61.70.Rj

I. INTRODUCTION

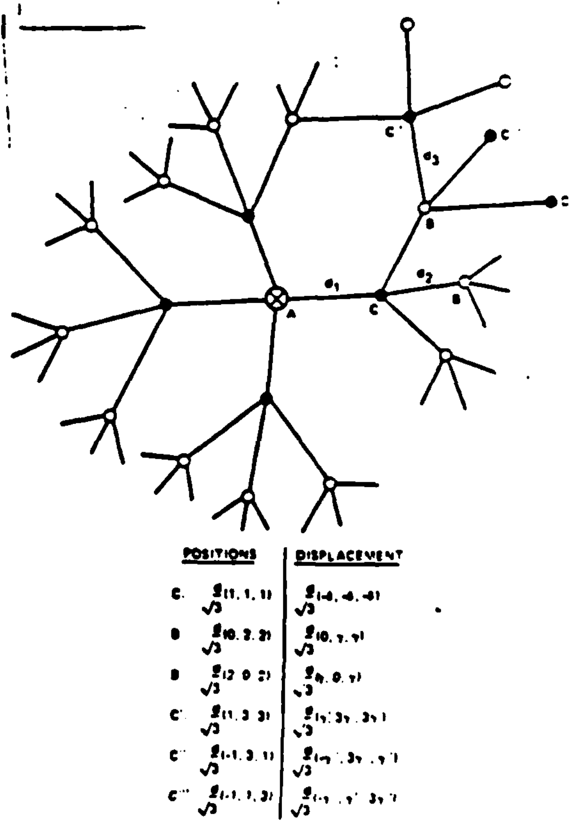
The discovery of a bimodal distribution of the nearest-neighbor bond lengths[1] in $Ga_xIn_{1-x}As$ has sparked considerable interest in the bonding nature of semiconductor alloys.[2–6] This finding has changed the conventional picture of the alloy crystal bond configuration, which has far-reaching implications about the electronic structure, structural stability, and thermodynamics of these materials. Because of the complexity of both the structural and the potential disorder in these alloys, *ab initio* band-structure techniques have not yet evolved to a stage suitable for direct calculations. Therefore, we have extended Harrison's bonding theory[7] to study the alloy structural properties.[5,6] In this paper, we apply an intermediate version of the theory to the dilute-limit case of an isoelectronic impurity.

A particularly useful application of the theory is its perturbation-expansion form, in which measured elastic constants are incorporated to obtain accurate results. This form is also useful for comparison with other previously published models[3,8,9] that are based on the valence-force-field (VFF) (Ref. 10) model alone. Thus, all the factors influencing bond-length relaxation, e.g., strains, boundary conditions, and chemical effects, can be studied. The ability to incorporate the chemical effects is one major difference between this theory and other VFF models.

The remainder of the paper contains the following sections: Sec. II describes a theory for calculating impurity substitution energies. Section III casts the theory into perturbation form and combines it with a valence force field and an elastic continuum. Several VFF models are derived in Sec. IV. The modifications of numerical results due to chemical effects on local bond lengths and alloy mixing enthalpies are summarized and discussed in Sec. V. Conclusions are drawn in the last section, Sec. VI.

II. IMPURITY-SUBSTITUTION ENERGY

Consider the problem of substituting an isoelectronic atom *A* for a *B* atom in a zinc-blende compound *BC* (e.g., In substitutes for Ga in GaAs, as shown in Fig. 1). In



	POSITIONS	DISPLACEMENT
C	$\frac{d}{\sqrt{3}}(1,1,1)$	$\frac{d}{\sqrt{3}}(-\delta,-\delta,-\delta)$
B	$\frac{d}{\sqrt{3}}(0,2,2)$	$\frac{d}{\sqrt{3}}(0,\gamma,\gamma)$
B	$\frac{d}{\sqrt{3}}(2,0,2)$	$\frac{d}{\sqrt{3}}(\gamma,0,\gamma)$
C'	$\frac{d}{\sqrt{3}}(1,3,3)$	$\frac{d}{\sqrt{3}}(\gamma,3\gamma,3\gamma)$
C''	$\frac{d}{\sqrt{3}}(-1,3,1)$	$\frac{d}{\sqrt{3}}(-\gamma,3\gamma,\gamma)$
C'''	$\frac{d}{\sqrt{3}}(-1,1,3)$	$\frac{d}{\sqrt{3}}(-\gamma,\gamma,3\gamma)$

FIG. 1. A sketch of the flattened picture of a single impurity *A* in a *BC* compound. The positions and displacements for those atoms labeled are used in Appendix B.

general, the bond lengths $d_1, d_2, d_3, \ldots$ for the first-, second-, and third-shell bonds surrounding the impurity are different from the equilibrium values of either the pure BC compound, denoted as d, or the "impurity" compound AC, denoted as $d_I = d(1-\delta_0)$. If A starts being a free atom and B also ends being a free atom, then the energy difference between the final and initial states is defined as the substitution energy and is given by

$$\Delta_s = (E_{\rm def} + \varepsilon_B) - (E_{\rm pure} + \varepsilon_A) , \tag{1}$$

where ε_A and ε_B are free-atom energies for A and B, respectively, $E_{\rm def}$ is the total energy of the semiconductor with a defect as sketched in Fig. 1, and $E_{\rm pure}$ is that of the pure host BC crystal. Equation (1) can be written as

$$\begin{aligned}\Delta_s &= (E_{\rm def} - E_{\rm dis}) - (\varepsilon_A - \varepsilon_B) + (E_{\rm dis} - E_{\rm pure}) \\ &= \Delta_R - (\varepsilon_A - \varepsilon_B) + \Delta_{\rm dis} ,\end{aligned} \tag{2}$$

where we have added and subtracted a term $E_{\rm dis}$, which is the total energy of a distorted BC compound with all the atoms held at the positions specified in Fig. 1, except that the central atom is a B atom. Clearly, $\Delta_R = E_{\rm def} - E_{\rm dis}$ is a replacement energy, and the distortion energy $\Delta_{\rm dis} = E_{\rm dis} - E_{\rm pure}$ is the energy required to deform a pure BC crystal from its equilibrium lattice to that specified in Fig. 1. Δ_R contains all the chemical terms that arise from different bond lengths and polarities between AC and BC bonds.

Δ_R can be treated most easily by Harrison's bonding theory.[7] In this theory, the energy per bond relative to the vacuum state is

$$E_b = 2\varepsilon_b + V_0 + 6\varepsilon_m^+ + 6\varepsilon_m^- , \tag{3}$$

where ε_b is the energy of the bonding state constructed from the two hybrid orbitals facing each other along the bond direction

$$E_b = \tfrac{1}{2}(\varepsilon_h^A + \varepsilon_h^C) - (V_2^2 + V_3^2)^{1/2} , \tag{4}$$

with ε_h^A and ε_h^C being the energies of the anion and cation hybrid orbitals, respectively. The antibonding energy ε_a has the same form as in Eq. (4), except with a plus sign. V_2 is called the covalent energy, which is the total electronic Hamiltonian matrix element between the two hybrids in question, and the polar energy V_3 is the difference $V_3 = \frac{1}{2}(\varepsilon_h^C - \varepsilon_h^A)$. The ε_m^+ and ε_m^-, the metallization energies, are the shifts of the bonding level caused by interactions with the neighboring antibonding states, where + and − indicate whether the common adjacent atom is a cation or anion. For example, this term for an AC bond labeled by d_1 in Fig. 1 due to an antibonding state labeled by d_2 is given by

$$\varepsilon_{m,C}^{B,A}(2,1) = \frac{|U_b^A(1)|^2 \, |U_a^B(2)|^2 \, |V_1^C|^2}{\varepsilon_b^A(1) - \varepsilon_a^B(2)} , \tag{5}$$

where A and B denote AC and BC bonds, respectively, $V_1^C = \frac{1}{4}(\varepsilon_s^C - \varepsilon_p^C)$, with ε_s^C and ε_p^C being the s- and p-term values of the common adjacent atom C. $U_b^A(1)$ is the probability amplitude of finding an electron in the hybrid orbital of the C atom in the bonding state of an AC bond with a bond length d_1, whereas $U_a^B(2)$ is the corresponding probability amplitude for the antibonding state of a BC bond of bond length d_2. Finally, V_0 is a repulsive pair potential required to prevent the crystal from collapsing and to guarantee a correct equilibrium bond length. The local perturbation, Eq. (5), is applicable because the square root of the numerator is much smaller than the separation between the antibonding and bonding levels and the valence band is completely filled, so the interaction between the bonding states only spreads the ε_b levels into bands without affecting the center of gravity of the occupied states.

The replacement energy Δ_R of Eq. (2) can now be written explicitly:

$$\begin{aligned}\Delta_R = 4[\,& 2\varepsilon_b^A(1) + V_0^A(1) + 6\varepsilon_{m,A}^{A,A}(1,1) + 6\varepsilon_{m,C}^{B,A}(2,1) \\ & - 2\varepsilon_b^B(1) - V_0^B(1) - 6\varepsilon_{m,B}^{B,B}(1,1) - 6\varepsilon_{m,C}^{B,B}(2,1) \\ & + 6\varepsilon_{mC}^{A,B}(1,2) - 6\varepsilon_{m,C}^{B,B}(1,2)] .\end{aligned} \tag{6}$$

The distortion energy $\Delta_{\rm dis}$ of Eq. (2) now involves only BC bonds of different bond lengths. It can be treated with exactly the same procedure for any given set of bond-length distributions. Thus, a straightforward energy minimization procedure can be carried out. The accuracy of this procedure, however, depends in turn on the accuracy of scaling rules for V_2 and V_0 and the input parameters. At present, Harrison's model[7] with $V_2 \propto 1/d^2$ and $V_0 \propto 1/d^4$ and his universal parameters are only semiquantitative. We are improving the quantitative nature of the theory so that the full theory will yield accurate predictions of the structural and thermal properties of semiconductor defects and alloys.

III. PERTURBATION EXPANSION, VALENCE FORCE FIELD, AND ELASTIC MEDIUM

As pointed out earlier, a perturbation expansion of the theory is instructive. This is feasible because the differences $\Delta V_2 = V_2(AC) - V_2(BC)$ and $\Delta V_3 = V_3(AC) - V_3(BC)$ are small compared to each individual value for many of the isoelectronic impurities in III-IV and II-VI compounds. To this end, Eq. (6) is rewritten as

$$\begin{aligned}\tfrac{1}{4}\Delta_R = E_b^A(1) - E_b^B(1) &+ 6[\varepsilon_{m,C}^{B,A}(2,1) - \varepsilon_{m,C}^{A,A}(1,1)] \\ &- 6[\varepsilon_{m,C}^{A,B}(2,1) - \varepsilon_{m,C}^{B,B}(1,1)] \\ &+ 6[\varepsilon_{m,C}^{A,B}(1,2) - \varepsilon_{m,C}^{B,B}(2,2)] \\ &- 6[\varepsilon_{m,C}^{B,B}(1,2) - \varepsilon_{m,C}^{B,B}(2,2)] ,\end{aligned} \tag{7}$$

where $E_b^A(1)$ and $E_b^B(1)$ are energies per bond in Eq. (3) for AC and BC compounds, respectively, with the relaxed bond length $d_1 = d(1-\delta)$. The difference between these energies E_b and the corresponding values at their respective equilibrium bond lengths d_I and d are just the strain energies per bond in uniform deformation:

$$\begin{aligned}E_b^A(1) &= E_b^A(d_I) + 2\sqrt{3}B_I d_I (d_1 - d_I)^2 , \\ E_b^B(1) &= E_b^B(d) + 2\sqrt{3}B d (d_1 - d)^2 ,\end{aligned} \tag{8}$$

where B_I and B are the bulk moduli for the impurity AC and host BC crystals. The rest of the terms in Eq. (7) are all due to changes in ε_m caused by the differences ΔV_3 and ΔV_2. We shall use Harrison's scaling rules to deduce them.[7] Expanding Eq. (7) to second order in ΔV_3 and $\Delta d = d_2 - d_1$, we write

$$6[\varepsilon_{m,C}^{B,A}(2,1)-\varepsilon_{m,C}^{A,A}(1,1)]$$
$$= f_I\,\Delta d - g_I\Delta V_3 + h_I(\Delta d)^2 - W_I\,\Delta d\,\Delta V_3 + U_I(\Delta V_3)^2\,, \tag{9}$$

where f_I, g_I, and so on, are appropriate derivatives evaluated for the impurity crystal AC. When similar expansions are made for the rest of the terms in Eq. (7), it becomes [with $d_1 = d(1-\delta)$, $d_I = d(1-\delta_0)$]

$$\tfrac{1}{4}\Delta_R = \Delta E_b + (f_I - f)\Delta d - (g_I - g)\Delta V_3 + (h_I + h)(\Delta d)^2$$
$$-(W_I + W)\Delta d\,\Delta V_3 + (U_I + U)(\Delta V_3)^2$$
$$+2\sqrt{3}B_I d^3(\delta-\delta_0)^2 - 2\sqrt{3}Bd^3\delta^2\,, \tag{10}$$

where

$$\Delta E_b = E_b^A(d_I) - E_b^B(d) - \tfrac{1}{4}(\varepsilon_A - \varepsilon_B) \tag{11}$$

is just the difference in the binding energy per bond between the "BC" and "AC" crystals. In Eq. (10) the coefficients f, g without a subscript are those for the host BC system. It is convenient to define an excess energy $\Delta E = \Delta_s/4 - \Delta E_b$, which is the extra energy per bond required for the impurity substitution over and above the binding-energy difference between the BC and AC crystals. The binding-energy difference accounts for much of the substitution energy; however, the correction measured by the excess energy can be significant. The excess energy results from strain energies and chemically driven charge redistributions around the defect. Using Eqs. (2) and (10) and defining $F = f_I - f$ and $G = g_I - g$, we can write ΔE up to second order in ΔV_3 and Δd as

$$\Delta E = 2\sqrt{3}B_I d^3(\delta-\delta_0)^2 - 2\sqrt{3}Bd^3\delta^2 + F\,\Delta d - G\,\Delta V_3$$
$$+ H(\Delta d)^2 + W\,\Delta d\,\Delta V_3 + U(\Delta V_3)^2 + \tfrac{1}{4}\Delta_{\text{dis}}\,, \tag{12}$$

where $H = h_I + h$, $W = w_I + w$, and $U = u_I + u$.

To treat the distortion energy Δ_{dis}, we divide the crystal into two regions. Inside a sphere of some radius R measured from the impurity, the strain energy is taken to be the valence-force-field[10] value:

$$\Delta_{\text{dis}}^{(\text{in})} = \frac{3}{8d^2}\sum_i \alpha_i[\Delta(\mathbf{d}_i\cdot\mathbf{d}_i)]^2$$
$$+\frac{3}{8d^2}\sum_i{}' \sum_{\substack{j\\(j<i)}} \beta_{ij}[\Delta(\mathbf{d}_i\cdot\mathbf{d}_j)]^2\,, \tag{13}$$

where i sums over all the bonds inside R and the pairs in the β terms include those that have adjacent atoms inside R and on the boundary. The parameters α and β are force constants to be considered later. $\Delta(\mathbf{d}\cdot\mathbf{d}_j) = \mathbf{d}_i\mathbf{d}_j - \mathbf{d}_i^{(0)}\cdot\mathbf{d}_j^{(0)}$ measures the change of the dot product between bond vectors due to distortions. Outside R we assume an elastic continuum with radial displacements inversely proportional to the square of the distance from the center. The elastic energy in this medium can be shown to be (see Appendix A)

$$\Delta_{\text{dis}}^{(\text{out})} = RCu^2\,, \tag{14}$$

where the effective shear coefficient is

$$C = 4\pi[0.4(C_{11}-C_{12}) + 1.2C_{44}]$$

and u is the displacement at R. In view of the fact that the bonds d_1 and d_2 are coupled through the chemical terms in Eq. (10), the smallest logical radius R is the second-shell atomic distance, namely $R = 2\sqrt{2}d\sqrt{3}$. Atoms on this boundary have displacements of the forms $\mathbf{u} = d(\gamma,\gamma,0)/\sqrt{3},\ldots$. Thus, $u = \sqrt{2}\gamma d/\sqrt{3}$ and the elastic energy in the continuum is

$$\Delta_{\text{dis}}^{(\text{out})} = \tfrac{1}{3}\sqrt{2/3}\,C\gamma^2 d^3\,. \tag{15}$$

The distortion energy represented by Eq. (14) contains six different contributions (see Appendix B): the bond-stretching energy of the four first-shell bonds $6\alpha\delta^2 d^2$, the β terms from the first-shell bonds, $\beta\delta^2 d^2$, the α terms from the second-shell bonds, $2\alpha(\delta+2\gamma)^2 d^2$, the β terms between the first- and second-shell bonds, $2\beta(\delta+\gamma)^2 d^2$, the β terms among the second-shell bonds, $2\beta\delta^2 d^2$, and finally the β terms between the second-shell bonds and those in the continuum,

$$\tfrac{1}{2}\beta d^2[(3\delta+\lambda_1\gamma)+(\delta+\lambda_2\gamma)^2]\,,$$

where $\lambda_1 = 40\sqrt{2}/(19\sqrt{19})$ and $\lambda_2 = 2\text{-}40\sqrt{2}/(11\sqrt{11})$.

To assemble all the contributions to Eq. (12), we need to consider the assignments of the elastic constants and the force constants α and β in VFF. While the experimental values[11] of C_{11}, C_{12}, and C_{44} can be used for the elastic constants, α and β have to be deduced. If Martin's original procedure[10] (also followed by Martins and Zunger[3]) is used, then Eq. (13) alone will not produce the correct (experimental) bulk moduli. There are small corrections due to Madelung terms, which are hard to treat in the case of nonuniform distortions. A simpler procedure is adopted here. We use the experimental bulk moduli for B_I and B in Eq. (12) and experimental elastic constants to calculate C of Eq. (14) and then force α and β in the VFF to produce the correct bulk moduli B and shear coefficients C_{11}-C_{12}. Such an approach is also consistent with Harrison's bonding theory[7] and other approaches in which the Coulomb forces are automatically incorporated in the band and bond energies, and do not need to be redundantly treated. With our procedure, the bulk modulus is simply related to the force constants by $B = (3\alpha+\beta)/(4\sqrt{3}d)$. Table I lists our α and β values. We want to point out in advance, however, that the numerical results deduced from our sets and those of Martin of α and β do not introduce differences more than the present experimental uncertainties in the local bond length (~0.01 A) and the mixing enthalpies (> 0.5 kcal/mole).

Using the above procedure and adding all contributions, the excess energy per bond from Eqs. (10) and (12) is the full perturbation theory (FPT) result

TABLE I. The bond lengths d (in Å), valence force constants α and β (N/m), shear coefficients C of the continuum (in 10^{11} ergs/cm^3), melting temperatures T_m (K), and Liedermann ratios χ_m for the compounds used in this paper.

Compound	d	α	β	C	T_m[a]	χ_m
AlP	2.367	44.323	8.068	122.396	1773	0.070
AlAs	2.451	40.849	8.717	112.695	1873	0.073
AlSb	2.656	34.073	6.900	85.351	1323	0.062
GaP	2.360	44.764	10.737	145.921	1510	0.064
GaAs	2.448	39.235	9.159	121.844	1738	0.071
GaSb	2.640	31.876	7.347	89.372	985	0.055
InP	2.541	40.363	6.543	91.785	1343	0.059
InAs	2.622	33.203	5.752	78.816	1215	0.061
InSb	2.805	28.557	4.891	60.721	798	0.049
ZnS	2.342	40.429	5.273	89.272	2123	0.081
ZnSe	2.454	32.200	4.562	82.687	1788	0.080
ZnTe	2.637	29.445	4.659	62.430	1511	0.077
CdTe	2.806	26.569	2.722	38.453	1371	0.067
HgTe	2.798	26.396	2.746	40.363	943	0.056

[a] Reference 30.

$$\Delta E=[3\alpha_l(\delta-\delta_0)^2/2+\beta_l(\delta-\delta_0)^2/2+\alpha(\delta+2\gamma)^2/2$$
$$+\beta(\delta+\gamma)^2/2+\beta\delta^2/4+\beta(3\delta+\lambda_1\gamma)^2/8$$
$$+\beta(\delta+\lambda_2\gamma)^2/4]d^2+\sqrt{2}C\gamma^2d^3/(3\sqrt{3})+\Delta E_{ch}\,, \tag{16}$$

where the chemical contribution is written as

$$\Delta E_{ch}=F_{ch}\,\Delta d+H(\Delta d)^2+\Delta E_p$$
$$=\tfrac{2}{3}F_{ch}(2\delta+\gamma)d+\tfrac{4}{9}H(2\delta+\gamma)^2d^2$$
$$+[U(\Delta V_3)^2-G\,\Delta V_3]\,, \tag{17}$$

where $\Delta F_{ch}=F-W\,\Delta V_3$ and $\Delta E_p=U(\Delta V_3)^2-G\,\Delta V_3$. ΔF_{ch} is a chemical force, which when it is positive tends to push the C atom away from the impurity atom A. This force arises from the difference in the bond tensions induced between the AC and BC bonds adjacent to C because the neighboring antibinding states are different from those of their respective host states. ΔE_p is due to the difference in the polarities ΔV_3 alone and is independent of the displacement. Finally, H can be regarded as a chemically induced force constant, which when it is positive tends to restrain the lattice from distortion and increases the elastic energy.

The equilibrium requirements $\partial(\Delta E)/\partial\delta=0$ and $\partial(\Delta E)/\partial\gamma=0$ then lead to the solution $\gamma=Q\delta$, and δ is given by

$$\delta=(\delta_0+\delta_0')/\{1+[\alpha(1-2Q)+\beta(17/4-\lambda Q)$$
$$+16H(1-2Q)/9]/(3\alpha_l+\beta_l)\}\,, \tag{18}$$

where the constant λ is $1+3\lambda_1/4+\lambda_2/2$, and

$$\delta_0'=-4F_{ch}(1-Q/2)/[3d(3\alpha_l+\beta_l)]\,, \tag{19}$$

with $Q=2J/K$, $J=\alpha+\lambda\beta/2+8H/9$, and $K=4\alpha+2\sqrt{2}Cd/(3\sqrt{3})+(1+\lambda_1^2/4+\lambda_2^2/2)\beta+8H/9$.

IV. VALENCE-FORCE-FIELD MODELS

In this section we consider several models based on the valence force field. These models have been used frequently to explain the impurity bond relaxation.[2,4,8,9] We shall first derive the explicit expressions for these models and then connect them with the existing results.

A. Model A: Third-shell atoms and beyond are fixed at their pure crystal positions

Let the bond lengths surrounding the impurity again be $d_1=d(1-\delta)$ and let the second-shell atoms have radial displacements of the forms $(d/\sqrt{3})$, $(\gamma,\gamma,0)$, etc. Beyond and including the third shell, all the other atoms are held at their pure-crystal positions. There are nine different contributions to the strain energy in VFF (see Appendix B): the α terms from the four bonds surrounding the impurity, $6\alpha_l(\delta-\delta_0)^2d^2$; the β terms among the six pairs of these bonds, $\beta_l(\delta-\delta_0)^2d^2$; the β terms between the first- and second-shell bonds, $2\beta(\delta+\gamma)^2d^2$; the α terms from the second-shell bonds, $2\alpha(\delta+2\gamma)^2d^2$; the β terms among the second-shell bonds, $2\beta\delta^2d^2$; the β terms between the second- and third-shell bonds, $\frac{4}{9}\beta\delta^2d^2+\beta(\delta+2\gamma)^2d^2$; the α terms from the third-shell bonds, $8\alpha\gamma^2d^2$; the β terms among the third-shell bonds, $4\beta\delta^2d^2$; and the β terms between the third- and fourth-shell bonds, $6\beta\gamma^2d^2$. Thus, the excess energy (in this case $\frac{1}{4}$ times the strain energy) becomes

$$\Delta E=[\tfrac{3}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{4}\beta_l(\delta-\delta_0)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2$$
$$+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{15}{8}\beta\delta^2+\tfrac{1}{4}\beta(\delta+2\gamma)^2$$
$$+2\alpha\delta^2+\tfrac{3}{2}\beta\delta^2]d^2 \,. \tag{20}$$

The minimization of ΔE with respect to δ and γ leads Eq. (20) to $\gamma=-\delta/4$, and

$$\delta=\delta_0/[1+(\alpha+17\beta/4)/(6\alpha_l+\beta_l)] \,. \tag{21}$$

We note that there is some ambiguity in the third contribution listed above for the β terms between the impurity and host bonds. The value of β could be chosen as one of these combinations β, β_l, $\frac{1}{2}(\beta+\beta_l)$, $\sqrt{\beta\beta_l}$, or other proper combinations. Because the values of β and β_l are comparable and β values are much smaller than α (see Table I), the results for δ and ΔE are not too sensitive to the choice. There is also some ambiguity in the values for $\mathbf{d}_i^{(0)}\cdot\mathbf{d}_j^{(0)}$ for the "undistorted" crystal. The $-d^2/3$ used is the simplest choice. A different choice will not affect the results for δ at all, but will make ΔE slightly different. In fact, model A was first used by Martins and Zunger.[3] However, their expression for δ is different from Eq. (21) because they made different choices of the two quantities just mentioned. Nevertheless, Sec. IV will show that these two expressions yield very similar results. These ambiguities do not occur in the full theory in Sec. III, where the impurity-host interactions are taken into account naturally by the replacement energy Δ_R [see Eq. (10)].

B. Model B: Second-shell atoms connect to a fixed boundary

This model corresponds to $\gamma=0$ in model A. So we have

$$\Delta E=[\tfrac{3}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{4}\beta_l(\delta-\delta_0)^2+\tfrac{1}{2}\alpha\delta^2+\tfrac{19}{8}\beta\delta^2]d^2 \tag{22}$$

and

$$\delta=\delta_0/[1+(\alpha+19\beta/4)/(3\alpha_l+\tfrac{1}{2}\beta_l)] \,. \tag{23}$$

This expression will be used to study the effect of truncation.

C. Model C: Simple spring model

If all the β's in Eqs. (22) and (23) are set equal to zero, we have the simple spring model with

$$\Delta E=[\tfrac{3}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{2}\alpha\delta^2]d^2 \tag{24}$$

and

$$\delta=\delta_0/(1+\tfrac{1}{3}\alpha/\alpha_l) \,. \tag{25}$$

The spring model recently discussed by Shih *et al.*[8] corresponds to Eq. (25) with $\alpha=\alpha_l$, so $\delta/\delta_0=\frac{3}{4}$.

D. Model D: VFF with the continuum connected to the second-shell atoms

1. Model D1

In this case, ΔE only contains the first five contributions listed for case A plus the elastic energy in the continuum. However, the β terms between the second- and third-shell bonds are modified because atoms outside R in the continuum now have radial displacements proportional to the inverse of the square of the radius. The result is

$$\Delta E=\left[\tfrac{3}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{4}\beta_l(\delta-\delta_0)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2+\tfrac{1}{2}\beta\delta^2+\tfrac{1}{8}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta+\lambda_2\gamma)^2+\frac{\sqrt{2}}{3\sqrt{3}}Cd\gamma^2\right]d^2 \,, \tag{26}$$

where λ_1 and λ_2 are the same as the constants that appear in Eq. (16). The corresponding equilibrium condition can be shown to be

$$\delta=\delta_0/\{1+[\alpha(1-2Q)+19\beta/4-\beta(1+3\lambda_1/4+\lambda_2/4)Q]/(3\alpha_l+\beta_l/2)\} \,, \tag{27}$$

where

$$Q=\frac{(2\alpha+\beta+3\lambda_1\beta/4+\lambda_2\beta/2)}{(4\alpha+2\sqrt{2}Cd/3\sqrt{3}+\beta+\frac{1}{4}\lambda_1^2\beta+\frac{1}{2}\lambda_2^2\beta)} \,. \tag{28}$$

2. Model D2

A comparison between Eq. (26) and the full perturbation theory, Eq. (16), shows two major differences. First, all the chemical terms are absent in Eq. (26). Secondly, the terms

$$\tfrac{1}{2}\beta_l(\delta-\delta_0)^2d^2+\tfrac{1}{4}\beta\delta^2d^2$$

in Eq. (16) become

$$\tfrac{1}{4}\beta_l(\delta-\delta_0)^2d^2+\tfrac{1}{2}\beta\delta^2d^2$$

in Eq. (26). This difference in the strain energy will mask the true effects of chemical forces if δ from Eq. (26) is compared with FPT. A better way to study the chemical effects is to use the following equation:

$$\Delta E=[\tfrac{1}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{2}\beta_l(\delta-\delta_0)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2+\tfrac{1}{4}\beta\delta^2+\tfrac{1}{4}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta-\lambda_2\gamma)^2]d^2-\frac{\sqrt{2}}{3\sqrt{3}}C\gamma^3d^3\,, \tag{29}$$

which is Eq. (16) with all the chemical terms neglected. The corresponding δ becomes

$$\delta=\delta_0/\{1+[\alpha(1-2Q)+\tfrac{17}{4}\beta-\beta(1+\tfrac{3}{4}\lambda_1+\tfrac{1}{4}\lambda_2)Q]/(3\alpha_l+\beta_l)\}\,, \tag{30}$$

with Q still given by Eq. (28).

E. Model E: Continuum connected to the first-shell atoms

In this case, $\gamma=-3\sqrt{3}\delta/(8\sqrt{2})$ and ΔE only includes the first three contributions listed in model A plus the strain energy of the continuum:

$$\Delta E=\left[\tfrac{1}{2}\alpha_l(\delta-\delta_0)^2+\tfrac{1}{4}\beta_l(\delta-\delta_0)^2+\tfrac{1}{2}\beta\left[1-\frac{3\sqrt{3}}{8\sqrt{2}}\right]^2\gamma^2+\tfrac{1}{4}Cd\delta^2\right]d^2\,. \tag{31}$$

The relaxation parameter is given by

$$\delta=\delta_0\Big/\left\{1+\left[\tfrac{1}{2}Cd+\left[1-\frac{3\sqrt{3}}{8\sqrt{2}}\right]^2\beta\right]\Big/(3\alpha_l+\tfrac{1}{2}\beta_l)\right\}\,. \tag{32}$$

We note that the continuum model used to estimate the bond-length relaxation by Baldereschi and Hopfield[9] corresponds to Eq. (32) without the β terms, which yields $\delta/\delta_0\approx 0.4$ to 0.5, rather than the proper values around 0.7 to 0.8.

V. ALLOY MIXING ENTHALPY

The impurity-substitution excess energies ΔE provide a first estimate of the mixing enthalpies of pseudobinary alloys. Most current thermodynamics theories of semiconductor alloys are based on an extension of the binary solution model.[12] In this model, the mixing Helmholtz energy of an $A_xB_{1-x}C$ alloy is defined as

$$\Delta F_m=F_{\text{alloy}}-(xF_{AC}+yF_{BC})\,, \tag{33}$$

where $y=1-x$, and F_{AC} and F_{BC} are the respective free energies of the pure AC and BC compounds at the same temperature. Because the C atoms occupy a sublattice, the nearest neighbors of A and B atoms in the alloy are the C atoms. Thus, the pair potentials that enter the binary solution theory are now the second-neighbor interactions. Let N_{AA}, N_{AB}, and N_{BB} be, respectively, the numbers of the second-neighbor AA, AB, and BB pairs, with corresponding pair interaction energies ϵ_{AA}, ϵ_{AB}, and ϵ_{BB}. For tetrahedral semiconductors, there are a total of $6N$ second-neighbor pairs for a crystal containing N unit cells. Denote the ratios N_{AA}, N_{AB}, and N_{BB} to $6N$ as r_{AA}, $r_{AB}=r$, and r_{BB}, respectively. Then those ratios are related to the alloy composition by $r_{AA}=x-r/2$ and $r_{BB}=y-r/2$. The mixing free energy has two terms,

$$\Delta F_m=\Delta E_m-T\,\Delta S\,, \tag{34}$$

where the mixing energy is given by

$$\begin{aligned}\Delta E_m&=E_{\text{alloy}}-(xE_{AC}+yE_{BC})\\&=6N(\varepsilon_{AA}r_{AA}+\varepsilon_{AB}r_{AB}+\varepsilon_{BB}r_{BB})-6N(x\varepsilon_{AA}+y\varepsilon_{BB})\\&=6Nr\,\Delta\varepsilon\,,\end{aligned} \tag{35}$$

where

$$\Delta\varepsilon=\varepsilon_{AB}-\tfrac{1}{2}(\varepsilon_{AA}+\varepsilon_{BB})\,. \tag{36}$$

The mixing entropy ΔS can also be written from a simple generalization of the random distribution.[12] For modest pressure, ΔE is the same as the mixing enthalpy ΔH_m.

Now the pair interaction energies can be approximately related to the impurity-substitution energies by

$$\Delta_s(A\text{ in }BC)\approx 12(\varepsilon_{AB}-\varepsilon_{BB}) \tag{37}$$

and

$$\Delta_s(A\text{ in }BC)\approx 12(\varepsilon_{AB}-\varepsilon_{AA})\,. \tag{38}$$

Thus, $\Delta\varepsilon$ of Eq. (36) becomes

$$\begin{aligned}\Delta\varepsilon&=\tfrac{1}{24}[\Delta_s(A\text{ in }BC)+\Delta_s(B\text{ in }AC)]\\&=\tfrac{1}{6}[\Delta E(A\text{ in }BC)+\Delta E(B\text{ in }AC)]\,.\end{aligned} \tag{39}$$

Usually, the experimental ΔH_m is written as

$$\Delta H_m=x(1-x)\Omega\,, \tag{40}$$

which is equivalent to assuming a random distribution, i.e., $r=2x(1-x)$. Using this expression for r and comparing Eqs. (40) and (35), we see that the mixing enthalpy parameter Ω is given by

$$\Omega=2[\Delta E(A\text{ in }BC)+\Delta E(B\text{ in }AC)]\,. \tag{41}$$

This connection provides a further check of the theory.

VI. NUMERICAL RESULTS AND DISCUSSION

A. Chemical terms

Table II lists $\delta_0=1-d_l/d$, $\delta=1-d_1/d$, the excess energy (per bond) ΔE for the full theory and its corresponding VFF model D2, and the terms derived from the metallization energies, δ_0' [Eq. (19)], F_{ch}, H, ΔE_r, and ΔE_{ch} [Eq. (17)]. The appropriate derivatives f, g, h, ... [see Eq. (9)] are computed using the atomic-term values that we have generated from impurity-level[13] and structural studies.[5]

For substitutions involving the cation pair (Ga,Al), F_{ch} has the same sign as δ_0, which means that F_{ch} prevents relaxation and thus tends to increase the strain energy. The chemical forces H are also significant. As a result,

TABLE II. Comparison between the full theory and the corresponding VFF Model D2 to study the effects of chemical terms. All ΔE's are in units of kcal/mole band.

		Model D2			Full theory						
Impurity	Host	δ_0	δ	ΔE	δ'_0	δ	F_{ch} (10^{-10}N)	H (N/m)	ΔE_p	ΔE_{ch}	ΔE
Ga	AlP	0.003	0.002	0.001	−0.001	0.001	0.150	4.581	−0.016	−0.016	−0.013
Al	GaP	−0.003	−0.002	0.001	0.001	−0.001	−0.150	4.581	−0.016	−0.016	−0.013
Ga	AlAs	0.001	0.001	0.000	−0.002	−0.001	0.243	5.733	−0.020	−0.021	−0.018
Al	GaAs	−0.001	−0.001	0.000	0.002	0.001	−0.243	5.733	−0.020	−0.021	−0.018
Ga	AlSb	0.006	0.004	0.005	−0.004	0.001	0.389	5.632	−0.054	−0.053	−0.039
Al	GaSb	−0.006	−0.004	0.005	0.004	−0.001	−0.389	5.632	−0.054	−0.053	−0.039
In	GaP	−0.077	−0.052	0.959	−0.006	−0.054	0.699	3.778	−0.188	−0.219	0.742
Ga	InP	0.071	0.056	0.734	0.005	0.057	−0.699	3.778	−0.188	−0.206	0.530
In	GaAs	−0.071	−0.048	0.752	−0.009	−0.050	0.804	4.778	−0.257	−0.283	0.472
Ga	InAs	0.066	0.052	0.592	0.007	0.054	−0.804	4.778	−0.257	−0.265	0.330
In	GaSb	−0.062	−0.043	0.554	−0.004	−0.042	0.352	5.201	−0.363	−0.308	0.247
Ga	InSb	0.059	0.046	0.445	0.004	0.044	−0.352	5.201	−0.363	−0.287	0.160
In	AlP	−0.074	−0.053	0.761	−0.007	−0.056	0.769	3.506	−0.035	−0.087	0.679
Al	InP	0.068	0.053	0.674	0.006	0.056	−0.769	3.506	−0.035	−0.083	0.596
In	AlAs	−0.070	−0.048	0.705	−0.010	−0.051	0.942	4.437	−0.948	−0.111	0.602
Al	InAs	0.065	0.052	0.576	0.008	0.054	−0.942	4.437	−0.048	−0.099	0.485
In	AlSb	−0.056	−0.039	0.440	−0.008	−0.041	0.689	4.979	−0.061	−0.073	0.369
Al	InSb	0.053	0.042	0.368	0.007	0.044	−0.689	4.979	−0.061	−0.061	0.310
Cd	ZnTe	−0.064	−0.048	0.432	−0.003	−0.050	0.202	−0.484	−0.005	−0.064	0.373
Zn	CdTe	0.060	0.050	0.314	0.002	−0.053	−0.202	−0.484	−0.005	−0.072	0.247
Hg	CdTe	0.003	0.002	0.001	0.004	9.005	−0.278	−0.753	−0.018	−0.026	−0.018
Cd	HgTe	−0.003	−0.002	0.001	−0.004	−0.005	0.278	−0.753	−0.018	−0.026	−0.018
Hg	ZnTe	−0.061	−0.045	0.392	−0.001	0.046	0.075	0.002	0.052	0.037	0.429
Zn	HgTe	0.058	0.048	0.286	0.001	0.049	−0.075	0.002	0.052	0.035	0.322
As	AlP	−0.035	−0.026	0.179	0.001	−0.025	−0.085	0.717	−0.005	0.008	0.187
P	AlAs	0.034	0.025	0.185	−0.001	0.025	0.085	0.717	−0.005	0.008	0.194
As	GaP	−0.037	−0.025	0.226	0.002	−0.024	−0.181	1.078	−0.011	0.012	0.240
P	GaAs	0.036	0.027	0.211	−0.001	0.025	0.181	1.078	−0.011	0.014	0.228
As	InP	−0.032	−0.023	0.136	0.001	−0.022	−0.057	0.919	−0.003	0.008	0.144
P	InAs	0.031	0.024	0.128	−0.001	0.024	0.057	0.919	−0.003	0.009	0.138
Sb	AlAs	−0.084	−0.058	1.024	0.008	−0.051	−0.815	0.644	−0.180	0.002	1.060
As	AlSb	0.077	0.059	0.919	−0.007	0.053	0.815	0.644	−0.180	0.027	0.984
Sb	GaAs	−0.078	−0.052	0.908	0.018	−0.040	−1.599	0.927	−0.363	−0.106	0.929
As	GaSb	0.073	0.055	0.823	−0.014	0.044	1.599	0.927	−0.363	−0.061	0.904
Sb	InAs	−0.070	−0.051	0.603	0.010	−0.042	−0.824	0.855	0.171	−0.009	0.645
As	InSs	0.065	0.051	0.551	−0.009	0.044	0.824	0.855	−0.171	0.008	0.613
Sb	AlP	−0.122	−0.085	2.007	0.010	−0.077	−9.944	0.645	0.241	0.074	2.127
P	AlSb	0.109	0.085	1.855	−0.008	0.078	0.944	0.645	−0.241	0.123	2.030
Sb	GaP	−0.119	−0.075	2.132	0.021	−0.061	−1.868	0.930	−0.505	−0.046	2.244
P	GaSb	0.106	0.083	1.806	−0.015	0.070	1.868	0.930	−0.505	0.093	2.084
Sb	InP	−0.104	−0.072	1.383	0.011	−0.063	−0.922	0.854	−0.214	0.059	1.501
P	InSb	0.094	0.077	1.191	−0.008	0.069	0.922	0.854	−0.214	0.123	1.379
Se	ZnS	−0.048	−0.036	0.231	−0.001	−0.036	0.077	0.645	0.003	0.000	0.231
S	ZnSe	0.046	0.037	0.221	0.001	0.037	−0.077	0.645	0.003	0.001	0.222
Te	ZnSe	−0.075	−0.056	0.550	0.000	−0.056	0.028	0.635	0.008	0.024	0.574
Se	ZnTe	0.069	0.054	0.532	0.000	0.054	−0.028	0.635	0.008	0.025	0.557
Te	ZnS	−0.126	−0.092	1.565	−0.001	−0.092	0.101	0.644	0.022	0.041	1.606
S	ZnTe	0.112	0.091	1.446	0.001	0.091	−0.101	0.644	0.022	0.051	1.496

all six cases involving this pair have nearly equal d_1 and d_2, i.e., the small bond-length differences are made even smaller in the alloy. The excess energies all become negative, mainly because ΔE_p is negative. For the systems involving the (Ga,In) and (In,Al) pairs, F_{ch} has the opposite sign from δ_0, so δ_0' and δ_0 have the same sign. The chemical force favors bond distortion. However, because H is positive and introduces an increase in the denominator of Eq. (18), most of the effect of δ_0' is cancelled. For cases involving (Ga,In), the polarity contributions ΔE_p are all negative. The $F_{ch}\,\Delta d$ term is negative, but $H(\Delta d)^2$ is positive, so they cancel to a certain degree and leave ΔE lowered primarily because of ΔE_p. While ΔE_p is still negative for the (In,Al) substitutions, its magnitude is reduced considerably. The other chemical energies $F_{ch}\,\Delta d + H(\Delta d)^2$ can be as large as ΔE_p, but the overall reductions of ΔE are only a fraction of those for the (Ga,In) cases. For the several II-VI systems studied, both F_{ch} and H are small and the net changes in δ have the same sign as δ_0. However, because δ_0 is small in the (Cd,Hg) substitutions, F_{ch} actually causes a reversal in which the short bond length becomes shorter and the longer one becomes longer. This is the only exceptional case of this type found for all the systems studied. The change of ΔE due to chemical terms in the (Hg,Zn) substitution is also peculiar—it increases mainly because ΔE_p is positive.

Next, we examine the anion substitutions. For the groups involving the (P,As) pair, the chemical shifts are all small, but the trend is less toward relaxation and larger ΔE. The groups involving (As,Sb) and (P,Sb) pairs behave very similarly: F_{ch} are significant and are opposing relaxations, i.e., δ_0' and δ_0 have opposite signs. At the same time, the H values are several times smaller than those for the corresponding III-V cation substitution case. Thus, most of δ_0' translates into a real reduction of the ratio δ/δ_0 and consequently introduces extra strain energy. Although the ΔE_p energies are significant and negative, $F_{ch}\,\Delta d$ are positive and the net ΔE_{ch} can be either positive or negative. However, the induced-strain energy due to reduction of the δ/δ_0 makes all ΔE positive for these two groups of systems. For II-VI systems, all the chemical effects again are small, but the net chemical changes on ΔE are slightly positive.

The above discussion can be summarized in Fig. 2, where the excess energies ΔE calculated from the full perturbation theory and Model D2 are plotted normalized to the results of the simple spring model of Shih, Spicer, Harrison, and Sher (SSHS) (Ref. 8), i.e., Eq. (24) with $\alpha_I = \alpha$, so $\Delta E / \frac{1}{4}\alpha d^2 \simeq \delta_0^2$. The calculated ΔE rises faster for $\delta_0 \geq 0$ than for $\delta_0 \leq 0$, mainly because $\alpha/\alpha_I \simeq 1$. In fact, if the relation[7,10] $\alpha/\alpha_I = (d_I/d)^S$ with S of order of 3 to 5 is used in Eq. (24), we obtain a percentage correction of $S\delta_0/4$ to the SSHS results, which explains the skewed behavior of the curve. It is also clear that ΔE rises faster than δ_0^2 for larger δ_0. However, the zeroth-order theory of SSHS is clearly an excellent representation of the global features of ΔE. The results from model D2 are closer to the parabolic form than those from FPT. The figure clearly shows the general trends; the chemical terms cause negative shifts in ΔE for cation substitutions and positive shifts for anion impurities. It is also clear that the chemical shifts can be very large. These effects will have important consequences on the alloy mixing enthalpies to be discussed later.

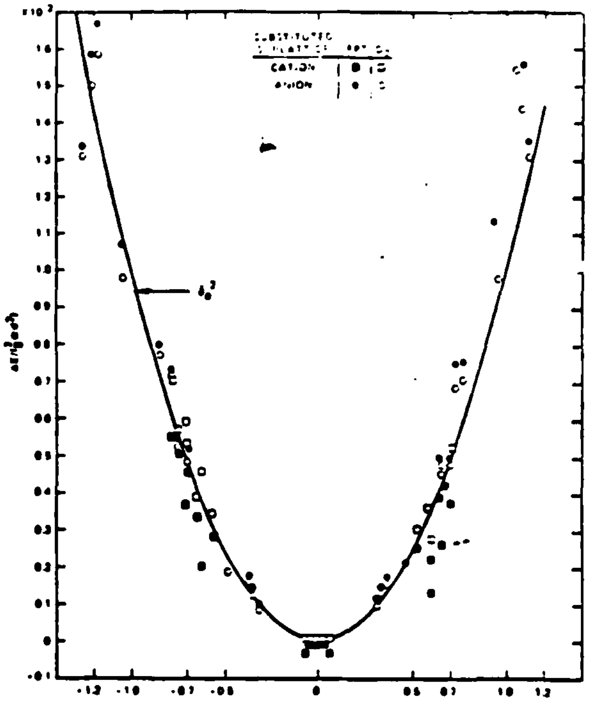


FIG. 2. The excess energies ΔE over $\frac{1}{4}\alpha d^2$ calculated from the full perturbation theory (FPT) and its corresponding valence-force-field model D2. The solid curve corresponds to Eq. (24) with $\alpha = \alpha_I$.

B. Impurity bond length

Table III lists the impurity bond lengths d_1 calculated from different models, while a comparison of theory and the available experimental data[1,14,15] is presented in Table IV. The actual size of changes in d_1 induced by the chemical terms can be seen by comparing model D2 with the full theory. Except for the systems involving the substitution of (Ga,As) and (P,Sb) pairs (where changes range from 0.01 to 0.03 Å), all the chemically induced changes are less than 0.01 Å. Comparison among models A, B, and C shows that, while extending the boundary helps the relaxation (compare model B to A), i.e., δ/δ_0 is closer to 1, the inclusion of the bond-bending forces (the β terms) (compare model B with C) prevents it. The simple spring model (model C), which contains neither of these terms, evidently represents a delicate cancellation of these effects and predicts results close to those of the full perturbation theory and experiment. Although the d_1 values of Model C are often very close to those of model A, there are cases [e.g., Ga(P,Sb)] in which Model C can differ from model

TABLE III. Calculated impurity local bond lengths (in Å) from the full theory and several valence force models discussed in Sec. III, and their comparison with the values calculated by Martins and Zunger (Ref. 3).

Impurity	Host	A	B	C	D1	D2	E	FPT	MZ
Ga	AlP	2.362	2.363	2.362	2.362	2.362	2.364	2.364	
Al	GaP	2.365	2.364	2.365	2.365	2.365	2.365	2.363	
Ga	AlAs	2.449	2.449	2.449	2.449	2.449	2.450	2.452	
Al	GaAs	2.450	2.450	2.450	2.450	2.450	2.450	2.447	
Ga	AlSb	2.645	2.646	2.644	2.645	2.644	2.649	2.653	
Al	GaSb	2.651	2.650	2.652	2.652	2.652	2.650	2.644	
In	GaP	2.477	2.462	2.492	2.479	2.483	2.435	2.487	2.474
Ga	InP	2.406	2.421	2.402	2.395	2.399	2.409	2.396	2.409
In	GaAs	2.559	2.544	2.573	2.561	2.565	2.518	2.570	2.556
Ga	InAs	2.492	2.506	2.486	2.482	2.485	2.496	2.481	2.495
In	GaSb	2.747	2.734	2.760	2.749	2.754	2.710	2.750	2.739
Ga	InSb	2.683	2.697	2.678	2.673	2.676	2.686	2.680	2.683
In	AlP	2.487	2.472	2.494	2.490	2.493	2.447	2.498	2.480
Al	InP	2.412	2.427	2.408	2.401	2.405	2.415	2.400	2.414
In	AlAs	2.561	2.546	2.572	2.563	2.568	2.523	2.575	2.553
Al	InAs	2.493	2.506	2.487	2.483	2.487	2.497	2.480	2.495
In	AlSb	2.754	2.741	2.763	2.756	2.760	2.721	2.765	2.746
Al	InSb	2.693	2.705	2.689	2.685	2.687	2.696	2.683	2.693
Cd	ZnTe	2.756	2.740	2.760	2.760	2.763	2.720	2.770	2.755
Zn	CdTe	2.673	2.688	2.676	2.660	2.665	2.671	2.658	2.674
Hg	CdTe	2.800	2.801	2.800	2.800	2.799	2.801	2.790	
Cd	HgTe	2.804	2.803	2.804	2.805	2.805	2.804	2.813	
Hg	ZnTe	2.750	2.735	2.754	2.753	2.757	2.715	2.758	2.748
Zn	HgTe	2.671	2.685	2.674	2.659	2.664	2.671	2.662	2.673
As	AlP	2.425	2.418	2.429	2.427	2.428	2.406	2.427	2.422
P	AlAs	2.392	2.399	2.387	2.387	2.389	1.394	2.390	2.395
As	GaP	2.417	2.409	2.424	2.417	2.420	2.396	2.416	2.414
P	GaAs	2.386	2.393	2.380	2.382	2.383	2.389	2.386	2.387
As	InP	2.596	2.589	2.599	2.598	2.600	2.579	2.598	2.595
P	InAs	2.561	2.568	2.558	2.557	2.558	2.563	2.560	2.562
Sb	AlAs	2.584	2.566	2.597	2.587	2.592	2.539	2.577	2.574
As	AlSb	2.506	2.522	2.496	2.495	2.498	2.511	2.514	2.510
Sb	GaAs	2.569	2.553	2.584	2.571	2.576	2.524	2.546	2.564
As	GaSb	2.501	2.516	2.489	2.492	2.495	2.508	2.525	2.505
Sb	InAs	2.747	2.730	2.754	2.750	2.754	2.705	2.733	2.739
As	InSb	2.669	2.683	2.663	2.658	2.662	2.672	2.683	2.667
Sb	AlP	2.555	2.529	2.569	2.561	2.568	2.488	2.550	2.542
P	AlSb	2.440	2.462	2.426	2.425	2.430	2.447	2.448	2.444
Sb	GaP	2.526	2.503	2.551	2.529	2.537	2.461	2.504	2.519
P	GaSb	2.431	2.451	2.414	2.418	2.422	2.440	2.454	2.436
Sb	InP	2.712	2.687	2.720	2.719	2.725	2.654	2.702	2.700
P	InSb	2.599	2.619	2.591	2.585	2.590	2.604	2.611	2.597
Se	ZnS	2.420	2.409	2.421	2.424	2.426	2.396	2.426	2.420
S	ZnSe	2.367	2.376	2.365	2.360	2.364	2.370	2.363	2.367
Te	ZnSe	2.586	2.569	2.588	2.589	2.592	2.540	2.591	2.584
Se	ZnTe	2.501	2.517	2.497	2.490	2.494	2.504	2.495	2.502
Te	ZnS	2.543	2.513	2.544	2.552	2.558	2.478	2.557	2.539
S	ZnTe	2.406	2.429	2.400	2.390	2.396	2.410	2.396	2.407

TABLE IV. Deviations of the calculated impurity bond lengths (in Å) as compared with the experimental values from EXAFS.

Impurity	Host	A	C	D1	FPT	MZ	SSHS[a]	Expt.
In	GaAs	−0.028	−0.014	−0.025	−0.017	−0.031	−0.009	2.587[b]
Ga	InAs	0.004	−0.001	−0.005	−0.006	0.008	0.005	2.487[c]
Cd	ZnTe	0.004	0.008	−0.008	0.018	0.003	0.012	2.75(2)[c]
Zn	CdTe	−0.01 to −0.02	0.0 to −0.01	−0.02 to −0.03	−0.02 to −0.03	−0.01 to −0.02	0.0 to −0.01	2.68 to 2.69[c]
Te	ZnSe	−0.009	−0.07	−0.003	−0.004	−0.011	−0.004	2.595[d]
Se	ZnTe	0.005	0.01	−0.006	−0.001	0.006	0.004	2.496[d]
Mean absolute deviation	0.011	0.006	0.012	0.012	0.012	0.012	0.006	

[a] SSHS, Reference 8.

[b] Reference 1.

[c] Quoted in Reference 3.

[d] Reference 14.

A by 0.025 Å. Model A produces about the same d_1 values as Model D1, where the maximum difference in d_1 is only 0.015 Å. Martins and Zunger[3] used the same model as model A; however, their analytic expression for δ is different from that given by Eq. (21). Nevertheless, numerical results indicate that these two calculations agree to 0.01 Å. The slightly different forms of strain energies used in models D2 and D1 only introduce a small change in d_1, with the largest difference being less than 0.01 Å. The first-shell continuum model (model E) allows too little relaxation, so while the other models produce a ratio δ/δ_0 ranging from 0.6 to 0.8, model E only ranges from 0.4 to 0.6. The reason that the fixed boundary in model A works is that the effective shear coefficient C (see Table I) characterizing the strain energy in the elastic continuum is large. However, model B is too rigid and does not provide enough buffer between the impurity bond and the fixed boundary.

The comparison of the theoretical results with the available experimental data in Table IV indicates that models B and E are the least accurate. Models A, D1, that of Martins and Zunger (MZ), and the full theory are comparable in that all have an average absolute deviation of 0.012 Å, which is about the experimental uncertainty in extended x-ray-absorption fine structure (EXAFS). The agreement between theory and experiment, however, is not uniform. The most surprising result in Table IV is that the simple spring model (model C) and its cruder version used by Shih *et al.*[8] ($\alpha=\alpha_I$ so $\delta/\delta_0=0.75$, labeled as SSHS) have the smallest variance in d_1, about 0.006 Å. We know this does not imply that the simple spring model represents the real picture of bond-length relaxation. For example, if we let all the shear coefficients be 0, i.e., $\beta=C=0$ in our model, then as the range of the boundary is gradually extended, the local bond length will eventually relax to the impurity bond length $d_1=d_I$, or $\delta=\delta_0$. This can be seen in model A from Eq. (21), where δ reduces to $\delta_0/(1+\alpha/6\alpha_I)$, rather than $\delta_0/(1+\alpha/3\alpha_I)$ as predicted by model C, and in model D from Eqs. (27) and (30), δ becomes δ_0, if the continuum is taken to be shearless. Considering that various effects are included that may mask the absolute accuracy of d_1 predictions (e.g., while low-temperature bond lengths are used in the calculation, the room-temperature values of elastic constants are adopted), the agreement of various models with experiments in Table IV should be regarded as excellent. There are, however, many other impurity systems in which the simple-spring-model predictions differ considerably from other models, as is shown in Fig. 3, where δ is plotted against δ_0 for the full theory. Those points that deviate significantly from the 0.75-slope line are the systems with (As,Sb) or (P,Sb) substitutions. Additional EXAFS measurements on these systems are needed to test these predictions.

C. Mixing enthalpies

Table V lists the mixing enthalpy parameters Ω (in kcal/mole) for a number of alloys estimated from Eq. (41) for all the models considered, along with other theoretical[3,16–18] and experimental values.[16,19] As already dis-

TABLE V. Mixing enthalpy parameter Ω (in kcal/mole) estimated from the full perturbation theory and several valence force models discussed in Sec. III, and comparison with experiments and other theories.

	A	B	C	D1	D2	E	FPT	MZ[a]	DL[b]	FM[c]	VV[d]	Expt.[e]
(Ga,Al)P	0.00	0.01	0.00	0.01	0.01	0.01	−0.05					
(Ga,Al)As	0.00	0.00	0.00	0.00	0.00	0.00	−0.07	0.02	0.02	0.03	0.11	0.0
(Ga,Al)Sb	0.02	0.03	0.02	0.02	0.02	0.03	−0.15	0.02	0.02	0.03		0.0
(Ga,In)P	3.76	4.79	3.0	3.29	3.39	5.24	2.54	4.56	3.63	2.94		3.25,3.5
(Ga,In)As	2.97	3.76	2.36	2.61	2.69	4.14	1.60	2.49	2.81	2.42	1.25	1.65, 2.0, 3.0
(Ga,In)Sb	2.22	2.83	1.77	1.95	2.00	3.09	0.81	2.53	1.85	1.83		1.47, 1.9
(In,Al)P	3.24	4.22	2.77	2.78	2.87	4.60	2.55					
(In,Al)As	2.86	3.65	2.32	2.49	2.56	3.93	2.17	3.60	2.81	2.37		2.5
(In,Al)Sb	1.81	2.33	1.49	1.57	1.61	2.50	1.36	2.06	1.46	1.45		0.6
(Cd,Zn)Te	1.80	2.43	1.73	1.43	1.49	2.45	1.24	2.12	1.97	1.63		1.34
(Hg,Cd)Te	0.00	-0.00	0.00	0.00	0.00	0.00	−0.07					0.7, 1.4
(Hg,Zn)Te	1.63	2.20	1.56	1.30	1.36	2.23	1.50	1.91	1.81	1.48		3.0
Al(P,As)	0.81	1.03	0.65	0.71	0.73	1.14	0.76					
Ga(P,As)	0.95	1.18	0.70	0.86	0.87	1.32	0.94	1.15	0.98	0.66	0.12	0.4, 1.0
In(P,As)	0.60	0.78	0.52	0.51	0.53	0.84	0.57	0.72	0.58	0.52		0.4
Al(As,Sb)	4.31	5.45	3.38	3.80	3.88	5.92	4.09					
Ga(As,Sb)	3.77	4.69	2.81	3.40	3.46	5.22	3.67	4.58	3.35	2.76		4.0, 4.5
In(As,Sb)	2.61	3.39	2.23	2.24	2.31	3.67	2.52	2.89	2.29	2.17	6.65	2.25, 2.9
Al(P,Sb)	8.60	10.99	6.99	7.54	7.73	12.00	8.32					
Ga(P,Sb)	8.54	10.61	6.36	7.72	7.88	11.66	8.66					
In(P,Sb)	5.87	7.64	5.08	4.99	5.15	8.04	5.76					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(Se,Te)	2.47	3.27	2.23	2.09	2.16	3.63	2.26	2.91	3.11	2.12	3.12	1.55
Zn(S,Te)	7.02	9.34	6.45	5.80	6.02	9.72	6.20					

[a] Reference 3, column A of Table III.
[b] Reference 16.
[c] Reference 18.
[d] Reference 17.
[e] References 16 and 19.

cussed, the chemical terms reduce the excess energies in the cation impurities and increase them for anion impurities. The corresponding changes in Ω are the differences between the FPT and Model D2. We note that the reductions of Ω for the (Ga,In) alloys are very large (>1 kcal/mole) and also significant for (In,Al) alloys. However, the increases in Ω for the anion substitutional alloys are not as large. Also, the Ω differences between models D1 and D2 are less than 10%. Model A produces Ω values about 20% larger than model D1, model B in turn is 20% higher than model A, and model E is 10% higher than model B. The Ω values in the simple spring model (model C) are seen to be about the same as model D1, although the differences among systems can be positive or negative. Although MZ used the same strain model as model A, their Ω values do not agree with our model A values because their way of estimating Ω is different. In fact, the values of MZ are closer to model B than to A.

To distinguish the quantitative nature of different theoretical models, we note that there are also important factors that may mask the comparison between theory and experiment for Ω. One important factor is that the mixing enthalpies extracted from phase-diagram analysis are sensitive to sample and experimental conditions. These Δ_{H_m} contain contributions from various nonideal structures such as vacancies, impurities, dislocations, grain boundaries, and surface conditions, in addition to the ideal ΔH_m considered here for solid solutions. Thus, our theoretical ΔH_m should represent a lower bound. Another complication comes from the version of the theory of solid solution adopted. The theory used for analysis so far assumes a regular solid solution with second-neighbor pair interactions as was outlined in Sec. V. Recent experimental[20-22] and theoretical[23] studies

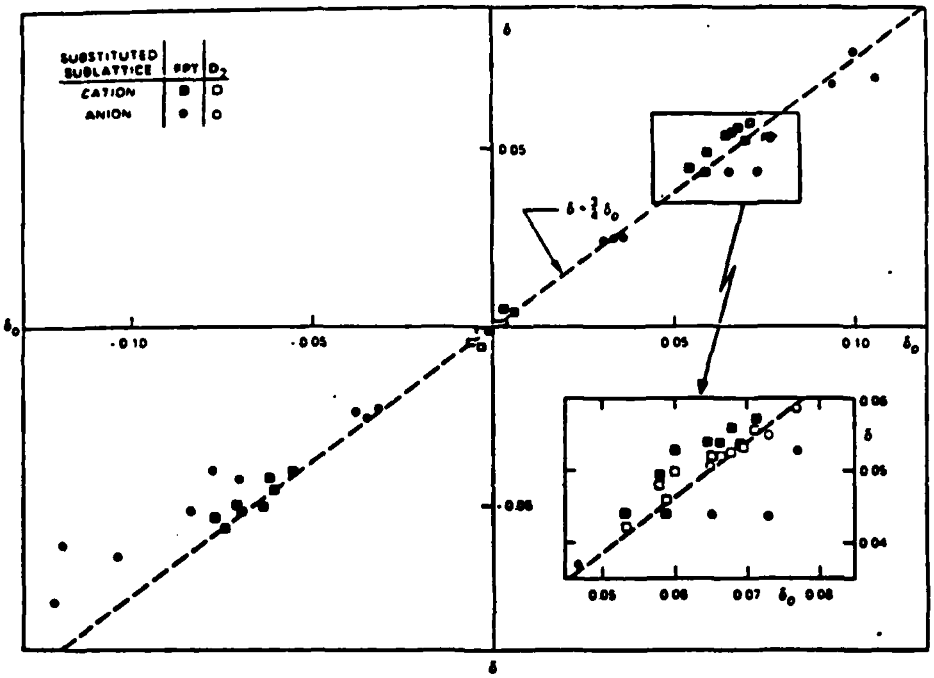
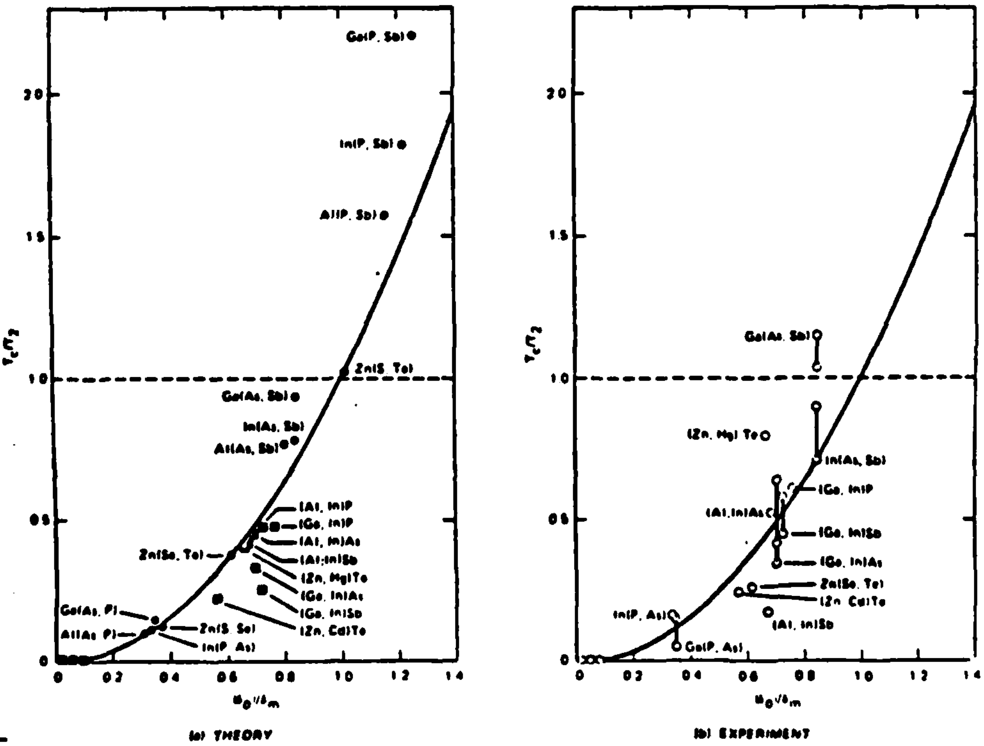


FIG. 3. Calculated bond-relaxation parameter δ from FPT and D2 as a function of δ_0. The $\delta = \frac{3}{4}\delta_0$ curve corresponds to theory of Shih, Spicer, Harrison, and Sher (Ref. 8).

have suggested the possibility of compositional clustering or long-range correlations in alloys. Extending the theory to include such effects will alter the simple results for ΔH_m from Eq. (41). Moreover, there is evidence from the composition variation of the alloy hardness[24] and from the optical-phonon frequencies[25] that the shear coefficients of alloys increase near the center of the composition range. This will cause the effective continuum shear coef-



(a) THEORY

(b) EXPERIMENT

FIG. 4. Plot of T_c/T_2 as a function $|\delta_0|/\delta_m$ for T_c obtained from (a) FPT and (b) the experimental Ω in Table V. In (a), the circles are for anion alloys, the squares for cation substitution. The solid lines in both parts correspond to the simple theory discussed in Appendix C. The dashed lines at $T_c/T_2 = 1.0$ separate the miscible from immiscible groups.

TABLE VI. Comparison of the critical temperature T_c of mixing and melting temperatures of the constituents T_1 and T_2, in the order of their appearance in the parentheses. Also tabulated are the averaged absolute values of $|\delta_0|$ and the ratio $|\delta_0|/\delta_m$.

System	$\lvert\delta_0\rvert$ (%)	T_c (K)	T_c/T_1	T_c/T_2	$\lvert\delta_0\rvert/\delta_m$
(Al,Ga)Sb	0.6	0	0	0	0.067
(Al,Ga)As	0.1	0	0	0	0.009
(Al,Ga)P	0.3	0	0	0	0.029
(Al,In)Sb	5.5	342	0.25	0.42	0.679
(Ga,In)Sb	5.8	204	0.21	0.25	0.716
(Al,In)As	6.8	547	0.29	0.45	0.687
(Ga,In)As	6.9	403	0.23	0.34	0.697
(Al,In)P	7.1	642	0.36	0.48	0.732
(Ga,In)P	7.4	639	0.43	0.48	0.763
In(P,As)	3.2	144	0.11	0.12	0.330
Ga(P,As)	3.7	236	0.14	0.15	0.352
Ga(P,As)	3.5	191	0.10	0.11	0.307*
In(As,Sb)	6.8	635	0.52	0.79	0.840
Ga(As,Sb)	7.6	924	0.53	0.94	0.844
Al(As,Sb)	8.1	1030	0.56	0.78	0.810
In(P,Sb)	9.9	1450	1.08	1.82	1.222
Ga(P,Sb)	11.3	2180	1.25	2.21	1.256
Al(P,Sb)	11.6	2095	1.19	1.58	1.116
(Cd,Hg)Te	0.3	0	0	0	0.033
(Zn,Hg)Te	6.0	377	0.25	0.40	0.659
(Zn,Cd)Te	6.2	312	0.21	0.23	0.564
Zn(S,Se)	4.7	226	0.11	0.13	0.362
Zn(Se,Te)	7.2	569	0.32	0.38	0.615
Zn(S,Te)	11.9	1560	0.74	1.03	1.017

ficient C in Eq. (14) to be composition dependent, which will cause Ω to increase. Despite these uncertainties, useful comparisons across the board in Table V can still be made.

Based on the above considerations, we can conclude that models B, E, and MZ predict Ω values that are too high. We should emphasize that all the Ω numbers for models from A through MZ are directly calculated without any adjustable parameters. The fact that models A, C, and D1,D2, and the FPT agree with the experiment as well as or even slightly better than the one-parameter theories[16,18] of DL and Fedder and muller (FM) is already quite an accomplishment. The few numbers taken from Van Vechten's calculations[17] (VV) indicate that the dielectric model predicts results at larger variance with experiments. There are two important implications about the FPT that can be drawn from Table III. First, the theory predicts a small but negative Ω value for several alloys. This not only means that there is no miscibility gap in these alloys but also implies a tendency toward ordering, in which the substitutional atoms tend to be surrounded by different second-neighbor species. For stoichiometric compositions, this implies a tendency toward compound formation. Secondly, the FPT tends to predict smaller Ω values than observed experimentally, which should be expected according to our discussion. To the extent that the FPT predicts the correct ΔH_m values for an ideal solution, the difference $\Delta H_m^{expt} - \Delta H_m$ may be attributed to imperfect conditions and deviations from the ideal solution theory.

Finally, the calculated Ω values in Table V provide some guidance in separating the completely miscible alloys from immiscible ones.[26,27] In a true random alloy, the criteria[28] for alloy mixing at a temperature T is that $T \geq T_c$, where the critical temperature T_c is given by $\Omega/(2R_g)$, with R_g being the universal gas constant.[29] For an $A_xB_{1-x}C$ alloy to be miscible throughout the whole concentration range, the requirement is that both the melting temperatures T_1 and T_2 of the pure AC and BC compounds be greater than T_c. Table VI lists the values of T_c associated with the Ω values in the FPT, the ratios T_c/T_1 and T_c/T_2, and the average absolute values of δ_0 for the alloys considered in Table V. In Table VI, T_c is set equal to zero if Ω is negative and T_2 is chosen to be the lower value of the two melting temperatures, so the criterion for not having a miscibility gap is $T_c/T_2 < 1$.

There is an empirical rule[26] stating that this will be satisfied if the lattice mismatch $|\delta_0|$ between the two alloy components is less than 7.5%. However, we find that (see Appendix C) a more precise rule is $|\delta_0| > \delta_m$, where $\delta_m = 1.63X_m$ and X_m is the ratio of the rms bond-length amplitude fluctuation to the bond length at the melting temperature T_2. The values of T_m for the compounds involved[30] and the associate X_m values estimated from Eqs. (C2) and (C3) are tabulated in Table I. The model used in Appendix C yields $T_c/T_2 = (\delta_0/\delta_m)^2$. This suggests that it is instructive to plot T_c/T_2 as a function of $|\delta_0|/\delta_m$, as is shown in Fig. 4 for the T_c calculated from FPT. This plot is similar to the ΔE versus δ_0 curve in Fig. 2 because, in fact, Ω is proportional to the sum of the ΔE values of the two constituents [see Eq. (41)]. However, if T_c/T_2 is plotted against $|\delta_0|$ alone, the FPT points are much more scattered, and those of SSHS would not even exhibit a smooth simple quadratic form because the lower melting temperature T_2 is not a smooth function of $|\delta_0|$. This result suggests that $|\delta_0|/\delta_m < 1$ is a better criterion than $|\delta_0| < 0.075$. Figure 3 also clearly shows the chemical effects: all the cation-substitution alloy points lie below the solid curve and all the anion-substitution alloys have (T_c/T_2) values on or above the curve, corresponding to negative and positive shifts in ΔE due to the chemical terms. Again, the curve based on the SSHS model is an excellent universal representation. From the figure, we see that all (P,Sb) alloys should have miscibility gaps and all (As,Sb) alloys are predicted to be miscible, although on the borderline, because the actual mixing enthalpies are larger than these ideal calculated values. The figure also shows that Zn(S,Te) has a miscibility gap but a smaller value of T_c/T_2 than the (P,Sb) alloys, despite the fact that its $|\delta_0|$ value is larger. All these predictions are consistent with the available experimental evidence.

VII. SUMMARY AND CONCLUSION

In this paper a simple theory of defect substitution energies is formulated. The substitution energy is compactly separated into a replacement energy Δ_R and a distortion energy of the pure host crystal [see Eq. (2)]. However, a rigorous application of this theory requires an improvement in certain quantitative aspects of Harrison's bonding theory,[7] particularly the elastic constants.[5] The most interesting application of this theory presented in this paper is its perturbation form which enables us to absorb the measured elastic coefficients into the calculation and, more importantly, to study the chemical effects. The origin of chemical influences on impurity bond relaxation can be attributed to three mechanisms [see Eq. (17)]: a chemical force F_{ch} that either helps or hinders lattice relaxation, depending on whether it has the same or opposite sign from the bond-length difference $d\text{-}d_I$ between the host and impurity, a chemical energy that depends on the difference of the polarities between the impurity and host bonds, ΔV_3, and an effective elastic force constant H that, when positive, also tends to restrain the lattice from distortion. To study the effect of boundaries between the core atoms around an impurity and the rest of the elastic medium, various models based on the valence force field[10] are derived and their results are compared with the full perturbation theory and available experimental data. We found at least five models, including the FPT, that produce the correct impurity bond lengths with variances for the compounds studied about equal to the experimental uncertainties in EXAFS[10-14] (~0.01 Å). However, some models are oversimplified and will certainly not predict other properties equally well. However, more experimental lattice constant measurements to further test the theory, particularly on (As,Sb) and (P,Sb) substitution systems for which there are larger differences between different models, are needed. It would also be instructive to see if the predicted reversal for HgCdTe is found.

The excess energies of impurity substitution are also shown to provide good estimates of the mixing enthalpies Ω of pseudobinary alloys. The chemical shifts are found to have a negative net contribution to Ω for most cation substitutions, but positive contributions for anion substitutions. The chemical reduction of Ω in (Ga,In) alloys is larger than 1 kcal/mole (30–100%). Several VFF models and the full perturbation theory produce results for Ω that are as good as the best theories with one adjustable parameter. However, the full theory tends to yield answers on the low side of the experimental values, which we argue is as it should be because there are nonideal structures that also contribute to Ω. The calculated Ω values and the melting temperatures are used to predict the existence of alloy miscibility gaps, and the results correlate well with experiments.

Finally, we wish to comment on the accuracy of the theories that are connected to the present model. The perturbation theory has already been stretched beyond its expected region of validity and predicts d_1 to within experimental uncertainties (~0.01 Å) even for cases with large bond-length differences ($\delta_0 \sim 0.1$). The accuracy can only be improved if the full nonperturbation theory outlined in Sec. II is used. This calculation is needed for the strong substitution cases that were not considered in this paper: examples are (B,Ga), (B,In), (N,P), (N,As), (N,Sb), (O,S), (O,Se), and (O,Te) substitutions. Although we believe that for the properties treated, the model with a continuum attached to the second shell is as accurate as the perturbation theory used, it remains to be seen if this is true for other properties, especially strain coefficients. Finally, the present theory has been extended to study alloys[5,6] by embedding clusters in an effective medium. This enables us to study the bond length and energy variations throughout the whole concentration range. However, a quantitative calculation still awaits an improvement of the accuracy of Harrison's theory. A similar procedure is also being extended to a study of the alloy electronic structure for which a cluster CPA (coherent-potential approximation) involving both potential and structural disorder[31] will be used.

ACKNOWLEDGMENTS

This work benefited from useful discussions with Professor W. A. Harrison and Professor A. Zunger. One of us (A.-B. C.) would like to thank Professor W. E. Spicer for his hospitality. The work is supported in part by U.S.

Air Force Office of Scientific Research Contract No. AFOSR-84-0284 and U.S. Defense Advanced Projects Agency DARPA Contract No. MDA-903-83-C-0108.

APPENDIX A: ELASTIC ENERGY IN CONTINUUM

In Sec. III the elastic energy outside a sphere of radius R centered at the impurity is assumed to be a continuum with a radial displacement $u\alpha\hat{r}/r^2$. If the displacement at R is u_0, then $\mathbf{u}(\mathbf{r})=u_0(R^2/r^2)\hat{\mathbf{r}}$. The energy density in the continuum is given by

$$\begin{aligned}\delta\varepsilon(\mathbf{r}) &= \tfrac{1}{2}C_{11}(e_{xx}^2+e_{yy}^2+e_{zz}^2)\\ &\quad +C_{12}(e_{yy}e_{zz}+e_{xx}e_{yy}+e_{xx}e_{zz})\\ &\quad +\tfrac{1}{2}C_{44}(e_{xy}^2+e_{yz}^2+e_{zx}^2)\,,\end{aligned}$$

where

$$e_{xx}=\frac{\partial u_x}{\partial x}=R^2u_0(r^2-3x^2)/r^5\,,$$

$$e_{xy}=\frac{\partial u_x}{\partial_x}+\frac{\partial u_y}{\partial x}=-6R^2u_0xy/r^5,\ldots.$$

Thus, the total elastic energy in the continuum is

$$\begin{aligned}\Delta_{\rm dis}^{\rm cont} &= \int_R^\infty \delta\varepsilon(\mathbf{r})d^3r\\ &=4\pi Ru_0^2(\tfrac{2}{5}C_{11}-\tfrac{2}{5}C_{12}+\tfrac{6}{5}C_{44})\\ &=CRu_0^2\,,\end{aligned}$$

where the effective shear coefficient is given by

$$C=\pi(1.6C_{11}-1.6C_{12}+4.8C_{44})\,.$$

APPENDIX B: DISTORTION ENERGY

In this Appendix we count the detailed contributions of the bond-stretching terms $\Delta(\mathbf{r}_i\cdot\mathbf{r}_i)$ and "bond-bending" terms $\Delta(\mathbf{r}_i\cdot\mathbf{r}_j)$ for $i\neq j$ in VFF [Eq. (13)] that enter Eq. (16) in FPT and in the VFF models in Sec. III.

1. *α* and *β* terms from the first-shell bonds

The four bond vectors pointing away from the central impurity according to Fig. 1 are

$$\mathbf{r}_1=(1-\delta,1-\delta,1-\delta)d/\sqrt{3}\,,$$

$$\mathbf{r}_2=(1-\delta,-1-\delta,-1-\delta)d\sqrt{3},\ldots.$$

Thus, $\Delta(\mathbf{r}_1\cdot\mathbf{r}_1)=-2\delta d^2$ and $\Delta(\mathbf{r}_1\cdot\mathbf{r}_2)=-\frac{2}{3}\delta d^2$. The α terms contribute

$$4\times3\alpha_I(-2\delta d^2)^2/8d^2=6\alpha_I\delta^2d^2\,,$$

and the β terms contribute

$$6\times3\beta_I(-\tfrac{2}{3}\delta d)^2/8d^2=\beta_I\delta^2d^2\,.$$

If an A atom is replaced by a B atom, as was done in FPT, the α_I and β_I are replaced by α and β, respectively.

2. *α* terms from the second-shell bonds, *β* terms between the first- and second-shell bonds and among the second-shell bonds

For these terms we need to consider the four bond vectors pointing away from $\mathfrak{C}$. They are

$$\mathbf{r}_1=(-1+\delta,-1+\delta,-1+\delta)d/\sqrt{3}\,,$$

$$\mathbf{r}_2=(-1+\delta,1+\delta+\gamma,1+\delta+\gamma)d/\sqrt{3}\,,$$

$$\mathbf{r}_3=(1+\delta+\gamma,-1+\delta,1+\gamma+\delta)d/\sqrt{3},\ldots.$$

Then $\Delta(\mathbf{r}_2\cdot\mathbf{r}_2)=\frac{2}{3}(\delta+2\gamma)d^2$, $\Delta(\mathbf{r}_1\cdot\mathbf{r}_2)=-\frac{2}{3}(\delta+\gamma)d^2$, and $\Delta(\mathbf{r}_2\cdot\mathbf{r}_3)=\frac{2}{3}\delta d^2$. Thus, the α terms from the second-shell bonds become

$$4\times3\times3\alpha[\tfrac{2}{3}(\delta+2\gamma)d^2]^2/8d^2=2\alpha(\delta+2\delta)^2d^2\,,$$

the β term between the first- and second-shell bonds are

$$4\times3\times3\beta[\tfrac{2}{3}(\delta+\gamma)d^2]^2/8d^2=2(\gamma+\delta)^2\beta d^2\,,$$

and the β terms among the second-shell bonds are

$$4\times3\times3\beta(\tfrac{2}{3}\delta d^2)^2/8d^2=2\delta^2\beta d^2\,.$$

3. *α* terms for the third-shell bonds, *β* terms between the second- and third-shell bonds and among the third-shell bonds adjacent to the second-shell atoms

For these terms we need to consider the bond vectors pointing away from B in Fig. 1. They are

$$\mathbf{r}_2=(1-\delta,-1-\gamma-\delta,-1-\gamma-\delta)d/\sqrt{3}\,,$$

$$\mathbf{r}_1=(1+\gamma',1+3\gamma'-\gamma,1+3\gamma'-\gamma)d/\sqrt{3}\,,$$

$$\mathbf{r}_3=(-1-\gamma'',1+3\gamma''-\gamma,-1+\gamma''-\gamma)d/\sqrt{3}\,,$$

and

$$\mathbf{r}_4=(-1-\gamma'',-1+\gamma''-\gamma,1+3\gamma''-\gamma)d\sqrt{3}\,.$$

Thus, we have

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_2)=-\frac{d^2}{3}(3\delta+5\gamma')\,,$$

$$\Delta(\mathbf{r}_2\cdot\mathbf{r}_3)=\frac{d^2}{3}(\delta+2\gamma-5\gamma'')\,,$$

$$\Delta(\mathbf{r}_3\cdot\mathbf{r}_4)=-\tfrac{2}{3}d^2\gamma''\,,$$

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_3)=\frac{d^2}{3}(-2\gamma-\gamma'+3\gamma'')\,,$$

$$\Delta(\mathbf{r}_3\cdot\mathbf{r}_3)=\frac{d^2}{3}(6\gamma'')\,,$$

and

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_1)=\frac{d^2}{3}(14\gamma'-4\gamma)\,.$$

For model A, $\gamma'=r''=0$, so the α terms of the third-shell bonds become

$$4\times3\times3\alpha[2(\Delta\mathbf{r}_3\cdot\mathbf{r}_3)^2+(\Delta\mathbf{r}_1\cdot\mathbf{r}_1)^2]/8d^2=8\alpha\gamma^2d^2\,,$$

the β terms between the second- and third-shell bonds are

$$4\times 3\times 3\beta[(\Delta \mathbf{r}_1\cdot\mathbf{r}_2)^2+2(\Delta \mathbf{r}_2\cdot\mathbf{r}_3)^2]/8d^2$$
$$=\tfrac{9}{2}\beta d^2[\delta^2+\tfrac{2}{9}(\delta+2\gamma)^2]=\tfrac{9}{2}\beta\delta^2 d^2+\beta(\delta+2\gamma)^2 d^2 ,$$

and the β terms among the third-shell bonds adjacent to the second-shell atoms are

$$4\times 3\times \beta[2(\Delta \mathbf{r}_1\cdot\mathbf{r}_3)^2/(8d^2)+(\Delta \mathbf{r}_3\cdot\mathbf{r}_4)^2]=4\beta\gamma^2 d^2 .$$

For continuum and the only contribution from this group are the β terms between the second- and third-shell bonds. Since the displacements in the continuum are proportional to $1/R^2, \gamma'=8\sqrt{2}\gamma/(19\sqrt{19})$ and $\gamma''=8\sqrt{2}\gamma/(11\sqrt{11})$. Thus, these β terms become

$$\frac{9}{2d^2}\beta\left[\frac{d^4}{9}(3\delta+5\gamma')^2+\frac{2d^4}{9}(\delta+2\gamma-5\gamma'')^2\right]$$
$$=\tfrac{1}{2}\beta\left[\left[3\delta+\frac{40\sqrt{2}}{19\sqrt{19}}\gamma\right]^2+2\left[\delta+2\gamma-\frac{40\sqrt{2}}{11\sqrt{11}}\gamma\right]^2\right]d^2 .$$

4. β terms for bonds adjacent to the third-shell atoms

These terms only enter model A, so $r'=r''=0$. There are two different groups, one like those adjacent to C' and another like those meeting at C''. The four bond vectors pointing away from C' are

$$\mathbf{r}_1=(-1,-1+\gamma,-1+\gamma)d/\sqrt{3} ,$$
$$\mathbf{r}_2=(-1,1,1)d/\sqrt{3} ,$$
$$\mathbf{r}_3=(1,-1,1)d/\sqrt{3} ,$$

and

$$\mathbf{r}_4=(1,1,-1)d/\sqrt{3} .$$

Thus, the only contribution from this group is

$$4\times 3\times 3\beta[\Delta(\mathbf{r}\cdot\mathbf{r}_2)]^2/8d^2=2\beta\gamma^2 d^2 .$$

The four bond vectors around C'' are

$$\mathbf{r}_1=(-1+\gamma,-1-\gamma,-1)d/\sqrt{3} ,$$
$$\mathbf{r}_2=(-1,1,1)d/\sqrt{3} ,$$
$$\mathbf{r}_3=(1,-1+\gamma,1+\gamma)d/\sqrt{3} ,$$

and

$$\mathbf{r}_4=(1,1,-1)d/\sqrt{3} ,$$

which only results in the first-order term $\Delta(\mathbf{r}_2\cdot\mathbf{r}_3)=\frac{2}{3}\gamma d^2$. Thus, the group contributes to

$$4\times 3\times 3\beta[2(\Delta \mathbf{r}_2\cdot\mathbf{r}_3)^2]/8d^2=4\beta\gamma^2 d^2 ,$$

and the combined contribution from these two groups is $6\beta\gamma^2 d^2$.

APPENDIX C

Starting with Eq. (24) and using the SSHS model $\alpha=\alpha_l$, one finds the mixing enthalpy parameter Ω to be

$$\Omega=\tfrac{3}{2}\bar{\alpha}(d_{AC}-d_{BC})^2 N_0 , \qquad \text{(C1)}$$

where N_0 is Avagodro's number and $\bar{\alpha}=\frac{1}{2}(\alpha_{AC}+\alpha_{BC})$. Then relate the mean-square bond-length fluctuation $\langle\zeta^2\rangle$ at the melting temperature T_m to T_m for a compound by equating the average potential energy per unit cell to half of the thermal energy:

$$\langle V_{\rm pot}\rangle\simeq 4\times\tfrac{3}{2}\alpha\langle\zeta^2\rangle=\tfrac{1}{2}(2\times 3k_B T_m) , \qquad \text{(C2)}$$

where k_B is the Boltzmann constant. Defining a Liedermann ratio of melting χ_m by

$$(\langle\zeta^2\rangle)^{1/2}\approx\chi_m d \qquad \text{(C3)}$$

and choosing the mixing criterion to be $T_c/T_m<1$, where T_m now is the smaller value of the two melting temperatures of the constituent compounds, we require that

$$\frac{T_c}{T_m}=\frac{\Omega k}{4R_g\alpha\chi_m^2 d^2}=\frac{3}{8}\,\frac{(d_{AC}-d_{BC})^2}{\chi_m^2 d^2}<1 \qquad \text{(C4)}$$

or

$$|\delta_0|/\delta_m<1 , \qquad \text{(C5)}$$

where $\delta_m=1.63\chi_m$ and $|\delta_0|$ is the percentage bond-length difference.

---

[*] Present address: Stanford University, Stanford, CA 94305.

[1] J. C. Mikkelsen and J. B. Boyce, Phys. Rev. Lett. 49, 1412 (1982); Phys. Rev. B 28, 7130 (1983).

[2] A. Zunger and J. E. Jaffe, Phys. Rev. Lett. 51, 662 (1983).

[3] J. L. Martins and A. Zunger, Phys. Rev. B 30, 6217 (1984).

[4] A. Balzarotti, A. Kisiel, N. Motta, M. Zimmal-Starnawska, M. T. Czyzyk, and M. Podgorny, Phys. Rev. B 30, 2295 (1984).

[5] A.-B. Chen and A. Sher, Microscience 3, xx (1984).

[6] A. Sher, A.-B. Chen, and W. E. Spicer, in Proceedings of the 13th International Conference on Defects in Semiconductors, August 12–17, 1984, Coronado, California (unpublished), p. 43.

[7] W. A. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980); Microscience 3, 35 (1983); Phys. Rev. B 27, 3592 (1983).

[8] K. Shih, W. E. Spicer, W. A. Harrison, and A. Sher, Phys. Rev. B 31, 1139 (1985).

[9] A. Baldereschi and J. J. Hopfield, Phys. Rev. Lett. 28, 171 (1972).

[10] R. M. Martin, Phys. Rev. B 1, 4005 (1970).

[11] The experimental elastic constants were taken from the tabulated values in Refs. 10 and 25 except AlP and AlAs, for which we used the estimated values by J. D. Wiley, in *Semiconductors and Semimetals*, edited by R. K. Willardson and

(1) Phenomena Influencing the Dislocation Density of Semiconductor Compounds and Alloys(a)

A. Sher, SRI International, Menlo Park, CA 94062
An-Ban Chen, Auburn University, Auburn, AL 36489
W.E. Spicer(b), Stanford University, Stanford, CA 94305

Abstract

The objective of this paper is to identify the principal microscopic phenomena controlling dislocation densities in bulk grown semiconductors. Then, based on this understanding, a strategy for selecting materials to reduce dislocation densities is offered. The relevant quantities are calculated from an extension of Harrison's bonding theory, which, with our improved accuracy relates properties of the solids to the constituent atoms' valence electron energy states and wave functions. We report on the alloy composition variation of bond energies, bond lengths, charge redistribution among constituents, vacancy formation energies, dislocation energies, and hardness. Several III-V and II-VI compound semiconductors are treated including, GaAs, GaInAs, HgCdTe, and ZnHgTe.

(a) This work was supported in part by AFOSR Contract F49620-81-K0012 and DARPA Contract MDA-903-83-C-0108.

(b) *Stanford Ascherman Professor of Engineering*

Introduction

According to a currently accepted model (1) of the mechanism leading to islocations in bulk-grown material at a given temperature and temperature radient behind the growth front; their density is governed by:

- The vacancy density that depends on the formation energies,
- The competition between vacancy annealing rates and vacancy interaction caused clusters,
- The condensation rate of these vacancy clusters into dislocation loops, and
- The subsequent growth rate of these loops.

he objective of this paper is to identify some of the principal microscopic henomena controlling these features in semiconductors compounds and their seudobinary alloys. If any of the foregoing steps can be inhibited, then here will be fewer dislocations in bulk-grown crystals.

A model of the bonding of tetrahedrally coordinated semiconductor comounds due to Harrison underlies this work (2). He has derived expressions or the bond energies and strain coefficients of pure semiconductor compounds n terms of the constituent atom's valence state wave functions and energies. here are four contributions to the bond energies:

- A covalent energy arising from the interaction between sp^3 hybrids on adjacent sites, which, accord .g to a universal rule deduced by Harrison, is related to the inverse square of the bond length, d,
- An ionic energy which is proportional to the energy difference between the sp^3 hybrid energies of the anion and cation,
- A metallic energy arising from the interaction between filled bonding and unfilled antibonding states on adjacent bonds, and
- An overlap repulsion energy which is taken to vary as d^{-4}.

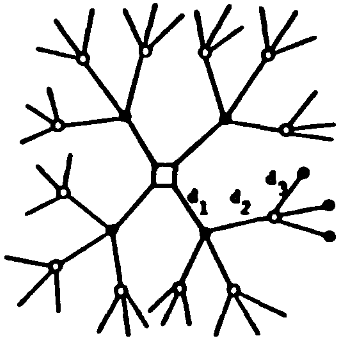
he shear coefficients, which play an important role in dislocation energies, re shown to vary for covalent solids as the covalent energy per cell unit olume or as d^{-5}. The ionic energy arising as it does from coulomb interctions, is insensitive to bond angle distortions and depends only on bond engths. Thus, in the ionic semiconductors the shear coefficients are reduced rom those of equivalent bond length covalent compounds, but also the power aw dependence on bond length increases to d^{-11} for the extreme ionic limit (3). he metallic interaction is bond angle sensitive and contributes to the d^{-11} ower law.

We have modified Harrison's theory in several ways (3) (4). First, nstead of using atomic term values as the input energies in the bonding alculation, we use correlated atomic energies that take account of all the tomic electronic energy changes associated with a state change. In our prior ublications these energies were calculated from pseudopotentials (3). In his paper we have modified the procedure and taken the ground state energies rom free atom ionization experiments, and only calculate the excitation nergies from the pseudopotentials. Second, the Harrison's theory has been xtended to alloys in which each type of the individual bond's energy and ength changes and the net alloy substitution energies are calculated. A heory of dislocation energies and hardness of semiconductors has been advanced.(4) n this paper the second item will be emphasized since item one is less elevant to this conference and these will be published elsewhere.

Bond Length and Energy Changes in an Alloy

Single Impurity

Start by considering the simplest defect, one cation of a host semiconductor compound AC is replaced by an isoelectronic impurity I, as sketched in Fig. 1. This will be generalized to an $AB_{1-x}C_x$ alloy later and the cation substituted case $A_{1-x}B_xC$ follows by symmetry. To study the principal effect of core lattice distortions, we allow the positions of the first shell of atoms labeled by bond length d_1 to move but fix the atoms of the second neighbors and beyond at their pure-crystal positions. Eventually, the effects of long range strain fields, generated by a point distortion, must be added to this core distortion. Here we allow only three different bond lengths $d_1 = (1-\delta)d$, $d_2 = (1 + 2/3\delta + \delta^2)^{1/2}d$, and $d_3 = d$, where d is the bond length of the pure host AC compound, and δ is a scale parameter.



Schematic representation of the bonds around an atom of interest (designated by the square). The first, second, and third neighbor bonds are designated d_1, d_2, and d_3

There are four bonds with energy E_1, twelve with energy E_2, and 36 with energy E_3. These energies can be written in closed form in the Harrison theory. The energies E_1 and E_2 differ from the host bond energy E_o because of bond length and angle distortions as well as difference in the ionic energy of I and C atom induced charge transfer. The energy E_3 differs from E_o only by the charge transfer couple through the metallization terms. There are two interesting energies we can calculate. The first is the energy Δ_s required to substitute the I atom for a C atom, i.e. bring a free I atom from infinity and replace a C atom that is taken from the crystal and removed to infinity.

$$\Delta_s = 4\,\Delta E_1 + 12\,\Delta E_2 + 36\,\Delta E_3 - (\epsilon_C - \epsilon_I) \qquad (1)$$

where $\Delta E_j = E_j - E_o$, $j = 1,2,3$, and ϵ_C and ϵ_I are the free atom valence electron binding energies. The second is the bond energy change of the impurity AI bond relative to the bond energy of a pure AI compound, denoted Δ_I

$$\Delta_I = \Delta E_1 - \frac{1}{4}(\epsilon_I - \epsilon_C) \qquad (2)$$

This energy tells us if the AI bond is stabilized ($\Delta_I < 0$) or destabilized ($\Delta_I > 0$) when it is in an AC host. If we define Δ_b as the bond energy difference between the IA and CA bonds each in their respective pure crystals

$$\Delta_b = BE(IA) - BE(CA) \qquad (3)$$

then one can write

$$\Delta_s = 4\Delta_b + \Delta_{sc} \qquad (4)$$

where Δ_{sc} is extra energy difference caused by strain and charge transfer.

The energyes E_j and bond lengths d_j, j = 1, 2,3, are calculated by minising Δ_0. If we let $d_I = d_0\ (1 - \delta_0)$ where d_I is the bond length in the ıre AI lattice, then the approximate expression

$$\delta \approx \frac{\delta_0}{1 + \frac{1}{3}\ \frac{B}{B_I} + \frac{8(C_{11} - C_{12})}{27\ B_I} + \text{(small terms due to charge transfer)}} \qquad (5)$$

n be deduced. The bulk modulii B of the host and B_I of the impurity lattices, d a shear coefficient $C_{11} - C_{12}$ of the host appear in the expression. While is expression is approximate, the effects of the various types of strains n be visualized using it. Our detailed calculations are done using the full eory. If one fixes the bond lengths $d_1 = d_2 = d_3 = d_0$, then $\delta = \delta_0$ and the rain energy Δ_{sc} is large and positive. If one allows only bond length distortions ($C_{11} - C_{12} = 0$) then δ is reduced and the bonds tend to adjust so $\approx d_I$. This reduces Δ_{sc} considerably from the undistorted lattice case. $d_I > d_0$ then both bonds are stretche somewhat, $d_1 - d_I > 0$ and $d_2 - d_0 > 0$. wever, this configuration produces large bond angle distortions. When the ear coefficients are turned on the lattice relaxation is modified and the nd lengths cannot adjust as much, so d_1 differs from d_I by a larger amount. e net result is that Δ_{sc} is increased since one pays the price of strain ergy either as a bond length or bond angle distortion.

If we calculate δ from the full theory, then we predict values that are o large. This occurs because the theory predicts $C_{11} - C_{12}$ which agree well th experiment but it predicts bulk modulii with the proper trends from one mpound to another but which are about a factor of 2.4 too small. If we use e experimental values of the strain coefficients in Eq. (5) then good agreement is found with the experiments (5) on $Ga_{1-x}In_xAs$ and $ZnSe_{1-x}Te_x$. The sults are quoted in Table I.

Table I. Bond Lengths in Å for Impurities in the Indicated Host

	Ga in InAs	In in GaAs	Se in ZnTe	Te in ZnSe
periment(5)	2.487	2.587	2.496	2.595
.(5) and perimental B&C	2.499	2.547	2.510	2.570
ll Theory	2.538	2.518	2.570	2.512

loy

In an $A_{1-x}B_xC$ alloy the four cations around a given C anion can be ranged in five different configurations, denoted by $A(4-\eta)$, $B(\eta)$, $\eta = 0,1,$ 3,4. An A(2) B(2) configuration, for example, is one in which the C atom s two A atom and two B atom neighbors. It is possible to solve the full loy problem for large clusters, but for now we have restricted the cluster near neighbors only, and in Fig. 1 the central atom is now an anion C and $= d_3 - d_{eff}$ are taken to be an effective medium bond length that is deterned self-consistently. The different bonds of type 1 no longer have the me length. The procedure is as follows: First assume a value for d_{eff}, y the virtual crystal bond length $(1-x)\ d_0(AC) + x d_0(BC)$. Next calculate e d_1 values for the various configurations by minimizing the energy of the nfiguration. Then, configuration average the various d values to find a w d_{eff}. Then iterate this procedure until it converges.

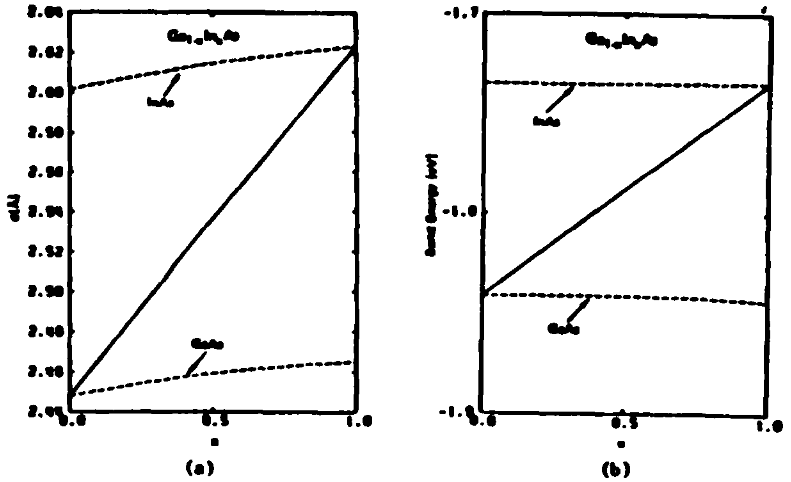
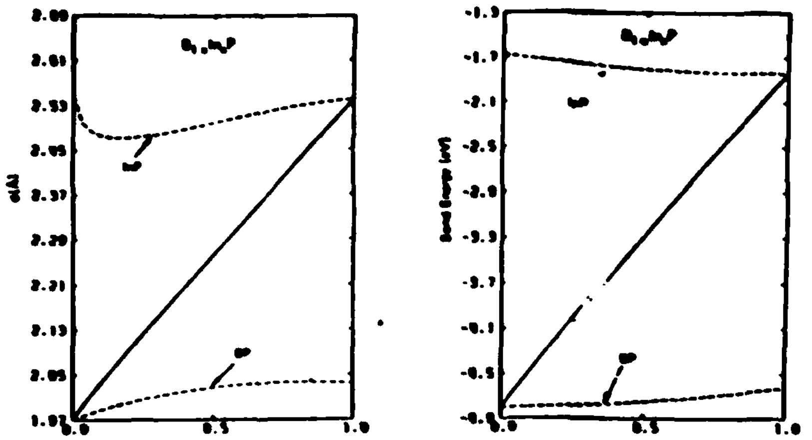


Figure 2: Bond length (a) and energy (b) of $Ga_{1-x}In_xAs$ as functions of x. The dashed curves are the average values for the designated bonds, and the solid curved are the alloy averages.

Because the bond angle distortion terms are unphysically large we have left them out of the present calculation. Hence bond length shifts will be somewhat too small.

Results for the bond energy and bond length shifts with concentration are given for $Ga_{1-x}In_xAs$ in Fig. 2 and $B_{1-x}In_xP$ in Fig. 3 as examples of two different behavior patterns. The predicted trends for GaInAs agree with experiment but the bond length changes are too small as expected with the bond angle distortion terms absent. Notice that the longer bond in this case decreases in length and the shorter one increases as expected. However, in the BInP case the charge shift terms are so large that the longer InP bond has a minimum as a fucntion of composition rather than a monotonic decrease. Because of the large bond length difference between BP and InP there is probably a missability gap in this alloy that prevents these compositions from being prepared. However, one may be able to examine the variation of the anomalous behavior of the InP bond in a BP host (x << 1).

The $Hg_{1-x}Cd_xTe$ system is completely anamolous. The bond lengths of CdTe (2.805 Å) and HgTe (2.797 Å) are nearly the same by an accident. CdTe bonds are dominated by covalent and ionic interactions, while HgTe is more covalent and the metallic terms are important. The mix of interactions in the two cases leads to the same bond lengths. When an alloy is made the charge shift terms dominate, and they cause the longer CdTe bond to become still longer and the HgTe to contract by amounts that are large compared to the pure crystal difference. Moreover, the already weak HgTe has its bond energy reduced still more by the presence of Cd. Since the melting point of HgTe increases as Cd is added and the local strength of HgTe bonds adjacent to a Cd decrease, vacancy densities will increase. All these and other observed phenomena are predicted by the theory. A complete catalog of results requires more pages than we have been allocated in this article, but we have tried to offer a representative group.



(a) (b)

Figure 3: Bond length (a) and energy (b) of $B_{1-x}In_xP$ as functions of x. The dashed curves are the average values for the designated bonds and the solid curves are the alloy averages.

Conclusions

We have demonstrated in agreement with experiment that dislocation energies and hardness of semiconductors are proportional to the shear coefficients which vary roughly as d^{-9}. Thus to decrease dislocations in a given semiconductor one should inhibit their formation by introducing some means of shortening bonds. This should be accomplished without also decreasing the ratio of the vacancy formation energy to the melt temperature. While detailed calculations to support the following contentions are still incomplete, a set of criteria on an impurity (denoted I) in a host semiconductor (denoted H) that are likely to satisfy these conditions are: the bond length of the impurity is smaller than that of the host $d_I < d_H$, the smaller bond energy is stabilized in the alloy min $|\Delta E_{bH}, \Delta E_{bI}| < 0$. In these circumstances, the average bond length will shrink and in the vicinity of each impurity the four surrounding bonds nearly have the length of the impurity b c are stretched slightly. The next neighboring host bonds are also stretched. The net effect is an arrangement which is more rigid than the unperturbed lattice and consequently the local shear strain energy increases, causing dislocation energies to increase. Examples of this case are $B_{1-x}Ga_xAs$, and $Zn_{1-x}Hg_xTe$. Both B in GaAs (6) and Zn in HgTe (7) have proven to be effective in reducing dislocation densities. Unfortunately, B has a low solubility in GaAs and it is not clear that enough can be gotten in to make it easy to prepare dislocation-free material.

A second case where an improvement occurs is if $d_I > d_H$, and again min $|\Delta E_{bH}, \Delta E_{bI}| > 0$. In this case each of the four longer impurity atom bonds are compressed by the surrounding host bonds and, more importantly, the twelve next neighbor host bonds are also compressed. Once again for small impurity concentrations the net dislocation energy should be increased. However, in this case the effect is competing against a net bond lengthening trend of the alloy which is tending to make it less rigid against a shear, so at higher concentrations the mechanism should cease to function. An example of this case is $Ga_{1-x}In_xAs$, where In has proven to be effective in dislocation reduction of GaAs (8).

We have demonstrated that our modification of the Harrison bonding theory accurately predicts the observed change of bond lengths in semiconductor alloys and offers guidance to means for dislocation reduction.

References

1. G. Schoeck and W.A. Tiller, Phil. Mag. 5, 43 (1960).
2. W.A. Harrison, Electronic Structure and Properties of Solids, (Freeman, San Francisco, 1980); R.C. Sokel Thesis, Stanford University, 1978; W.A. Harrison Microscience , Vol. 4 (limited distribution SRI International publication 1983) p. 34.
3. A. Sher, An-Ban Chen, and W.E. Spicer "Effects Influencing the Structure Integrity of Semiconductors and Their Alloys", 1984 Workshop on the Physics and Chemistry of Mercury Cadmium Telluride, San Diego, 1984.
4. A. Sher, An-Ban Chen, and W.E. Spicer, "Dislocations Energies and Hardness of Semiconductors", (submitted to Appl. Phys.Lett.).
5. J.C. Mikkelsen, Jr. and J.B. Boyce, Phys. Rev. Lett. 19, 1412 (1982), and private communications.
6. G. Jacob, J. Crystal Growth 59, 669 (1982); Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. 42, 822 (1983).
7. S.L. Bell and S. Sen, IRIS Detector Specialty Group Meeting, Boulder Colorado (1983); T.W. James and B.F. Zuck, ibid.
8. R.N. Thomas, H.M. Hobgood, D.L. Barrett, G.W. Eldridge, M.M. Sopira, and M.C. Driver, "Large Diameter, Low-Dislocation In Doped GaAs:..." Third Conference on Semi-Insulating III-V Materials, Warm Springs, Oregon 1984.

(a) Effects Influencing the Structural Integrity of Semiconductors and Their Alloys[a]

Arden Sher
SRI International, Menlo Park, California 94025
An-Ban Chen
Auburn University, Auburn, Alabama 36489
W. E. Spicer[b] **and C-K Shih**
Stanford University, Stanford, California 94305

(Received 1 June 1984; accepted 21 August 1984)

The bond length and energy changes of the constituents of alloys relative to their pure crystal values are calculated from an extension of Harrison's method. It is demonstrated that the already weak HgTe bonds are destabilized by adjacent CdTe, HgS, or HgSe, but are stabilized by ZnTe. It is also argued that dislocation energies and the hardness of semiconductors vary as a high inverse power of the bond length of the constituents. Hence, the shorter ZnTe bond as an additive should improve the structural properties of HgTe and CdTe. Experiments that support these predictions are noted. The electronic transport properties of 0.1 eV band gap HgZnTe are about the same as those of HgCdTe, and the structural properties of the Zn compound are superior; thus, we conclude that HgZnTe is likely to be the better material for IR devices.

I. INTRODUCTION

The objective of this work is to understand the microscopic mechanisms that govern the stability of $Hg_{1-x}Cd_xTe$ alloys, and then to suggest changes in the material that improve its strength without adversely affecting its electronic behavior. In pursuit of this goal we have extended Harrison's bonding theory,[1] which is applicable to all tetrahedral structured semiconductors, to calculate bond length and energy changes in an alloy—including charge shift and reconstruction effects—relative to their pure crystal values, vacancy formation energies, dislocation energies, and hardness.

In this paper, we review the experimental situation in several of these areas and compare some of the results with theory for all the group IV, III–V compound, and II–VI compound semiconductors and their alloys. For example, the theory properly predicts the observed inverse ninth power bond length (d^{-9}) dependence of the hardness[2] of semiconductors and, with no adjustable parameters, their correct magnitude. We demonstrate that the weak HgTe bond is destabilized by alloying it with CdTe, HgS, or HgSe; however, the bond is stabilized by ZnTe. Moreover, because the bond length of ZnTe (2.406 Å) is 14% shorter than that of HgTe (2.797 Å) or CdTe (2.804 Å), the dislocation energy per unit length and hardness of the alloys $Hg_{1-x}Zn_xTe$ and $Cd_{1-x}Zn_xTe$ are predicted in agreement with some experiments to be significantly higher than those of the compounds with $x = 0$. Measurements indicate that the electron and hole mobilities of $Hg_{1-x}Zn_xTe$ with $x \sim 0.16$, corresponding to a 0.1 eV band gap, are comparable to those of $Hg_{1-x}Cd_xTe$ for $x = 0.2$.[3] This is to be expected, because the electron effective mass in a narrow-gap material is small; hence, the electron wave functions at Γ_{1C} are distributed over many atoms, with the result that alloy scattering rates are small at the band edge. The hole mobilities are large for a somewhat different reason: The valence band edge is mostly composed of Te p states; however, there is also some cation p state contribution. Because the p state energies of Hg and Zn are nearly the same, the alloy scattering strength at the valence band edge is small. We conclude that, because of its structural advantage, HgZnTe may be superior to HgCdTe for infrared devices

II. ALLOY MODIFICATION OF THE BOND ENERGY AND LENGTH

A. Pure compounds

The first task is to recast Harrison's bond energy and bond length formalism into a structure suitable for generalization to alloys. Focus attention on the nth bond of a pure zinc blende structured compound semiconductor; for the time being, suppress any notation identifying it. Then, in a II–VI compound, the bond energy is[1]:

$$E_b = \frac{\epsilon_{p+} - \epsilon_{p-}}{2} + |V_{1+} + V_{1-}| - 2(V_2^2 + V_3^2)^{1/2} + 2\sum_{i=0}^{3}(\epsilon_{m-} + \epsilon_{m+}) + \frac{\gamma}{d^4}, \tag{1}$$

where the first term is the energy per bond needed to transfer two electrons from a p state on the group VI atom (the anion) to a p state on the group II atom (the cation), so that both start with four electrons (ultimately in the final bonding arrangement there is a net electron transfer from the cation to the anion), the second term is the promotion energy per bond to form sp^3 hybrids on both atoms, the third term is the bond formation energy owing to the covalent and ionic terms, the fourth term is Harrison's metallization energy, and the fifth term is the repulsive overlap energy that prevents the bonds from collapsing. The various symbols are defined below:

$\epsilon_{p|\pm|}$ is the $\left\{\begin{smallmatrix}\text{cation}\\\text{anion}\end{smallmatrix}\right\}$ p state correlated term values,[4]
$V_{1|\pm|} \equiv (\epsilon_{p|\pm|} - \epsilon_{s|\pm|})/4$ are hopping integrals between two adjacent bonds coupled through a $\left\{\begin{smallmatrix}\text{cation}\\\text{anion}\end{smallmatrix}\right\}$,
$\epsilon_{s|\pm|}$ are the $\left\{\begin{smallmatrix}\text{cation}\\\text{anion}\end{smallmatrix}\right\}$ s state correlated term values,
$V_2 \equiv -24.5/d^2$ (eV) is the covalent energy,
d is the bond length,

$$V_3 \equiv (\epsilon_+ - \epsilon_-)/2,$$

$\epsilon_{(\pm)} \equiv (\epsilon_{s(\pm)} + 3\epsilon_{p(\pm)})/4$ are the sp^3 hybrid energies,

$$\epsilon^o_{m(\pm)} \equiv \sum_{j=1}^{3} (U^o_{b(\pm)} U^j_{a(\pm)} V^2_{1(\pm)})/(\epsilon^o_b - \epsilon^j_a)$$

is the contribution to the reduction of the bonding state energy ϵ^o_b calculated by second-order perturbation theory of the bond in question (denoted by superscript o), owing to its interaction with the three neighboring ($i = 1$–3) antibonding states sharing the same $\left\{ {\text{cation} \atop \text{anion}} \right\}$

$$U_{b+} = \sqrt{(1 - \alpha_p)/2} = U_{a-},$$

$$U_{b-} = \sqrt{(1 + \alpha_p/2} = U_{a+},$$

are the probability amplitudes of finding the cation and anion sp^3-hybrid contributions to the bonding and antibonding states, e.g., U_{b+} is the probability amplitude of finding the cation sp^3 hybrid in the bonding state, etc.,

$\alpha_p \equiv V_3/\sqrt{V_2^2 + V_3^2}$ is the polarity,

$\alpha_c \equiv |V_2|/\sqrt{V_2^2 + V_3^2}$ is the covalency,

$$\epsilon_{\left({b \atop a}\right)} = \frac{\epsilon_+ + \epsilon_-}{2} (\mp)\sqrt{V_2^2 + V_3^2},$$

are the bonding and antibonding one electron energy states owing only to covalent and ionic interactions,

$$\epsilon^i_{m(\pm)} \equiv \frac{(U^i_{b(\pm)} U^o_{a(\pm)} V_{1(\pm)})^2}{\epsilon^i_b - \epsilon^o_a}, \quad i = 1,2,3$$

are the contributions to the bond in question arising from its antibonding state interactions with a neighboring bonding state i, and γ is an adjustable coefficient chosen so the measured bond lengths are reproduced.

Harrison has written Eq. (1) without explicit reference to the surrounding bonds. Equation (1) in its present form is rather easily used for alloy calculations by replacing host atom terms by one appropriate to the alloy.

The parameter γ for each compound is fixed by the condition that there is no net force on atoms at their equilibrium lattice positions. This condition is satisfied if the gradient of the total energy with respect to all the bond lengths vanishes:

$$\nabla_{\{d_n\}} E_T = 0, \quad \text{where } E_T = \sum_n E^n_b.$$

The condition is equivalent to the requirement that the net bond tension $T_{on} + T_{mn}$ for the nth bond equals the repulsive force

$$T_{on} + T_{mn} - \frac{4\gamma_n}{d^5_n} = 0, \tag{2}$$

where

$$T_{on} \equiv \frac{\partial}{\partial d_n}(-2\sqrt{V_2^2 + V_3^2}) = 4\alpha_c\sqrt{V_2^2 + V_3^2}/d, \tag{3}$$

$$T_{mn} \equiv 2\frac{\partial}{\partial d_n}\sum_{i=0}^{3}(\epsilon^i_{m+,n} + \epsilon^i_{m-,n}),$$

$$\equiv \sum_{i=0}^{3}(T^i_{m+,n} + T^i_{m-,n}). \tag{4}$$

TABLE I. Correlated term values and bond energies for three II–VI compounds. The experimental bond lengths, polarity, and metallic contribution to the bond energy are also plotted.

	Correlated term values (eV)		
Atom	s	p	sp^3
Zn	− 10.224	− 4.920	− 6.246
Cd	− 9.611	− 4.784	− 5.991
Hg	− 10.946	− 4.872	− 6.391
Te	− 19.620	− 9.824	− 12.273

			Bond energies	
Compound	d (Å)[a]	α_p	E_{metallic} (eV)	E_b (eV)
ZnTe	2.643	0.652	− 0.625	− 0.985
CdTe	2.805	0.710	− 0.466	− 0.944
HgTe	2.797	0.685	− 0.614	− 0.480

[a] Reference 12.

The last identity in Eq. (4) defines the indicated bond tensions. Collecting all the results from the earlier paragraphs, one finds (dropping the bond designation n):

$$T^o_{m\pm} = -\frac{4}{d}\left[(1 + \alpha_p)\alpha_p + \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^o_b - \epsilon^1_a}\right]\epsilon^o_{m\pm}, \tag{5a}$$

$$T^1_{m\pm} = T^2_{m\pm} = T^3_{m\pm}$$

$$= \frac{4}{d}\left[(1 - \alpha_p)\alpha_p - \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^1_b - \epsilon^o_a}\right]\epsilon^1_{m\pm}, \tag{5b}$$

$$T^o_{m\mp} = \frac{4}{d}\left[(1 - \alpha_p)\alpha_p - \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^o_b - \epsilon^1_a}\right]\epsilon^o_{m\mp}, \tag{5c}$$

$$T^1_{m\mp} = T^2_{m\mp} = T^3_{m\mp}$$

$$= -\frac{4}{d}\left[(1 + \alpha_p)\alpha_p + \frac{\alpha_c^2\sqrt{V_2^2 + V_3^2}}{\epsilon^1_b - \epsilon^o_a}\right]\epsilon^1_{m\mp}. \tag{5d}$$

Finally combining Eqs. (1)–(5), the bond energy is

$$E_b = -1(1 + \alpha_p^2)(V_2^2 + V_3^2)^{1/2}$$

$$- \frac{9}{8}(1 - \alpha_p^2)^2\frac{(V^2_{1+} + V^2_{1-})}{(V_2^2 + V_3^2)^{1/2}}$$

$$+ |V_{1+} + V_{1-}| + \frac{\epsilon_{p+} - \epsilon_{p-}}{2}. \tag{6}$$

Equation (6) was first derived by Harrison using a much simpler argument.[1] The bond energies quoted in Table I and on Fig. 5 are calculated from this expression, using as input the correlated term values from Table I and the experimental bond lengths.

B. Alloy calculations

There are several layers of sophistication that can be used to compute the bond energy and length changes between pure compounds and their alloys. However, all the calculations are motivated by the EXAFS experimental result found first for $Ga_{1-x}In_xAs$ and later in other compounds.[5] It was found that, while the average bond length in the alloy follows Vegard's rule and varies linearly between those of the pure constituents, the individual bond lengths of GaAs and

InAs are nearly the same as those of the pure materials—changing by only about 4%. The shorter GaAs bond increases its length when it is an impurity in an InAs host by about 4%, and the longer InAs bond shrinks by about 4% in the opposite extreme where it is the impurity in GaAs. The lattice fits together by having the metal sublattice nearly retain its fcc structure with a lattice constant following Vegard's rule, and each As atom adjusts its local position and accommodates to which metal atoms occupy its four neighboring sites. Thus there are five local arrangements; in the first an As atom is surrounded by four Ga atoms [a Ga(4)In(*o*) configuration], in the second there are three Ga atoms and one In [a Ga(3)In(1) configuration], and the last three configurations are Ga(2)In(2), Ga(1)In(3), and Ga(*o*)In(4). In a general cation $A_{1-x}B_xC$ alloy or anion $CA_{1-x}B_x$ alloy, the generalizations are obvious, and one can examine $A(4-\eta)B(\eta); \eta = 1,2,3,4$ configurations.

1. Method 1

The lowest level of approximation that takes account of these findings calculates [from Eqs. (1) and (2)] a given bond's energy and length shift when its surrounding bonds have different constituents. The calculation is done assuming that the surrounding bonds retain their pure crystal bond lengths. This model is somewhat unrealistic, in that it ignores the strains that must be present to allow the lattice to fit together. However, it is nonetheless useful because it treats the largest terms, viz. those arising from charge transfer between the substituted species in the alloy, and results in analytic expressions, while the successively more complicated treatments must be done numerically. It is also useful later because, taken together with the more complete calculations, it allows one to judge the relative contribution from different physical effects. From Eq. (1) in a cation-substituted alloy one can deduce an expression for the energy shift of an AC bond when η surrounding bonds are replaced by BC, ΔE^{η}_{b} as

$$\Delta E^{1}_{b} = \frac{2[U^{o}_{b-}U^{1}_{a-}(\mathrm{BC})V_{1-}]^2}{\epsilon^{o}_{b}-\epsilon^{1}_{a}(\mathrm{BC})} - \frac{2[U^{o}_{b-}U^{1}_{a-}(\mathrm{AC})V_{1-}]^2}{\epsilon^{o}_{b}-\epsilon^{1}_{a}(\mathrm{AC})} + \frac{d\partial}{4\partial d}2\left\{\frac{[U^{o}_{b-}U^{1}_{a-}(\mathrm{BC})V_{1-}]^2}{\epsilon^{o}_{b}-\epsilon^{1}_{b}(\mathrm{BC})} - \frac{[U^{o}_{b-}U^{1}_{a-}(\mathrm{AC})V_{1-}]^2}{\epsilon^{o}_{b}-\epsilon^{1}_{a}(\mathrm{AC})} + \frac{[U^{1}_{b-}(\mathrm{BC})U^{o}_{a-}V_{1-}]^2}{\epsilon^{1}_{b}(\mathrm{BC})-\epsilon^{o}_{a}} - \frac{[U^{1}_{b-}(\mathrm{AC})U^{o}_{a-}V_{1-}]^2}{\epsilon^{1}_{b}(\mathrm{AC})-\epsilon^{o}_{a}}\right\}, \tag{7}$$

where the partial derivative is taken with respect to the central bond length (a term designated with a superscript *o*) and in this simple approximation

$$\Delta E^{\eta}_{b} = \eta E^{1}_{b}, \quad \eta = 0,1,2,3. \tag{8}$$

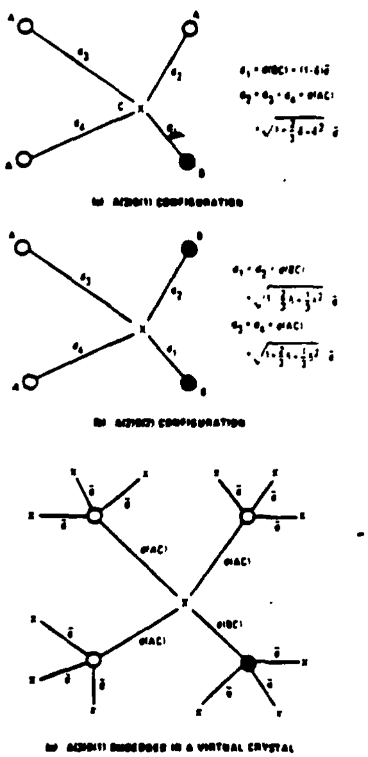


FIG. 1. Schematic view of various bond configurations. (a) A(3)B(1) configuration, (b) A(2)B(2) configuration, and (c) A(3)B(1) configuration embedded in a virtual crystal.

The bond length shifts are found from Eq. (2). The modification of these expressions for an anion-substituted alloy follows from symmetry.

2. Method 2

A more realistic calculation includes strain and permits the lattice to fit together. The simplest non-self-consistent, near-neighbor version of this calculation proceeds as follows. In a cation-substituted alloy $A_{1-x}B_xC$ (e.g., $Hg_{1-x}Cd_xTe$ or $Zn_{1-x}Hg_xTe$), the cations are assumed to occupy their regular fcc sublattice sites, while the anions accommodate to their local configuration. The average lattice constant is taken to follow Vegard's rule:

$$\bar{d} = (1-x)d_0(\mathrm{AC}) + xd_0(\mathrm{BC}), \tag{9}$$

where d_0 and $d_0(\mathrm{BC})$ are the pure crystal lattice constants. Then in an A(3)B(1) configuration [see Fig. 1(a)], the anion will be displaced along the BC bond. For definiteness, suppose $d(\mathrm{AC}) > d(\mathrm{BC})$; then the anion will be shifted away from the tetrahedral site toward the B atom. Because $d_0(\mathrm{AC}) > \bar{d} > d_0(\mathrm{BC})$, the $d(\mathrm{BC})$ bond will be elongated and the $d(\mathrm{AC})$ bonds will also be locally elongated in a strained configuration. We can for this case [see Fig. 1(a)] write

$$d(\mathrm{BC}) = (1-\delta)\bar{d} \tag{10}$$

and for the three AC bonds

$$d(\mathrm{AC}) = \sqrt{1+\frac{2\delta}{3}+\delta^2}\,\bar{d}. \tag{11}$$

In an A(2)B(2) configuration [see Fig. 1(b)], we have for both BC bonds

$$d(\mathrm{BC}) = \sqrt{1 - \frac{2}{3}\delta + \frac{1}{3}\delta^2}\,\bar{d}, \tag{12}$$

and for both AC bonds

$$d(\mathrm{AC}) = \sqrt{1 + \frac{2}{3}\delta + \frac{1}{3}\delta^2}\,\bar{d}. \tag{13}$$

The values of δ are determined by the condition $\partial E_T/\partial\delta = 0$, or

$$\sum_{n=0}^{3}\left(T_{on} + T_{mn} - \frac{4\gamma_n}{d_n^5}\right)\frac{\partial d_n}{\partial \delta} = 0. \tag{14}$$

Equation (14) is solved by

(1) Assigning the γ_ns their pure crystal values.

(2) Supposing that the four tetrahedrally configured anions are embedded in the virtual crystal [see Fig. 1(c)].

(3) For each x, δ is varied and the bond tensions are calculated and inserted into Eq. (14) until a value of δ is found that satisfies this equation.

The computer outputs are δ and the AC and BC bond energies. This method can be extended to take explicit account of the bond length variations of larger clusters, and $\bar{d}$ can be calculated explicitly (instead of using Vegard's rule) so the results can be made to be self-consistent. Also, effects arising from bond angle distortions can be included. The major effects that are currently included in the calculation arise from the bond length dependence of V_2, and charge shifts driven by hybrid energy differences and coupled through the metallization terms. However, V_2 is also sensitive to bond angle distortions, and these effects are not included in the present results. We expect their inclusion will modify the quantitative results by 10%–30%, but not the trends.

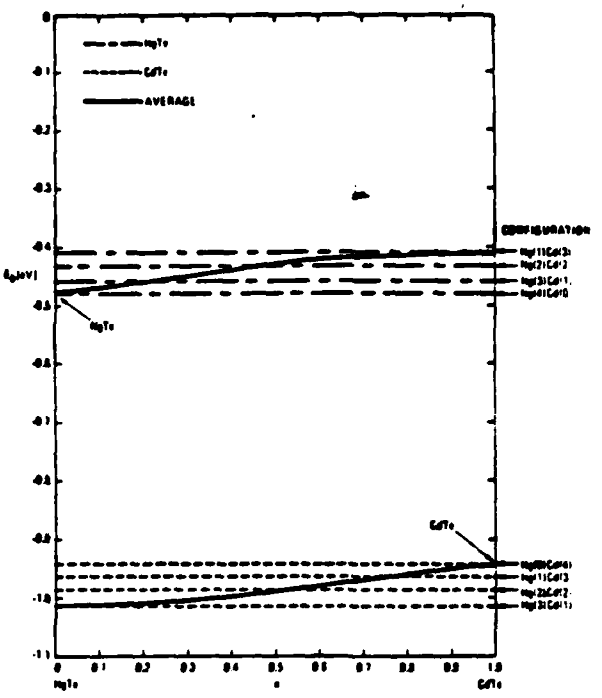


FIG. 3. Bond energy as a function of concentration x for the $Hg_{1-x}Cd_xTe$ alloy in different configurations. The short dashed lines are for the CdTe bonds and the long–short dashed lines represent the behavior of the HgTe bonds. The solid lines are the concentration weighted averages of the respective bond energies.

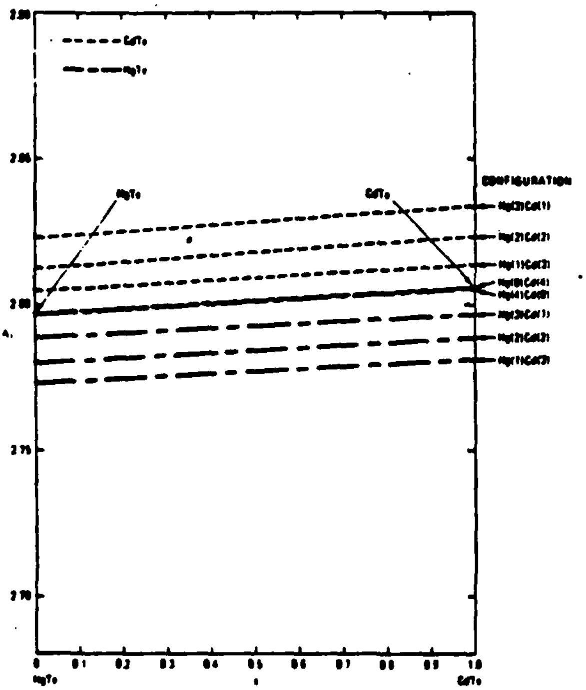


FIG. 2. Bond lengths as a function of concentration x for the $Hg_{1-x}Cd_xTe$ alloy in different configurations. The short dashed lines are for the CdTe bonds and the long–short dashed lines represent the behavior of the HgTe bonds.

C. Results

The correlated atomic term values[4] for the elements in the more common II–VI compounds are listed in Table I, along with the bond energies of their nine compounds. The valence s state energies for Zn are large, become smaller as for Cd, and then (in an unusual occurrence) the trend reverses and the Hg s levels are deeper again. This is caused by the relativistic terms that become important in Hg. The p state levels have a similar trend; however, the differences among the atoms are much smaller. The levels of the anions are deeper than those of the cations. Thus, there is a net electron transfer from the cations to the anions, which is responsible for the ionic contribution to the bonding. The polarity, also listed in Table I, reflects the relative contribution of ionic and covalent character to the bonds. Notice that α_p is small for ZnTe larger for CdTe, and smaller again for HgTe. By contrast, the metallic contribution E_m is large for ZnTe and HgTe and small for CdTe. In fact, HgTe would not be bound if it were not for the metallic contribution to the bond, not a terribly surprising result, because HgTe is a semimetal. The bond lengths of CdTe and HgTe are very nearly the same, but this is an accident resulting from the balance of contributions to their bonds: CdTe is dominated by the covalent and ionic contributions to its bond, while HgTe has bonds that are mostly covalent and metallic. We should therefore not be too surprised if their alloy displays some unusual bond length variations.

Figures 2 and 3 illustrate the $Hg_{1-x}Cd_xTe$ alloy concentration variations of the length and energy for the five configurations. These results were calculated using the second method discussed in Sec. II B. In Fig. 2 we see the unexpected result that the shorter HgTe bond becomes still shorter in the alloy, while the CdTe bond length increases. Moreover, the changes are large compared to the original length difference. Thus, one can expect to find local microstrains, even in this nearly lattice-matched material. The bond energy variations shown in Fig. 3 also display an important result. The already weak HgTe bond is destabilized by the presence of Cd. The flat concentration variation of the bond lengths and energies is peculiar to HgCdTe; much more structured behaviors are exhibited by other compounds.

Several considerations to keep in mind when dealing with alloys are depicted schematically in Fig. 4 for a Hg(3)Cd(1) configuration. Due to the charge shift alone, the HgTe bond shrinks by 0.030 Å, the minimum in the configuration diagram moves to a smaller bond length and the depth of the well is decreased. Then the bond is stretched by 0.011 Å to its final length. Thus, the bond is not centered at a local minimum. For example, this implies that the force constant of local phonon modes will be related to the second derivative of this configuration diagram evaluated at the strained position away from the well minimum. This effect is even more pronounced for materials with a larger lattice mismatch. Notice that the CdTe bond is also stretched relative to its local minimum in this configuration.

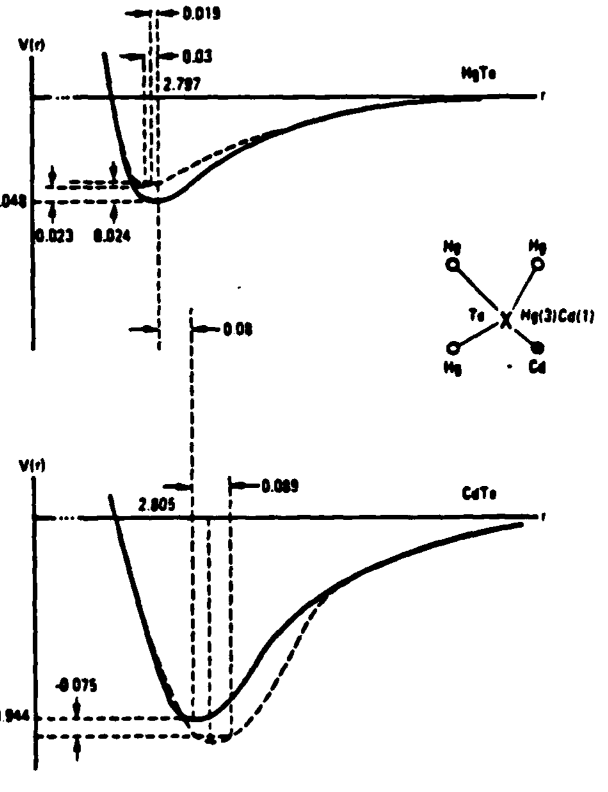


FIG. 4. Schematic picture of a potential configuration diagram for the HgTe and CdTe bonds in an Hg(3)Cd(1) configuration. For the HgTe bond the energy and position of the bond center are calculated by method 1 of Sec. II B and the final stretched position and energy by method 2. The results quoted for the CdTe bond are those determined by method 2.

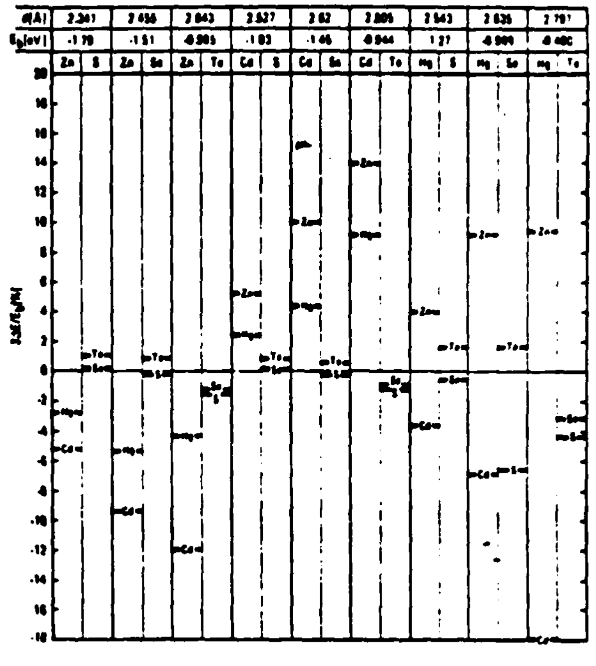


FIG. 5. AC bond energy shifts for the II–VI compounds in A(1)B(3) configurations expressed as a percentage of the pure crystal bond energy. The B atom is designated in the line that shows the shift. The pure crystal AC bond energies and lengths are given at the top of the figure. These numbers were calculated by method 1 in Sec. II B.

The result for bond energy variations for all the II–VI compound alloys is gathered in Fig. 5. Here the energy shifts of an AC bond in a A(1)B(3) configuration $\Delta E^3_b = 3\Delta E^1_b$ are presented as a percentage of the pure material bond energy for cation substitute alloys along with equivalent results for anion substantial alloys, e.g., $HgS_{1-x}Te_x$. The energy shifts are calculated following the first method presented in Sec. II B. We have not had the opportunity to run all the compounds following the more sophisticated calculational procedure. The bond at the top of a column is the one whose energy shift is calculated, and the other constituent is identified in the line that indicates the energy shift. A negative $3\Delta E_b/E_b$ is destabilizing and a positive one stabilizing. Focus on HgTe. It is destabilized by CdTe, HgS, and HgSe, but is stabilized by ZnTe. The CdTe destabilizes the HgTe bond by 18% according to the simple calculation and by 15% according to the better one. The difference between Zn and Cd occurs because in the Cd case there is a net electron transfer from the Cd to the Hg, while in the case of Zn there is practically no such transfer. These extra transferred electrons must occupy antibonding states, because the bonding states on the HgTe are full, so the net effect (including other energy state shifts) is to destabilize the HgTe bond. This predicted destabilization of the HgTe bond in HgCdTe alloys

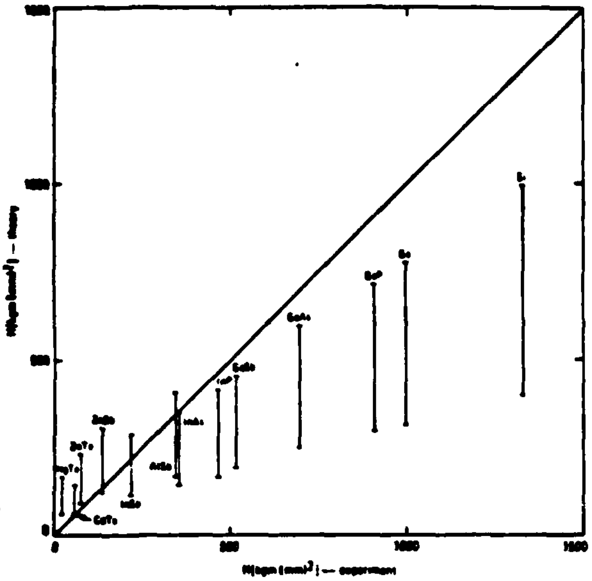


FIG. 6. Theoretical vs experimental hardness. The two values at the top and bottom of the vertical lines designated for each compound were calculated for two different possible dislocation arrangements. The lower one corresponds to the minimum energy configuration and the upper one to a likely higher energy configuration. The experimental numbers were taken from Ref. 2.

agrees with the *d* state energy shift reported by Spicer *et al.* at last year's workshop,[6] and by the oxygen uptake experiments presented by the same group at this conference.[7]

III. DISLOCATION ENERGIES AND HARDNESS

The dislocation energies and hardness of semiconductor are calculated by an extension of Harrison's method.[1] It is demonstrated, in agreement with experiment, that dislocation energies per unit length, depending on the covalency, are proportional to d^{-9}, where *d* is the bond length and hardness is proportional to d^{-5} to d^{-11}. The low powers are for pure covalent materials. The hardness is related to the interaction energies among dislocations. The detailed theory will be published elsewhere.[8]

Figure 6 shows the theoretical hardness of several semiconductors plotted against experimental values.[2] The theory is calculated for two different dislocation configurations caused by the extra material pushed by the indentor into the crystal. The lower end of each vertical line is a realistic hardness for the minimum energy configuration, and the upper ends of the lines are the values found for a higher-energy configuration. In most cases, one would expect real configurations to have hardnesses lying somewhere between these two values, once the effect of heat generated by the indention process is subtracted from the experimental numbers. This theory has no adjustable parameters, yet it predicts the right trends and magnitudes.

IV. NATIVE IMPERFECTIONS GROWN INTO ALLOYS

There are two kinds of imperfections we shall discuss: vacancies and dislocations. At present, the conclusions are tentative, because the detailed theory is still incomplete. Vacancy formations energies and the melting temperature are both proportional to the bond energies of the constituents. Thus, one would expect the vacancy density just below the melting point to be about the same for all materials with a similar coordination number; in our semiconductor case, the coordination number is four. Dislocation energies are mostly owing to long-range bond angle distortions[9] and accordingly are sensitive to the sheer coefficients, which vary roughly as d^{-7} for many compounds.[1] Based on this general argument, one would expect to find fewer dislocations in materials with short bond lengths.

Dislocations appear in bulk grown materials through a sequence of steps.[10] We shall only discuss the simplest case, where there are no mechanical stresses on the growing material. In the temperature gradient behind the growth front, there is usually a supersaturated vacancy content. These excess vacancies may diffuse to surfaces where they do no harm, or they may condense into vacancy clusters that can then organize into dislocation loops. If a loop is not parallel to the growth front, it can climb toward the front by absorbing more vacancies. If it climbs faster than the growth progresses, then the loop will reach the front. It can then spread and grow along with the crystal. Observed from the growth surface, such a loop will appear to be two edge dislocations. If this dislocation formation process can be inhibited in any of its steps, the resulting material will have fewer dislocations.

Zn in $Zn_{1-x}Cd_xTe$ or $Zn_{1-x}Hg_xTe$ will serve this function in two ways. The first is that Zn stabilizes the weaker bond in these alloys; accordingly, vacancy formation energies in the alloy should be larger than those in the pure constituent CdTe or HgTe. The melting temperature will also be raised somewhat; which effect "wins" is a matter still to be decided. The second effect is clear cut: The shorter bond lengths of the ZnTe will introduce stiff struts into the system that will increase the dislocation energies and thus should inhibit dislocation loop formation and subsequent climb. The large dislocation density reduction found by Bell and Sen in $Zn_{1-x}Cd_xTe$ with only 4% Zn supports this speculation.[11].

V. CONCLUSIONS

We have argued that, in an alloy, the bond energies and bond lengths are modified from their pure crystal values. Bond energies can be strengthened or weakened in an alloy depending on the relative sp^3 state energies and the consequent net electron shift between the constituents. For most compounds, the average bond length modifications for each type of bond are generally small compared to those suggested by Vegard's rule, in agreement with experiment. However, HgCdTe is an exceptional case, because the bond lengths for HgTe and CdTe are nearly the same (by accident). The bond length and energy shifts also depend sensitively on the local configuration and the bonds are locally strained. These effects must be taken into consideration in a proper theory of alloy vacancy formation energies, phonon frequencies, dislocation energies, etc.

The conclusions most relevant to the initial question posed in the introduction are that for the narrow gap allows:

Cd destabilizes the already weak HgTe bond and has little effect on the dislocation energy.

Zn both stabilizes the HgTe bond and increases the dislocation energy.

Zn also decreases the dislocation density of CdTe to improve substrates.

HgZnTe electron and hole mobilities are about the same as those of comparable band gap HgCdTe.

HgZnTe is likely to be a superior material to HgCdTe for detectors operating in the 8–12 μm regime.

---

[a] This work was supported in part by AFOSR Contract No. F496280-81-K-0012 and by DARPA Contract No. MDA-903-83-0-0108.

[b] Stanford Ascherman Professor of Engineering.

[1] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); R. C. Sokel, thesis, Stanford University, 1978; W. A. Harrison, *Microscience* (SRI International, Menlo Park, California, 1983), Vol. IV, p. 34 limited-distribution.

[2] N. A. Goyunova, A. S. Borshchevskii, and D. N. Fretaikov, *Semiconductors and Semimetals*, edited by R. L. Willardson and A. C. Beer (Academic, New York, 1958), Vol. IV, Chap. I.

[3] O. Mukhina, A. M. Gulyaev, A.S. Shnitnikov, and I. V. Yamshchikova, Tr. Mosk Energ. Inst. 443, 16 (1980).

[4] A. B. Chen, A. Sher, and W. E. Spicer, J. Vac. Sci. Technol. A 1, 1674 (1983).

[5] J. C. Mikkelsen, Jr. and J. B. Boyce, Phys. Rev. Lett. 19, 1412 (1982), and private communications.

[6] W. E. Spicer, J. A. Silberman, I. Landau, A. B. Chen, A. Sher, and J. A. Wilson, *J. Vac. Sci. Technol.* A 1, 1735 (1983).

[7] J. A. Silberman, D. Laser, I. Landau, and W. E. Spicer (these proceedings).

[8] A. Sher, A. B. Chen, and W. E. Spicer, Appl. Phys. Lett. 46, 54 (1985)

[9] J. P. Hirth and J. Lothe, *Theory of Dislocations*, 2nd ed. (Wiley, New York, 1982).

[10] G. Schoeck and W. A. Tiller, Philos. Mag. 5, 43 (1960).

[11] S. L. Bell and M. Sen (these proceedings).

[12] R. Zallen, *Handbook of Semiconductors*, edited by T. S. Moss and W. Paul (North-Holland, Amsterdam, 1982), Vol. I, Chap. 1.

PHYSICAL REVIEW B VOLUME 31, NUMBER 10 15 MAY 1985

Sensitivity of defect energy levels to host band structures and impurity potentials in CdTe

A.-B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849

A. Sher
SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

(Received 10 December 1984)

The sensitivity of defect energy levels in semiconductors to the host band structures and impurity potentials has been studied for approximately 30 impurities in CdTe using four different band-structure models. The discrepancies in the defect levels between two different sets of band structures and impurity potentials are found to range from less than 0.1 eV to the whole band gap (1.6 eV). The band-structure effects are analyzed here in terms of detailed partial densities of states. Examples of contradictory predictions from different band structures are illustrated, and ways to improve the theory are suggested.

I. INTRODUCTION

In several of our recent papers,[1-7] we have applied a method to calculate the band structure of semiconductors that is both efficient and accurate. Because the procedure involves casting the basis functions into orthonormal local orbitals[6] (OLO), our method has the advantages common to empirical tight-binding (ETB) calculations,[8-10] except that the Hamiltonian matrix elements to all ranges are retained. The inclusion of these higher coefficients makes it possible to produce excellent band structures including conduction bands and effective masses. The method also yields wave functions for optical property calculations.[7] Moreover, its OLO description also permits its extension, through the coherent-potential approximation, to alloys.[2-5]

The recent attention focused on defects in semiconductors motivated us to apply our method to this problem. The theories of defects have ranged from very sophisticated self-consistent density-functional theory[11-13] (SCDF) to simple ETB calculations. It is generally recognized that SCDF is as accurate in defects for the ground-state properties as it is for pure semiconductors, but less certain in assigning excited energy levels. ETB, because it can produce results for many systems in one study, claims to predict the trends of deep levels[10] even if the accuracy for a given impurity may be poor. However, this contention remains to be verified.

To assess this concern, we ask the following question: "How sensitive are defect levels to host band structures and impurity potentials?" To this end, we have adopted the simple yet nontrivial defect model, that of site-diagonal substitutional defects often used in ETB studies. CdTe was selected in this study because its band structure has been examined in great detail by us, and there are three published band-structure models[8-10] that we could easily generate for comparison. There is also a considerable body of experimental data on deep states in this system.[14-17]

II. CALCULATIONAL PROCEDURE

In the simple site-diagonal substitutional defect model, the impurity energy levels E are determined by the equation

$$1-v_\alpha g_\alpha(E)=0\ , \tag{1}$$

where α designates the symmetry of a local state, e.g., Γ_6, Γ_7, and Γ_8 on an atomic site in the zinc-blende structure, and g_α is the real part of the diagonal matrix element of the host-crystal Green function. g_α can be calculated from the partial density of states (PDOS) by

$$g_\alpha(E)=\int \rho_\alpha(\epsilon)/(E-\epsilon)d\epsilon\ . \tag{2}$$

The PDOS is given by

$$\rho_\alpha(\epsilon)=\sum_{n,k}|a_n^\alpha(k)|^2\delta(\epsilon-\epsilon_n(k))\ , \tag{3}$$

where $\epsilon_n(k)$ are band energies and $a_n^\alpha(k)$ are the probability amplitudes of the band state in the Bloch basis constructed from the OLO labeled by α. The Brillouin-zone integration in Eq. (3) is calculated using an accurate ray scheme.[18]

Because a principal concern of this paper is the sensitivity of impurity levels to the host band structures, we should emphasize the difference between our method and ETB. Our method consists of four steps.

(1) We start with four Gaussian orbitals per atom and empirical pseudopotentials,[19] and compute the Hamiltonian matrix $H(k)$ and overlap matrix $S(k)$ as was done by Kane[20] and Chadi.[21]

(2) The Gaussian orbitals are transformed into OLO,[6] so $H(k)$ is transformed into $H_0(k)$ and S into the identity matrix. The band structures calculated from $H_0(k)$ are accurate to 5% as compared to more sophisticated methods using the same potential.[1]

(3) A spin-orbit Hamiltonian in the OLO basis[4] is incorporated to deal with this interaction.

(4) To compensate for the effects of truncated basis and

nonlocal potentials, a perturbation Hamiltonian H_1 is added. H_1 has the same form as a truncated ETB Hamiltonian. The parameters in H_1 are adjusted to fine tune the important band energies and effective masses.[1-4]

Although both ETB and our methods are empirical, there are two major differences.

(1) While most ETB retains the H matrix elements only to the first- or second-neighbor shell, ours extends to all ranges, so that the high Fourier components needed to produce the sharp band curvatures are properly given.

(2) Our method can directly generate wave functions for calculation of other properties.

Thus, while our method yields more accurate band structures, it retains much of the advantage of ETB, namely the computational speed and a simple direct-space description of the Hamiltonian.

III. BAND STRUCTURES AND PARTIAL DENSITIES OF STATES

Figure 1 depicts the four band structures to be considered for CdTe. Our result is in panel (a); panels (b) (Ref. 8) and (c) (Ref. 9) are two ETB band structures with the Hamiltonian matrix elements truncated at second neighbors. (Because different parameters were selected, these two band structures are not identical.) Panel (d) (Ref. 10) results from the use of five basis orbitals per atom; the extra one is an excited s state. All these band structures are adjusted to have the proper fundamental band gap of 1.6 eV. The principal differences one sees on first inspection are in the band curvatures, especially the conduction bands. The effective mass at the bottom of the conduction band in panel (a) is 0.1 times the free-electron mass, in agreement with experiment,[17] while in other panels it is more than twice as large.

Figure 2 shows the densities of states (DOS) for each of the band structures in Fig. 1. While the valence bands at least exhibit general common features, the conduction bands are almost unrecognizable as representing the same compound. In panels (c) and (d), for example, there is a second band gap above the fundamental gap. Also note that there are two extra narrow peaks associated with the two extra excited s orbitals (one for Cd and the other for Te) included in the calculation.

To analyze the band effects on defect levels [see Eqs. (1) and (2)], the DOS is further decomposed into partial densities of states for $\Gamma_6(s)$, $\Gamma_7(p^{1/2})$, and $\Gamma_8(p^{3/2})$ states on the Cd and Te sites, as shown in Figs. 3–6. The Γ_8 PDOS are not shown because they are nearly the same as Γ_7 with only a slight upward energy shift. These PDOS show how the "atomic" levels evolve into band states. These curves contain useful information about many properties, e.g., the relation between the crystal bonding and atomic energies, and how potential disorder in alloys affects different parts of the bands,[2-5] in addition to defect levels studied here.

The Γ_6(Cd) PDOS shown in Fig. 3 split between the conduction and valence bands. It is generally assumed that the cation s states in III-V and II-VI compounds

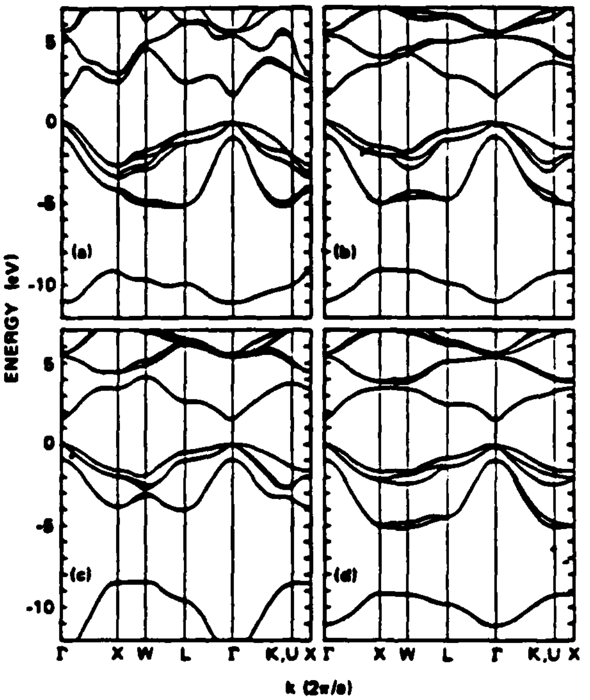


FIG. 1. Four band structures of CdTe used for comparative studies: (a) present work, (b) Ref. 8, (c) Ref. 9, and (d) Ref. 10.

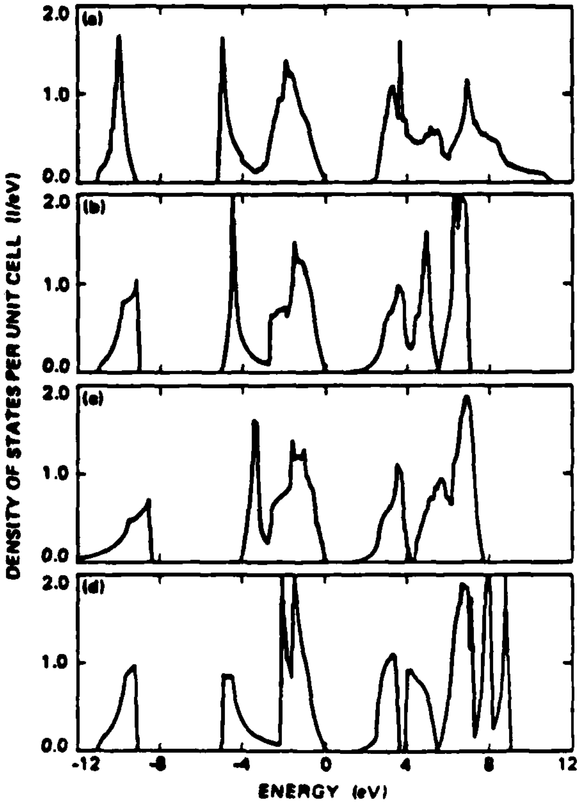


FIG. 2. Densities of states calculated from the four band structures in Fig. 1.

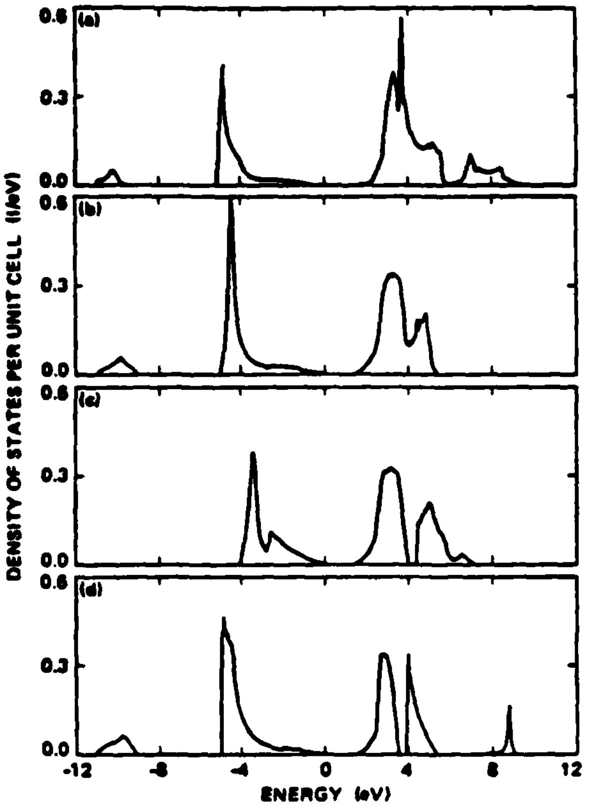


FIG. 3. The Cd Γ_6 partial densities of states.

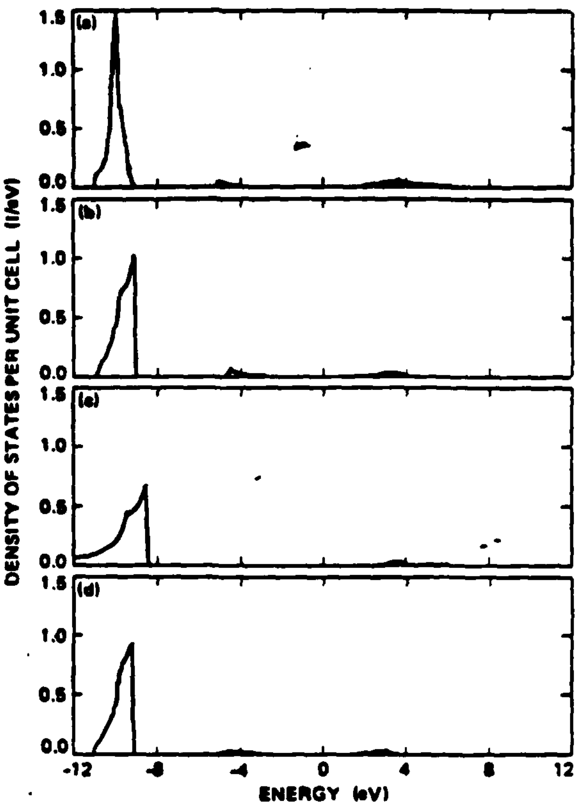


FIG. 5. The Te Γ_6 partial densities of states.

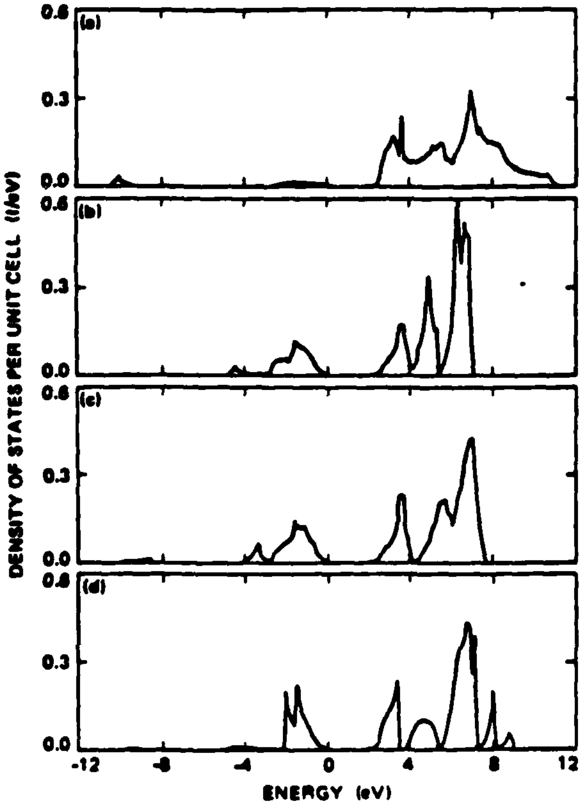


FIG. 4. The Cd Γ_7 partial densities of states.

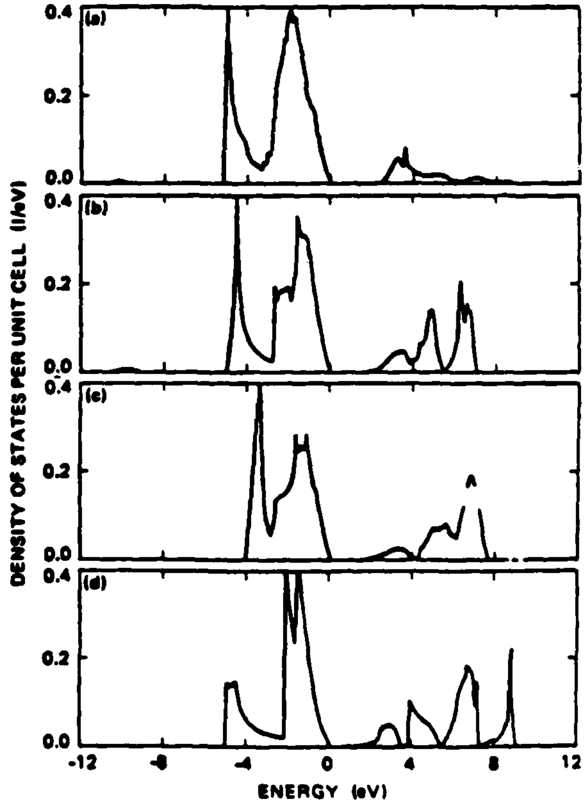


FIG. 6. The Te Γ_7 partial densities of states.

evolve into the conduction bands, while the anion p states make up most of the major valence bands just below the gap. Thus it is perhaps a surprise to see a prominent peak derived from the cation s states at the bottom of the major valence-band structure. However, this is a general feature for all sp^3-based compound semiconductors. These are the states responsible for the first observed breakdown of the virtual-crystal approximation for a semiconductor alloy: $Hg_{1-x}Cd_xTe$ (which is caused by the large s-energy shift between the Cd and Hg sites).[4,5,22]

A more detailed examination draws attention to some important differences among the four panels in Fig. 3: the valence-band peak in panel (c) is about 2 eV higher than the rest, and it is also high compared to experiment.[22] Our conduction-band PDOS in panel (a) is broader than the others. The ratio of the integrated PDOS in the conduction bands to that in the valence bands in our model is larger than those in other panels. Also our PDOS just below the valence-band edge is obviously smaller than that found in other models.

Figure 4 shows that the Cd p states are concentrated in the conduction-band states. This is particularly true in panel (a), where their contribution to the valence-band states shrinks almost to nothing. In other panels, there are still sizable (~20%) valence-band states. In contrast, all four panels in Fig. 5 show that the Te s states are confined to the deep valence-band states, as generally recognized. Finally, Fig. 6 shows that the Te p states dominate the upper valence-band states. Panel (a) has much less conduction-band content than the other three panels. As we will see, these differences can result in quantitatively or even qualitatively different predictions about the deep levels.

IV. IMPURITY-LEVEL DETERMINATION

A convenient way to study the impurity energy levels using Eq. (1) is to rewrite it as $v_\alpha = 1/g_\alpha(E)$ and plot E as a function of v. Once this E-v curve is deduced for each α, the deep levels E_α for a given impurity can be read off the curve by drawing a vertical line at the appropriate value of v_α for the impurity. We set the zero of energy at the top of the valence bands. Because the gap is 1.6 eV, we will focus on levels in the energy range from -0.5 to 2.0 eV.

Calculations have been performed for all neutral impurities listed in Table I. Because we do not believe that there exists a uniformly accepted table for v we have adopted a table that we used for structural studies.[23,24] Table I lists the term values, which we obtained from to-

TABLE I. s- and p-state correlated term values in units of $-$eV. The top entry is the s-state, the second the $p_{1/2}$-state, and the third the $p_{3/2}$-state energy. (All energies are *negative*.)

I	II	III	IV	V	VI	VII
Li	Be	B	C	N	O	F
5.390	9.320	14.003	19.814	26.081	28.551	36.229
	5.412	8.300	11.260	14.540	13.613	17.484
	5.412	8.300	11.260	14.540	13.610	17.420
Na	Zn	Al	Si	P	S	Cl
5.140	9.390	11.780	15.027	19.620	21.163	25.812
	4.237	5.980	8.150	10.610	10.449	13.136
	4.011	5.980	8.150	10.550	10.360	13.010
K	Cd	Ga	Ge	As	Se	Br
4.340	8.990	13.230	16.396	20.015	21.412	24.949
	4.313	6.000	7.880	10.146	10.188	12.353
	4.097	5.850	7.694	9.810	9.750	11.840
Rb	Hg	In	Sn	Sb	Te	I
4.180	10.430	12.032	14.525	17.560	19.120	21.631
	4.998	5.780	7.340	9.391	9.951	11.470
	4.031	5.453	6.879	8.640	9.010	10.450
Cs			Pb			
3.890			15.250			
			7.410			
			5.979			
Cu						
7.720						
Ag						
7.570						
3.647						
3.487						
Au						
9.220						
4.349						
3.688						

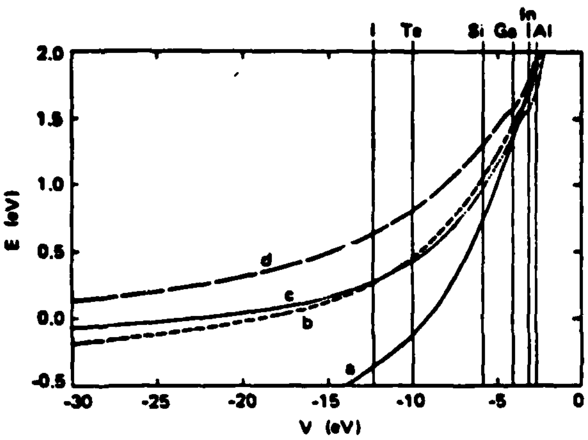


FIG. 7. The E-v curves for the Γ_6 states on a Cd site.

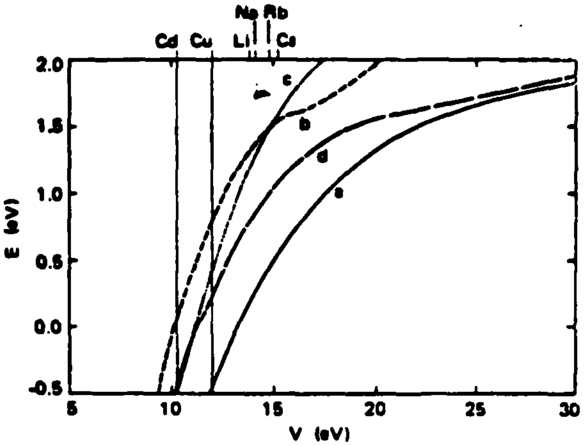


FIG. 9. The E-v curves for the Γ_6 states on a Te site.

tal energy differences between atomic configurations calculated using the norm-conserved pseudopotentials[25] and self-consistent charge-density-functional theory, with the first ionization energies adjusted to be the experimental values.[26] These term values are found to yield consistently better structural properties[23] in Harrison's theory[27,28] than those based on Mann's values[29] adopted by Harrison.[28] The impurity-potential parameters will then be taken as the difference of the term values between the impurity atom and Cd (or Te). To study the sensitivity of E_a to v_a, we shift v_a by ±0.5 eV and compute the corresponding changes in the energy levels.

Figures 7–10 display the E-v curves for several α. Each figure has four curves, corresponding to the four panels of PDOS in each of Figs. 3–6. The functional behavior of these curves can be understood qualitatively using Eq. (2) and Figs. 3–6. If E lies in the gap, the contribution from conduction bands is negative, but positive from the valence bands. The closer the PDOS to the E in question, the larger will be its influence. Applying this argument to the Γ_6(Cd) representation, we see that the curves in Fig. 7 are negative in the gap region because the PDOS in Fig. 3 near the bottom of the conduction bands are much larger than those near the valence-band top. Thus, on the Cd site, only impurities with an s energy below the Cd s level (-8.99 eV) will produce a Γ_6 level in the gap. However, we note that in Fig. 7, $g_a(E)=0$ for models (b) and (c) just below the valence-band edge because of cancellation between the conduction- and valence-band contributions. At this E value, the E-v curve switches from $v=-\infty$ to $v=\infty$ (not shown); an ideal vacancy level (corresponding to $v_a=\infty$) is located at this E. A similar consideration, but with the conduction and valence bands interchanged, leads to an understanding of the curves in Fig. 10. Using the same principle, we can easily understand why all curves in Fig. 9 for the Γ_6(Te) representation are positive, but the reasons for the large displacements between these curves are not easy to deduce. In Fig. 8, the curve labeled a is distinctively different from other curves, because the PDOS in panel (a) in Fig. 4 is completely dominated by the conduction band; however, for the other panels the PDOS just below the valence-band edge are as large as those just above the conduction-band edge. This produces a very sharp negative E-v curve for (a), but split behavior for (b), (c), and (d).

These E-v curves provide a clear picture of how different host band structures may affect the deep levels. Numerical values for the impurity levels can be obtained from these figures by drawing vertical lines at the ap-

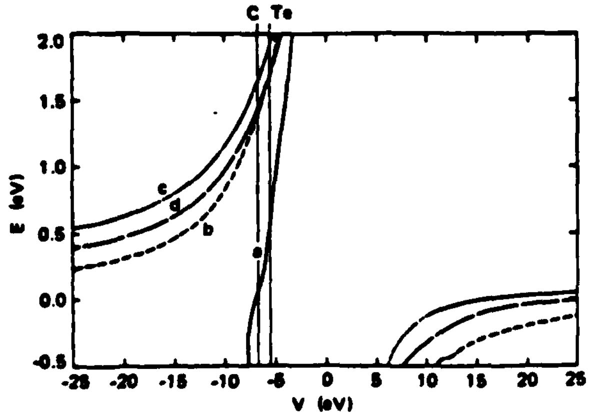


FIG. 8. The E-v curves for the Γ_7 states on a Cd site.

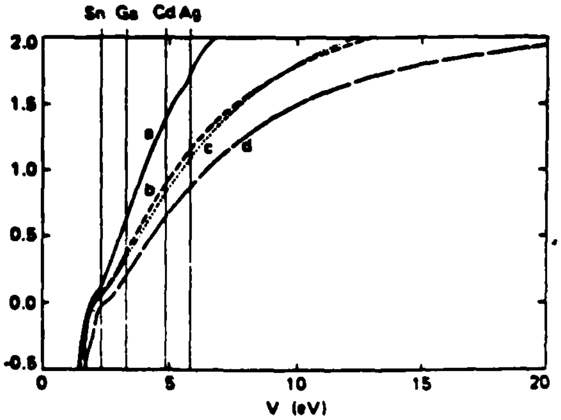


FIG. 10. The E-v curves for the Γ_7 states on a Te site.

propriate impurity potentials (i.e., differences between the term values listed in Table I), as has been shown for several representative impurities. To provide a more quantitative comparison, Table II lists some calculated impurity levels E_a and the corresponding changes ΔE_a due to the 1-eV change in v_a.

V. RESULTS AND CONCLUSION

To summarize we recall that band models (b) and (c) are the same second-neighbor ETB with two different sets of parameters, and model (d) is a first-neighbor ETB with one extra *s* orbital per atom. Our model [model (a)] has the form of ETB but is derived in a very different manner and includes all the long-range interactions. Therefore, we expect that the results from models (b) and (c) will be close, model (d) will have larger discrepancies from (b) and (c) than that between (b) and (c), and model (a) will differ even more. This is evident from Figs. 7–10 and Table II. We found the energies for the Γ_6(Cd), Γ_7(Te), and Γ_8(Te) states produced by models (b) and (c) agree within 0.1 eV. For the other states, i.e., Γ_6(Te), Γ_7(Cd), and Γ_8(Cd), the energies from (b) and (c) are qualitatively similar, but the difference can be as large as 0.4 eV. The largest discrepancy between models (d) and (b) [or (c)] is more than 0.5 eV, and that between (a) and other models is more than 1 eV. The largest difference comes from the *p* levels on a Cd site. For example, the filled *p* level of C on a Cd site in model (a) is a resonance state just below the valence-band edge but is a donor state in the other models. Similarly, model (a) puts the neutral Te antisite defect *p* levels at about $\frac{1}{3}$ and $\frac{2}{3}$ of the gap [$E(\Gamma_7)=0.48$ eV and $E(\Gamma_8)=0.95$ eV], while other models assign them as resonance states inside the conduction bands. We also note that the discrepancies between different models are not uniform, but vary with v_a. Consider Γ_6(Cd) for example. All four models yield the same ordering and about the same energies for the group-III impurities Al, In, and Ga. However, as v becomes more negative, the splitting between the curves increases, so the discrepancies become larger [~ 1 eV difference between models (a) and (d) for I impurity]. Similarly, for the Γ_7(Te) states, all four models put the Sn impurity energies close to the valence-band edge, but the agreement deteriorates as v_a increases.

Regarding the sensitivity of energy levels to impurity potentials, Table II shows that a 1-eV shift in v_a produces a change in E_a ranging from less than 0.1 to 0.65 eV. Very little is known about the size or trends in errors introduced in v_a from the use of atomic term values. However, we know that the discrepancy of v_a between two different tables of atomic term values can be larger than 2 eV. This discrepancy translates into an uncertainty of less than 0.1 to more than 1 eV in the impurity energy levels,

TABLE II. Defect energy levels E and changes ΔE due to a 1-eV change in the impurity-potential parameter. All energies are in units of eV. V_0 stands for ideal vacancy.

	Model (a)		Model (b)		Model (c)		Model (d)	
Defect	E	ΔE	E	ΔE	E	ΔE	E	ΔE
				Γ_6 on Cd site				
Ga	1.29	0.39	1.42	0.24	1.33	0.23	1.57	0.18
C	−0.21	0.09	0.38	0.09	0.36	0.13	0.74	0.08
Si	0.67	0.30	1.02	0.10	0.93	0.19	1.27	0.15
P	−0.19	0.11	0.39	0.09	0.38	0.08	0.75	0.08
O	<-0.5		−0.02	0.02	0.04	0.01	0.32	0.02
Te	−0.13	0.13	0.44	0.10	0.42	0.08	0.79	0.09
Cl	<-0.5		0.06	0.03	0.10	0.02	0.41	0.04
V_0	<-0.5		<-0.5		−0.30		−0.20	
				Γ_7 on Cd site				
C	−0.02	0.37	1.32	0.22	1.59	0.20	1.39	0.19
Si	1.57	0.65	>2.0		>2.0		>2.0	
P	0.16	0.38	1.48	0.26	1.73	0.23	1.52	0.21
O	<-0.5		0.89	0.14	1.22	0.13	1.03	0.12
Te	0.48	0.55	1.60	0.29	1.88	0.23	1.66	0.24
Cl	<-0.5		0.96	0.17	1.29	0.14	1.09	0.14
V_0	<-0.5		0.00		0.21		0.06	
				Γ_6 on Te site				
Li	0.14	0.29	1.28	0.22	1.15	0.35	0.76	0.25
Cu	<-0.5		0.54	0.42	0.12	0.52	0.03	0.32
				Γ_7 on Te site				
Ag	1.89	0.32	1.26	0.22	1.21	0.23	0.99	0.20
Cd	1.66	0.34	1.11	0.26	1.05	0.26	0.85	0.22
Ga	0.98	0.49	0.61	0.33	0.55	0.32	0.40	0.30
Si	−0.07	0.40	−0.11	0.36	−0.13	0.38	−0.38	0.72
Sn	0.28	0.47	0.15	0.31	0.13	0.24	0.02	0.28

which is comparable to that due to different host band structures.

Putting this large uncertainty in the deep levels against a band gap of 1.6 eV, we are left with great doubts about the predictability of this oversimplified theory. Unfortunately, the experimental means available for identifying microdefects in semiconductors are still very limited, and the *ab initio* band theory is still not capable of accurately predicting the energy levels. Thus, there is a great temptation to use simple theories like the one carried out here to help with the identifications. To illustrate this point, consider the following examples: Table II shows that Li on a Te site has an s level of 0.14 eV in model (a), so one may be tempted to relate it to the acceptor state identified experimentally.[14] However, this is not the hydrogenic acceptor state on a Cd site, as one might anticipate. One might also want to assign the $\frac{1}{3}$ and $\frac{2}{3}$ gap states for the Te antisite p levels on the Cd site found from model (a) as those seen in experiments.[15,16] Because of the large uncertainty in the calculation, these results should be regarded as suspicious surprises rather than theoretical confirmations.

The results presented here should not discourage continued research on the ETB approach, but improvement is clearly needed. Work ranging from universal[23,27,30] to specific[24,31,32] structural studies to our band calculations and alloy studies[1-7] indicates that the ETB type of theory is practical for both bonding properties and electronic structures. The reason that ETB works well for some properties, e.g., photoemission spectra and bonding properties, but not for impurity levels, is that the former depend only on the gross total density of states, while the latter have been shown to be sensitive to the details of the partial densities of states.

To establish the credibility of ETB in defect studies, one needs to look at the problem more seriously. The most difficult and yet important task is to develop a better way for determining the Hamiltonian matrix elements. Haas *et al.*[8] and Harrison[27,28] have suggested using the atomic term values as the diagonal matrix elements. Our work[1-4] has suggested using a universal long-range interaction to improve the accuracy of the conduction bands. Several studies[1,27,28,33] have also pointed out scaling rules of the matrix elements. A combination of these ideas may lead to an acceptable model. Secondly, both the bonding and deep-level states of impurities should be studied at the same time in order to provide correlated information for defect identification. Finally, more realistic models should be examined. Besides the substitutional site-diagonal defects, one should consider the possibility of interstitial, paired, and even more complex defects. One also needs to deal with long-range impurity potentials, possible charge shifts, and lattice distortions. Progress in all these areas can be expected if the calculation is constantly correlated with experiments and available *ab initio* theory.

ACKNOWLEDGMENTS

This work was supported by AFOSR Contract No. F49620-81-K0012 and Grant No. AFOSR-84-0282. A.-B. Chen would like to thank Professor W. E. Spicer of Stanford University for his hospitality.

---

[1] A.-B. Chen and A. Sher, Phys. Rev. B 22, 3886 (1980).

[2] A.-B. Chen and A. Sher, Phys. Rev. B 23, 5645 (1981).

[3] A.-B. Chen and A. Sher, Phys. Rev. B 23, 5360 (1981).

[4] A.-B. Chen and A. Sher, J. Vac. Sci. Technol. 121, 138 (1982).

[5] W. E. Spicer, J. A. Silberman, J. Morgan, J. Lindau, J. A. Wilson, A.-B. Chen, and A. Sher, Phys. Rev. Lett. 149, 948 (1982).

[6] A.-B. Chen and A. Sher, Phys. Rev. B 26, 6603 (1982).

[7] A.-B. Chen, S. Phokaichaipatana, and A. Sher, Phys. Rev. B 23, 5360 (1981).

[8] K. C. Hass, H. Ehrenreich, and B. Velický, Phys. Rev. B 27, 5360 (1983).

[9] C. A. Swarts, M. S. Daw, and T. C. McGill, J. Vac. Sci. Technol. 121, 199 (1982).

[10] A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 25, 6367 (1982).

[11] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. B 21, 3545 (1980).

[12] G. A. Baraff, E. O. Kane, and M. Schluter, Phys. Rev. B 21, 5662 (1980).

[13] G. B. Bachelet, M. Schluter, and G. A. Baraff, Phys. Rev. B 27, 2545 (1983).

[14] K. Zanio, in *Semiconductors and Semimetals*, edited by R. Willardson and A. C. Beer (Academic, New York, 1978), Vol. 13.

[15] C. E. Jones, V. Nair, J. Lundquist, and D. L. Polla, J. Vac. Sci. Technol. 21, 187 (1982).

[16] R. T. Collins and T. C. McGill, J. Vac. Sci. Technol. A1, 1633 (1983).

[17] K. K. Kanazawa and F. C. Brown, Phys. Rev. 135, A1757 (1964).

[18] A.-B. Chen, Phys. Rev. B 16, 3291 (1977).

[19] D. J. Chadi, J. P. Walter, and M. L. Cohen, Phys. Rev. B 5, 3058 (1972).

[20] E. O. Kane Phys. Rev. B 13, 3478 (1976).

[21] D. J. Chadi, Phys. Rev. B 16, 3572 (1977).

[22] J. A. Silberman, P. Morgan, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. 21, 142 (1982).

[23] A.-B. Chen and A. Sher, Microscience 3, 1 (1984).

[24] A. Sher, A.-B. Chen, and W. E. Spicer, in *Thirteenth International Conference on Defects in Semiconductors*, edited by L. C. Kimerling and J. M. Parsey, Jr. (The Metallurgical Society of AIME, 1985), p. 335.

[25] G. B. Bachelet, D. Hamman, and M. Schulter, Phys. Rev. B 26, 4199 (1982).

[26] C. Kittel, *Introduction to Solid State Physics, 5th edition* (Wiley, New York, 1976), p. 75, Table 2.

[27] W. A. Harrison, *Electronic Structure and Properties of Solids*

(Freeman, San Francisco, 1980).

[28]W. A. Harrison, Microscience 3, 35 (1983).

[29]J. B. Mann, *Atomic Structure Calculations, I: Hartree-Fock Energy Results for Elements Hydrogen to Lawrencium* (Clearinghouse for Technical Information, Springfield, Virginia, 1967).

[30]D. G. Pettifor and R. Podloucky, Phys. Rev. Lett. 53, 826 (1984).

[31]D. C. Allan and E. J. Mele, Phys. Rev. Lett. 53, 826 (1984).

[32]D. J. Chadi, Phys. Rev. Lett. 52, 1911 (1984).

[33]O. K. Andersen, W. Klase, and M. Nohl, Phys. Rev. B 17, 1209 (1978).

1 Dislocation Energies and Hardness of Semiconductors

A. Sher
SRI International, Menlo Park, California 94025
A.-B. Chen
Auburn University, Auburn, Alabama 36849
W. E. Spicer
Stanford University, Stanford, California 95305

(Received 13 July 1984; accepted for publication 5 October 1984)

The dislocation energies and hardness of semiconductors are calculated by an extension of Harrison's method. It is demonstrated in agreement with experiment that dislocation energies per unit length are proportional to d^{-3}– d^{-9}, where d is the bond length and hardness is proportional to d^{-5}– d^{-11}. The hardness is related to the interaction energies among dislocations. It is argued that dislocation densities of semiconductors will be reduced if they are alloyed with a second constituent that has a shorter bond length. Experimental evidence supporting this strategy is noted.

Dislocations in semiconductors are detrimental to device function; they serve as channels for impurity migration and trapping, which cause nonuniform doping and degrades p-n junctions.[1] They also decrease the material's resistance to plastic deformation. The aim of this letter is to provide insights into the underlying physical mechanisms controlling dislocations and semiconductor hardness, and then to suggest strategies for decreasing dislocation densities. It is well established that the hardness of tetrahedrally coordinated semiconductor materials—groups IV, III-V, and II-VI compounds—exhibits a sharp variation with their near-neighbor distance d, approximately proportional to d^{-9} for one group of seven compounds.[2] Thus, semiconductors with small lattice constants tend to be harder materials. These same materials have larger stiffness coefficients[3] and have fewer dislocations in as-grown crystals.[4-6]

The shear coefficients (combinations of C_{44} and $C_{11} - C_{12}$ in the Schoenflies notation)[7] depend on crystal orientation and (in Harrison's notation[8]) are proportional to $V_2^3/d^3(V_2^2 + V_3^2)^{3/2}$, where $V_2 \propto d^{-2}$ is the covalent and V_3 is the ionic energy. The metallic interaction modifies the functional dependence of the shear coefficient on V_2 and V_3, but introduces no explicit dependence on the hopping integrals, denoted V_1 by Harrison.[3,8] In a pure covalent material, the bond energy is proportional to V_2 (or d^{-2}), and the bond volume is $\propto d^3$; hence, in this case, the shear coefficient varies as d^{-5}. In the limit, $V_3 > V_2$, $C_{11} - C_{12} \propto d^{-11}$. For most polar semiconductors, d^{-9} is a good approximation.

Hardness is determined by applying a known force F to a probe of a prescribed shape driving it into the surface of the sample.[7] The area A of the resulting indentation is measured, and the hardness is the force per unit indented area. Many dislocations must be formed to allow the probe to indent the semiconductor. If the indenter is a rectangular pyramid, then the hardness is $H = F/A = Fh/Ah = \epsilon_T/Ah$, where ϵ_T is the work required to cause the indenter to penetrate to a depth h. A side view of the indentation in a cut through its center is illustrated schematically in Fig. 1. The top of the indentation has side length W; thus, $A = W^2$. The Burger's vector has magnitude b, proportional to the bond length d. The number of dislocations N_h required to accommodate an indentation to depth h is $N_h = h/b = \frac{1}{2}W \cos \vartheta /b$, where ϑ is the angle between the normal to the tip of the indenter and a side. Figure 1 also shows a model of one possible configuration of the dislocations. The edges of the extra atom planes that are driven from the indented volume into the bulk of the semiconductor are shown as lines terminated by dots. The dotted ends of these lines are the positions of the dislocations, which are perpendicular to the plane of the figure. The planes driven to the sides each have a finite extent and a trapezoidal shape. The planes driven down under the indenter have a square shape.

Much of the work done on the indenter goes into the energy to form the indicated dislocation configuration, although some certainly goes into heat. There are two major contributions to this formation energy. The first is the enegy needed to generate each dislocation as an isolated entity, and the second is the interaction energy among these dislocations. Because the interaction term dominates H, approximations made to simplify the first term are relatively unimportant. The extra planes driven to the sides of the indentation have a finite extent; accordingly (in this idealized picture), there are both edge dislocations at their base and screw dislocations associated with their termination. The square planes driven below the indentation have edge dislocations around the sides and screw dislocations at the corners to make the turns. Moreover, there are interactions

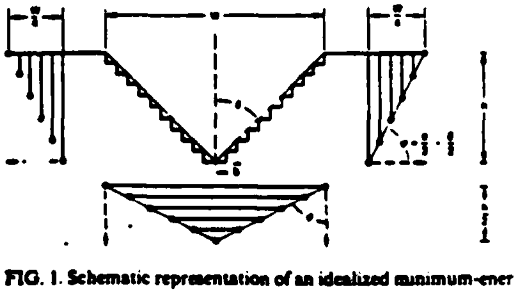


FIG. 1. Schematic representation of an idealized minimum-energy dislocation configuration produced by a square cross-section indenter in a hardness measurement. W is the side length and h is the depth of the indentation. In this ideal case, half the material from the indented region is displaced along the glide planes (indicated by the dashed arrows) to the sides and half is displaced below the indenter.

among the dislocations, which can produce a minimum-energy configuration. For the arrangement depicted in Fig. 1, it always costs energy to position a second dislocation on a parallel glide plane to one already present. However, the magnitude of this extra energy can be minimized and, for proper configurations of the dislocations, there are attractive forces along the glide planes that will tend to position the dislocations into the minimum-energy configuration. The minimum configuration arises when half of the atoms from the indented volume go respectively to the side and below the indenter. Then, in both regions, the maximum angle made between successive close-spaced dislocation lines and their glide planes is $\vartheta = \pi/2 - \vartheta/2$, as shown in the figure. This is the minimum realistic energy configuration. If the dislocations are separated more than shown in Fig. 1, then there is more volume of strained material and the interaction energy would be larger still.

An approximate expression for the energy required to indent the material is[7]

$$\epsilon_T = 2\cdot4\left[\sum_{i=1}^{N} E_i L_i + \frac{1}{2}\sum_{i,j=1}^{N}{}' E_{ij}\left[\min(L_i,L_j)\right]\right] \qquad (1)$$

In the first term, using an isotropic medium approximation and neglecting core terms, the energy per unit length to form an edge dislocation is[7]

$$E_i = \frac{Gb^2}{4\pi(1-\nu)}\ln\frac{R}{r_0}, \qquad (2)$$

the shear coefficient is G, the Burger's vector $b = d/3$ for an indentation along a $\langle 100\rangle$ axis, the range of the elastic deformation of a dislocation R is taken equal to W (for want of a better approximation), $r_0 \sim d$ is the dislocation core radius, the Poisson ratio is $\nu \sim 0.2$ for most semiconductors, and L_i is the length of the ith dislocation. In the second term, E_{ij} is the interaction energy per unit length between dislocations i and j. Assuming they have parallel glide planes and their Burger's vectors have the same sign, E_{ij} is given by[7]

$$E_{ij} = \frac{Gb^2}{2\pi(1-\nu)}\left[\ln\left(\frac{R}{r_{ij}}\right) + \cos^2\varphi_{ij}\right], \qquad (3)$$

where r_{ij} is the separation, and φ_{ij} is the angle that a line perpendicular to and joining the dislocations makes with the glide plane, as shown in Fig. 1. Because the various dislocations in a region have different lengths, the net interaction energy is approximated by multiplying the energy per unit length by the length of the shorter one. The upper limits on the sums N are the number of dislocations in one region (side or bottom) associated with one edge. For the minimum-energy configuration, $N = N_h/2$. The four that multiplies the bracket accounts for the four sides, and the two for the two regions for each side.

We now encounter our first surprise. As we can see from Eq. (2) and (3), E_{ij} and E_i have comparable magnitudes. Because there are approximately N^2 terms in the interaction energy sum, only N terms in the formation energy sum, and N is a large number ($N \gg 1$), the interaction energy completely dominates the hardness. In fact, N is typically of the order 10^5. Hence, terms owing to screw dislocations, core energies, heat dissipation as the dislocations propagate to their places, and other effects associated with the first term are unimportant. However, care must be taken with the interaction terms: Eq. (1) neglects a number of secondary interactions, some positive and others negative; these will be added later. The principal neglected terms are the interactions between the dislocations in the different regions on each side (positive), and the interactions between adjacent side and bottom regions (positive), and the integraction between the opposite sides in the bottom regions (negative). Comparison of the results with experiment will indicate how important these neglected terms are likely to be. The length L_i is $W(N-i)/N$ for i from 1 to N. This is the largest length of the side-inserted planes and its choice partially accounts for interacitons between the otherwise neglected screw dislocations. The distance $r_{ij} = r_i - r_j$ is given by $2\sqrt{2b}\,|i-j|$ for i and j ranging from 1 to N for the minimum-energy configuration and a tetrahedrally bonded semiconductor. Finally, in the indicated configuration, $\varphi_{ij} = \varphi = \pi/2 - \vartheta/2$. Inserting these expressions into Eq. (1) and retaining only terms of order N^2 yields

$$H_{\min} = \frac{G\cot\vartheta}{6\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\frac{\vartheta}{2}\right]. \qquad (4)$$

One can also get a number for the hardness of a dislocation in which all the material is pushed along the same glide plane, e.g., to the side, to the bottom, or normal to the face of the indenter (a possibility not depicted in Fig. 1). In this case, the factor of 2 in front of Eq. (1) is removed, $N = N_h$ and $\phi = \pi/2 - \vartheta$. Then a higher nonequilibrium hardness in the context of this model (denoted H_1), is obtained.

$$H_1 = \frac{G\cot\vartheta}{3\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\vartheta\right]. \qquad (5)$$

The proper answer for most materials, and depending on crystal orientation, probably lies somewhere between $H_{\min}$ and H_1. For an indenter with $\vartheta = \pi/4$, we have $H_{\min} = 0.0969G/(1-\nu)$ and $H_1/H_{\min} = 2.39$. Harrison[8] has shown that one contribution to the shear coefficient (actually $C_{11} - C_{12}$) is $G = 2.38\,\hbar^2 m\alpha_c^3/md^5$, where m is the free-electron mass, α_c is the covalence, $\alpha_c = V_2/(V_2^2 + V_3^2)^{1/2}$, and d is the bond length. We will approximate G by this expression. Using this G and $\nu = 0.2$, and changing the dimensions to those in terms of which experimental hardness numbers are customarily quoted gives $H_{\min} = 2.38 \times 10^4\,(\alpha_c^3/d^5)$ kgm/mm^2, where d is in angstroms. Calculated values of $H_{\min}$ and H_1 are plotted against experimental results in Fig. 2 for a number of semiconductors.

Figure 2 has the theoretical $H_{\min}$ and H_1 values connected by arrows from $H_{\min}$ to H_1 for each compound, plotted as a function of the corresponding experimental values.[2,9] If the theory were perfect and the experimental values were accurate, the points would fall on the indicated unity slope line. Several conclusions can be drawn. Firstly, the order of magnitude of the predicted and measured values are the same, a result obtained with no adjustable parameters in the theory. Secondly, the trends from one compound to another are properly given by the theory. Although the $H_{\min}$ values are generally too small, they fit the soft materials better, and the H_1 values fit the harder materials better. Thirdly, from Eqs. (1) and (3), H is given in a rough but revealing

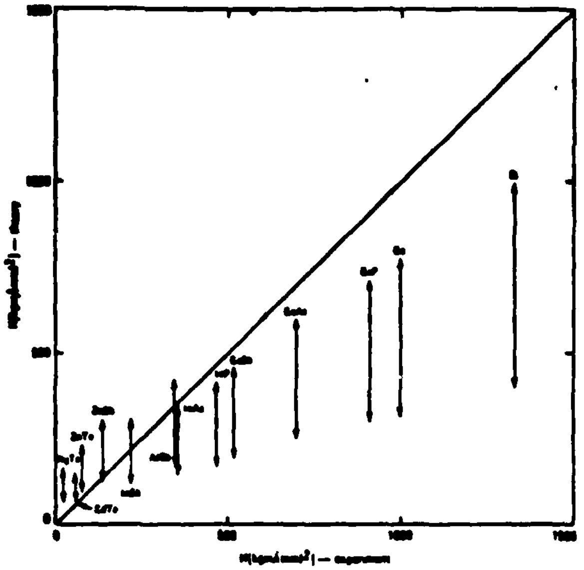


FIG. 2. Theoretical vs experimental hardness of several semiconductors in Refs. 2 and 9. The two theoretical values for each semiconductor are connected by an arrow from H_{max} to H_t. Perfect agreement would correspond to points being on the unity-slope solid line.

approximation by $4N^2\bar{E}_{ij}\bar{L}/Ah$, where the averge dislocation length $\bar{L} = W/2$ and $\bar{E}_{ij}$ is the average dislocation pair interaction energy per unit length. Notice that N (or W) cancels from this expression; thus, H is independent of W (or F) and therefore, H is truly a measure of the properties of the material. This result would not be found if the dislocation energies ($\propto N$) were to dominate H rather than the pair interaction energy ($\propto N^2$). Finally, the Berger's vector cancels from the leading term and appears only in the argument of the logarithm in Eq. (3). Thus, the answers are also insensitive to its choice.

Dislocations are often found in materials as they are grown. Their density is determined by the thermal and mechanical stresses to which they are subjected in the growth process. A dislocation constitutes a metastable excitation relative to the perfect crystal ground state. At the elevated growth temperatures and temperature gradient behind the growth front, the number of dislocations present is controlled by the relative rate at which vacancies anneal or condense into dislocations.[10] The dislocation formation rate will be slower in a material grown at the same temperature if E_i is higher. If an alloy is formed from a material of interest and a second constituent with a shorter bond length, one expects the average bond energy (and thus both the melt temperature and vacancy formation energy) to increase proportional to a low inverse power of the average bond length.[8] Hence, the equilibrium vacancy density just below the melting point tends to be the same in lowest order for all materials, independent of the bond lengths of the constituents. However, the shear coefficient and dislocation energy per unit length will increase with much higher inverse powers of the bond length. Consequently, dislocation densities should be reduced in such alloys relative to those found in the longer bond length pure constituent. This expectation is confirmed in the recent work reported on $Zn_{1-x}Cd_xTe$ bulk material.[4] The best CdTe that has ever been grown has dislocation densities in excess of 5×10^5 cm^{-2}. The addition of only 4% Zn reduced the dislocation count to less than 5×10^4 cm^{-2}. The ZnTe bond length is 2.643 Å, while the CdTe bond length is 2.805 Å, a 6% difference. This 6% difference in bond length translates into a 2% difference in the dislocation energy per unit length for $1-x = 0.04$. Dislocation energies per unit length are typically 10 eV per lattice spacing; accordingly, a 2% increase can be expected to slow their formation rate considerably.

The argument just presented naturally leads to a strategy for decreasing dislocations in other semiconductors. If an alloy is made of the material of interest with another compound with a shorter bond length, then the dislocation density should be reduced. For example, this suggests that the addition of a small amount of GaP ($d = 2.359$ Å) may significantly reduce the dislocation density of bulk grown GaAs ($d = 2.448$ Å). It has been demonstrated that the addition of approximately 1% GaN ($d = 1.946$ Å)[5,6] or of a 10^{19} cm^{-3} BAs ($d = 2.069$ Å) concentration[11] to GaAs can yield a large volume of dislocation-free material. An InAs additive with its longer bond length ($d = 2.623$ Å) serves the same function indirectly, by causing GaAs bonds in its neighborhood to be compressed. This indirect mechanism should be less effective than substituting short bond length additives.

We have demonstrated that the dislocation energies and hardness of tetrahedrally bonded semiconductors are rapid functions of the reciprocal of the bond length. This rapid d dependence of dislocation energies provides a rationale for the dramatic decrease of the dislocation density in bulk grown $Zn_{0.04}Cd_{0.96}Te$ material relative to that found in CdTe, and suggest means for accomplishing the same ends in other materials.

The authors are indebted to J. P. Hirth, W. A. Harrison, and T. N. Casselman for helpful comments. This work was supported in part by DARPA contract MDA 903-83-C-0108 and AFOSR contract 49620-81-K-0012.

[1] E. M. Swiggard, Proceedings of GaAs IC Symposium, Phoenix, AZ 1983, p. 26.
[2] N. A. Groyunova, A.S. Borshchevskii, and D. N. Fretiakov, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, NY, 1968), Vol. 4, Chap. 1.
[3] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); R. C. Sokel, thesis, Stanford University, 1978; W. A. Harrison, *Microscience* (limited distribution SRI International publication, Menlo Park, 1983), Vol. 4, p. 34.
[4] S. L. Bell and S. Sen, presented at Infrared Imaging Systems (IRIS) Detector Specialty Group Meeting, Boulder CO, 1983; T. W. James and B. F Zuck, *ibid*.
[5] Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. 42, 822 (1983).
[6] G. Jacob, J. Cryst. Growth 59, 669 (1982).
[7] J. W. Christian, *Theory of Transformations in Metals and Alloys*, 2nd ed (Pergamon, NY, 1975), Chap. 7. For a discussion of the interaction energies, see J. P. Hirth and J. Lothe, Theory of Dislocations, 2nd ed (Wiley, NY, 1982), pp. 262–263.
[8] W. A. Harrison, Phys. Rev. B 27, 3592 (1983).
[9] S. Cole and A. F. W. Willoughby, J. Cryst. Growth 59, 370 (1982)
[10] G. Schoeck and W. A. Tiller, Philos. Mag. 5, 43 (1960)
[11] S. Miyazawa, 1983 European Patent Application 833021652, filing date 4/18/83, and references therein.

IV.8

PHYSICAL REVIEW B VOLUME 31, NUMBER 2 15 JANUARY 1985

a. Bond-length Relaxation in Pseudobinary Alloys

C. K. Shih, W. E. Spicer, and W. A. Harrison
Stanford University, Stanford, California 94305

Arden Sher
SRI International, Menlo Park, California 94025

(Received 17 September 1984)

The bond-length relaxation in pseudobinary alloys can be predicted by a simple radial force model. In tetrahedral structure alloys the bond-length deviation of the solute in the dilute solution is a quarter of the bond-length difference between the two components. This result agrees with the experimental work done on the $Ga_{1-x}In_xAs$ system performed by Mikkelsen and Boyce.

In calculating pseudobinary alloy properties, the nearest-neighbor distance is a very important parameter. For many years people have used the virtual-crystal approximation (VCA) which assumes that all atoms occupy the average lattice positions. In 1982 Mikkelson and Boyce used extended x-ray-absorption fine structure (EXAFS) to study the atomic scale structure of the $Ga_{1-x}In_xAs$ system.[1] They found that the Ga–As and In–As bond lengths did not follow the VCA; instead, they only slightly deviated from their natural bond lengths even in the dilute solutions, and the magnitude of the bond-length deviation at the dilute limit was about a quarter of the difference of the natural bond lengths of the two components (GaAs and InAs). These results can be explained by a simple radial force constant model in which the weaker angular forces are neglected.

Consider a compound GaAs in which one of the Ga atoms is replaced by an In atom. This configuration represents the dilute limit of In-As in a GaAs host (see Fig. 1). In this configuration, four In–As bonds are identical and the tetrahedral symmetry is preserved at the In atom. We regard any difference between Ga and In as a first-order difference. Let us call d_{GaAs} the natural bond length of the host and K the stretching force constant for the host. Adding In will change K by an amount we regard as first order, but that alone will not give the distortion. However, because the natural bond length of InAs, d_{InAs}, is larger than d_{GaAs} by a first-order amount the As atoms around the In impurity move outward by a first-order distance; we hold the second neighbors fixed and to first order can take K to be the value for the host. Let us call u the outward relaxation displacement of the As. Then Δd_{InAs}, the amount that the final In–As bond length deviates from d_{InAs}, is $d_{GaAs}-d_{InAs}+u$. The bond lengths of the neighboring Ga-As are decreased by $-u\cos\theta$ equal to $-u/3$ to first order. The energy associated with the distortion is

$$E=4[(k/2)(d_{GaAs}-d_{InAs}+u)^2+3(k/2)(u/3)^2] .$$

By minimizng the energy with respect to u, we obtain

$$u=\frac{1}{4}(d_{InAs}-d_{GaAs}) ,$$

or the deviation of the In–As bond length

$$\Delta d_{InAs}=\frac{1}{4}(d_{GaAs}-d_{InAs}) .$$

Similarly, the bond-length deviation of dilute Ga-As in an InAs host is

$$\Delta d_{GaAs}=\frac{1}{4}(d_{InAs}-d_{GaAs}) .$$

This predicts exactly Mikkelsen and Boyce's result.[1] Note that this displacement leaves all metallic atoms in their original sites; only the As atoms move. To a good approximation this is true in the concentrated alloy, where the metal atoms form a face-centered-cubic lattice of the weighted average lattice parameter and the As atoms move to fit local bond lengths.

One can apply this principle to the rocksalt-structure systems, e.g., $K_{1-x}Rb_xBr$, and predict that the bond-length deviation of the solute in the dilute solution is half of the

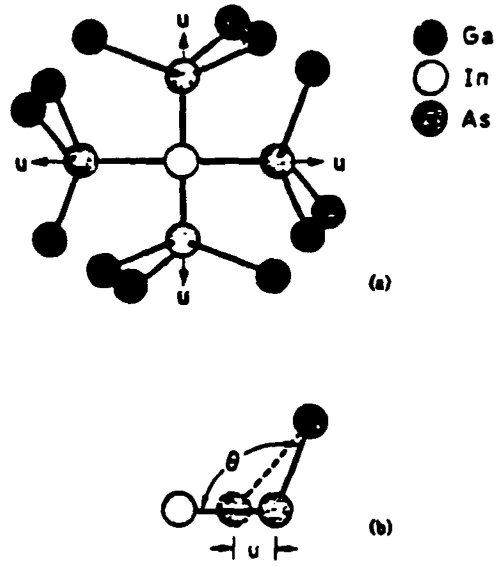


FIG. 1. (a) An indium atom substitutes for gallium in GaAs; the tetrahedral symmetry is preserved at the In atom. (b) The displacement of As atoms, u, results in the Ga–As bonds being decreesed by $u\cos\theta$. The In–As bond lengths are varied by $d_{GaAs}+u-d_{InAs}$.

difference between the two components:

$$\Delta d_{\mathrm{KBr}} = \frac{1}{2}(d_{\mathrm{RbBr}} - d_{\mathrm{KBr}}) \ ,$$

$$\Delta d_{\mathrm{RbBr}} = \frac{1}{2}(d_{\mathrm{KBr}} - d_{\mathrm{RbBr}}) \ .$$

However, because of the lattice geometry, the approximation that all the cations stay in their original sites might not be as good. In rocksalt structure, the restoring force on the displaced first neighbor results from the atom right behind the first neighbor; one would expect that the relaxation extends more than one interatomic distance and the value of the bond-length deviation would be somewhat smaller than what is predicted by this model. This agrees with the experimental result.[2]

We thank Dr. Alex Zunger for valuable discussions. This work was supported by the Defense Advanced Research Projects Agency under Contract No. 800989-B8.

---

[1] J. C. Mikkelsen, Jr. and J. B. Boyce, Phys. Rev. B 28, 7130 (1983).

[2] J. B. Boyce and J. C. Mikkelsen, Jr., Bull. Am. Phys. Soc. 29, 271 (1984).

IV.[illegible]9

Generalized Brooks' formula and the electron mobility in Si_xGe_{1-x} alloys

Srinivasan Krishnamurthy and A. Sher
SRI International, Menlo Park, California 94025
An-Ban Chen
Auburn University, Auburn, Alabama 36830

(Received 19 March 1985; accepted for publication 7 May 1985)

A formula for alloy-scattering-limited electron mobility in semiconductors is obtained for indirect gap systems with multiple band minima. All the input parameters needed are defined explicitly. The drift mobility of Si_xGe_{1-x} which has a dip at $x \sim 0.13$ and a broader minimum at $x \sim 0.5$ is calculated by adding alloy scattering to other scattering mechanisms and correlates well with the measured Hall mobility.

The electron and hole mobilities in semiconductors are determined by the band structure and various scattering mechanisms, predominately impurity and phonon scattering. For alloys, the mobility is also affected by disorder arising from aperiodic atomic potentials and atomic positions. Many years ago, Nordheim[1] and Brooks[2] obtained an expression for alloy-scattering-limited electron mobilities in metals and semiconductors, respectively. Brooks' well-known formula reads

$$\mu_A = \frac{\sqrt{2\pi} e \hbar^4 N_0}{3x(1-x) m^{*5/2}} \frac{1}{(\Delta E)^2 \sqrt{KT}}, \qquad (1)$$

where N_0 is the number of atoms per unit volume, m^* is a band-edge effective mass, x is the fractional concentration of one of the species, and ΔE is an energy parameter characterizing the alloy potential fluctuations. Although this formula has been widely and, to some extent, successfully used for direct gap materials,[3-5] the identification of the alloy disorder parameter ΔE remains uncertain. Various suggestions have previously been made for ΔE, e.g., and band-edge discontinuity[6] or band-gap differences.[7] Any of these simple choices is bound to fail when one applies Eq. (1) to more complicated indirect gap systems such as Si_xGe_{1-x} alloys, where one encounters conduction-band minima transferring between the X and L points of the Brillouin zone. For example, if ΔE is taken to be the difference in corresponding band edges, then one finds that $\Delta E \sim 0.1$ eV for the X (Δ) valley and ~ 1.2 eV for the L valley. The values that fit the experiment are about half this value for L and ~ 0.5 eV for X.[6] The purpose of this letter is to resolve the identity of ΔE for indirect gap materials.

Moreover, there is a problem with the m^* that enters Eq. (1). For direct gap alloys, the band-edge effective mass at Γ naturally enters Eq. (1). For the indirect gap alloys, the effective mass is anisotropic and hence an appropriate mass must be chosen. Previous authors[6,7] have chosen m^* to be the effective conductivity mass m_c^*. We shall show that different masses enter for different cases.

The first unambiguous assignment for ΔE in a direct gap alloy was given by Hass *et al.*[5] To estimate the limiting electron mobility in $Hg_xCd_{1-x}Te$ based on a tight-binding (TB) band description, they defined ΔE to be $f_s \Delta E_s$, where f_s is the s fraction in the density of states and ΔE_s is the difference between the s atomic term values of the Hg and Cd atoms. By extending this approach to alloys with indirect gaps and multiple bands, we show that all the uncertainties identified above are resolved. Our generalized Brooks' formula will then be applied to Si_xGe_{1-x} systems to explain their observed mobility.[7,8]

Because Brooks' formula has never been derived explicitly in the literature, we rederive it first and then generalize it. Consider the case of a single band with an isotropic effective mass. The dc electronic conductivity based on the linear response theory[9] is given by

$$\sigma = \int \sigma(\epsilon) \left(\frac{-df}{d\epsilon} \right) d\epsilon, \qquad (2)$$

where the energy-dependent $\sigma(\epsilon)$ in the weak alloy scattering limit is

$$\sigma(\epsilon) = (e^2/3) v^2(\epsilon) D(\epsilon) \tau(\epsilon). \qquad (3)$$

$D(\epsilon)$ is the density of states (DOS) per unit volume for both spins, so $D(\epsilon) = 2N_A \rho(\epsilon)$, with $N_A = N_0/2$ being the number of unit cells per volume (for the diamond structure, half the number of atoms N_0 per unit volume) and $\rho(\epsilon)$ being the DOS per unit cell per spin. The mean square velocity $v^2(\epsilon)$ for carriers with energy ϵ is given by

$$v^2(\epsilon) = \sum_{\mathbf{k}} |v(\mathbf{k})|^2 \frac{\delta[\epsilon - \epsilon(\mathbf{k})]}{\rho(\epsilon)}. \qquad (4)$$

The scattering lifetime for carriers with energy ϵ, $\tau(\epsilon)$, is related to the alloy broadening $\Delta(\epsilon)$ by $\tau(\epsilon) = \hbar/2\Delta(\epsilon)$, where the energy $\Delta(\epsilon)$ is the imaginary part of the self-energy in the averaged alloy Green's function. For weak scattering $\Delta(\epsilon)$ is

$$\Delta(\epsilon) = \pi x(1-x)(\Delta E)^2 \rho(\epsilon), \qquad (5)$$

where in a tight binding (TB) description ΔE is the difference in the term values of the constituents. Then the mobility is $\mu_A = \sigma/ne$ with the electron density given by

$$n = 2\int f(\epsilon) \rho(\epsilon) d\epsilon. \qquad (6)$$

For a nondegenerate semiconductor, $f(\epsilon)$ is the Boltzmann distribution and $f(\epsilon) \propto e^{-(\epsilon - \epsilon_0)/kT}$. Furthermore, for a parabolic band $\epsilon(\mathbf{k}) = \hbar^2 k^2/2m^*$, $\rho(\epsilon) = (2m^*)^{3/2} \epsilon^{1/2}/4\pi^2\hbar^3$, then all the above equations can be combined to arrive at Eq. (1).

For a real semiconductor alloy in a TB description, the alloy scattering can be characterized by two parameters ΔE_s and ΔE_p, the differences in s and p term values between two substitutional atoms. Then an effective broadening is given

0003-6951/85/140160-03$01.00

by

$$\Delta(\epsilon) = (\Delta_s \rho_s + \Delta_p \rho_p)/\rho, \tag{7}$$

where ρ_s, ρ_p are partial density of states (PDOS) and Δ_s and Δ_p are similar to Eq. (5), with ρ replaced by ρ_s and ρ_p, respectively. For $Hg_{1-x}Cd_xTe$, the s disorder is predominent[3,10] and one can neglect ΔE_p. Defining $\rho_s = f_s\rho$ (and $\rho_p = f_p\rho$), one arrives at

$$\Delta \simeq \pi x(1-x)(f_s \Delta E_s)^2 \rho.$$

Thus as was pointed out by Hass *et al.*,[5] $f_s \Delta E_s$ plays the role of ΔE in this special case where ΔE_p can be neglected.

For an alloy with a single indirect gap minimum, one has to consider both s and p contributions to the alloy broadening and the masses that enter ρ and v^2. Again, Eqs. (2)–(6) can be combined to yield

$$\mu_A = \frac{(e\hbar^4 N_0 \sqrt{2\pi})}{[3x(1-x)m_c^* m_t^* (m_l^*)^{1/2} (kT)^{1/2} N_v (f_s^2 \Delta E_s^2 + f_p^2 \Delta E_p^2)]}, \tag{8}$$

where m_l^* and m_t^* are respectively the longitudinal and the transverse mass at the band edge, and N_v is the number of equivalent minima, e.g., 6 for Si. The conductivity mass m_c^* comes from averaging v^2 in Eq. (2) and is given by $3(2/m_t^* + 1/m_l^*)^{-1}$. Equation (8) clearly identifies the masses and the energy parameter that enter Brooks' formula.

Next we consider a still more complicated case where the contribution to the mobility comes from more than one band. For example, in Si_xGe_{1-x} the X and L minima cross near $x = 0.15$.[11] There are now two contributions to the net conductivity, so $\sigma = \Sigma\sigma_i$, where i is X or L. The quantities $v_i^2(\epsilon)$, $D_i(\epsilon)$, and $N_i(\epsilon)$ now take different values for different bands. The structure of $\tau_i(\epsilon)$ requires more careful consideration. The complication comes from the fact that the effective broadening Δ is still given by Eq. (7), but ρ_s, ρ_p, and ρ contain contributions from both the bands. The proper expressions are $\rho = \Sigma_i \rho_i N_v^i$ and $\rho_\alpha = \Sigma_i f_{\alpha i} N_v^i \rho_i$ where $i = X$ or L, $\alpha = s$ or p, and $N_v^X = 6$, $N_v^L = 4$. The equation for Δ is

$$\Delta(\epsilon) = \pi x(1-x) \sum_\alpha \Big(\sum_i f_{\alpha i} N_v^i \rho_i(\epsilon) \Delta E_\alpha \Big)^2 \Big/ \Big(\sum_i \rho_i(\epsilon) N_v^i \Big). \tag{9}$$

The mobility associated with the ith band is defined as $\mu_i = \sigma_i/(n_i e)$, then

$$\mu_i^A = \frac{\pi e\hbar^4 N_0}{3x(1-x)} \frac{1}{[m_c^* m_t^* (2m_l^*)^{1/2}]_i} I_i, \tag{10}$$

$$I_i = \int_0^\infty \frac{\epsilon \rho_i(\epsilon) \Big(\sum_j N_v^j \rho_j(\epsilon) \Big) e^{-\epsilon/kT}}{\sum_\alpha \Big(\sum_j f_{\alpha j} N_v^j \rho_j(\epsilon) \Delta E_\alpha \Big)^2} d\epsilon \Big/ \int \epsilon^{1/2} e^{-\epsilon/kT} d\epsilon. \tag{11}$$

Thus, the generalized formula no longer has the explicit x and T dependences of the original Brooks' form. However, all the quantities needed—the masses, the scattering parameters ΔE_α, the band gaps, and the fractions $f_{\alpha i}$—can be evaluated theoretically without resorting to experimentally fitted parameters. To demonstrate, we shall apply Eq. (10) to Si_xGe_{1-x}. The band quantities are obtained from our CPA calculation.[11] We found that the effective masses vary weakly with the concentration, so m_l^* and m_t^* are assumed to be constant and assigned the values 0.97 and 0.19 for the X minima and 1.64 and 0.082 for the L minima, respectively. The calculated energy gaps for the $X(\Delta)$ follows the functional form $E_g^{(X)} = a + bx + cx^2$ and for L is given by $E_g^{(L)} = A + Bx + Cx^2$. All the parameters of our calculations are listed in Table I.

TABLE I. Calculation parameters.

Parameter	Si_xGe_{1-x} systems
$m_l^*(X)$	$0.97m_0$
$m_t^*(X)$	$0.19m_0$
$m_l^*(L)$	$1.64m_0$
$m_t^*(L)$	$0.082m_0$
$E_g^X(x)$	$0.8941 + 0.0421x + 0.1691x^2$
$E_g^L(x)$	$0.7596 + 1.0860x + 0.3306x^2$
$f_{sX}(x)$	$0.333 + 0.05x \quad (0<x<0.3)$ $0.339 + 0.03x \quad (0.3<x<1.0)$
$f_{sL}(x)$	$0.632 + 0.13x$

To correlate the calculation with the measured mobilities, we need to have an estimate of scattering rates $1/\tau_0$ due to impurities and phonons. A crude approximation is to assume $1/\tau_0$ for a given valley to be the same as the appropriate constituent's values and add to it the alloy scattering rate $1/\tau_A$. Then the average mobility and the mobility from the ith minimum in the alloy are

$$\mu = \sum_i n_i \mu_i \Big/ \sum_i n_i,$$

$$(\mu_i)^{-1} = (\mu_i^0)^{-1} + (\mu_i^A)^{-1}, \tag{12}$$

μ_i^A is given by Eq. (10) and μ_i^0 are the measured drift mobilities for Si or Ge.[12] The drift mobility, calculated from Eq. (12), is plotted as a function of alloy concentration x in Fig 1.

For $x<0.05$ and $x>0.20$, the energy difference between the X and L edges is large enough so there is a negligible contribution to the mobility from the higher minima. In the Si_xGe_{1-x} system, the s scattering is predominent. Because

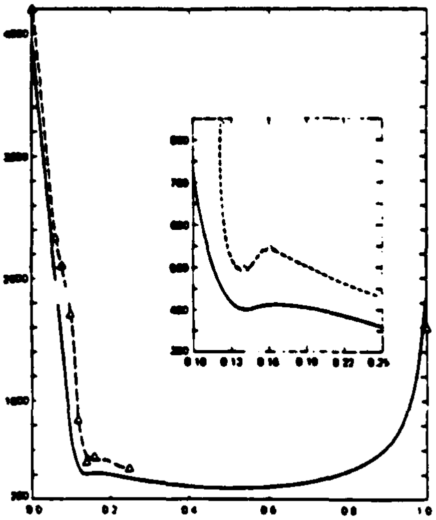


FIG. 1. Calculated (solid line) electron drift mobility and the experimental Hall mobility (dashed line) from Refs 6 and 8 are plotted as a function of alloy concentration.

the L edges have a larger s content, alloy scattering reduces the average mobility substantially for small x. Even though the s content is almost the same for all $x > 0.20$ at the X edge, the mobility still decreases to $x \simeq 0.5$ as shown in Fig. 1 because of the $x(1-x)$ term in Eq. (10).

An interesting feature is obtained for the compositions $0.13 < x < 0.18$. The average mobility attains a local minimum near $x = 0.14$ and a smaller maximum at $x = 0.17$. This feature occurs because of the X to L crossover.[11] For $x < 0.14$, the major contribution to μ comes from L minima. Near $x = 0.14$, the density of states increases because the X and L minima merge. So the alloy scattering increases there and the average mobility decreases. For $x > 0.14$, the X bands have the lower minima. As the s content is small at the X minima, the reduced alloy scattering increases the average mobility. For larger values of x, the $x(1-x)$ term takes over and the mobility varies as shown. The values of measured Hall mobility in $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ systems are also plotted in Fig. 1. The interesting feature near $x = 0.14$ is clearly seen. Since the experimental drift mobility μ_D for $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ is not available and the generalization of Eq. (10) to Hall mobility μ_H is less clear, we present the calculated μ_D and experimental μ_H (Ref. 7,8) here. While we do not expect quantitative agreement, because μ_H/μ_D can range from 1 to 2,[13,14] we do expect them to display the same qualitative x dependence. It is rewarding to note the similarity in the trend in Fig. 1. Previous authors explained the dip in the mobility curve by including intervalley scattering with an arbitrary adjustable coupling constant.[8] Our calculations automatically include that portion of intervalley scattering that results from alloy disorder with a coupling constant set by the atomic properties of the constituents. However, the additon of intervalley scattering mediated by phonons and impurities is expected to increase the dip near crossover.

In summary, an expression for alloy-scattering-limited charge carrier mobilities is derived for indirect gap alloys with multiple bands. This expression reduces to Brooks' formula for direct gap alloys. The quantities m^* and ΔE can be calculated exactly. Alloy scattering accounts for the observed mobility features in the $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ alloy, including the anomaly near the L to $X(\Delta)$ crossover.

A.-B. would like to thank Professor W. E. Spicer for his hospitality at Stanford University. This work was supported in part by DARPA contract MDA 903-83-C-0108 and grant AFOSR-84-0282.

[1] L. Nordheim, Ann. Phys. 9, 607 (1931); 9, 641 (1931).
[2] H. Brooks (unpublished). A discussion of this formula can be found, for example, in L. Makowski and M. Glicksman, J. Phys. Chem. Solids 34, 487 (1973).
[3] A. Chandra and L. F. Eastman, J. Appl. Phys. 51, 2669 (1980).
[4] D. Chattopadhyay and B. R. Nag, Phys. Rev. B 12, 5676 (1975).
[5] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B 27, 1088 (1983).
[6] J. W. Harrison and J. R. Hauser, J. Appl. Phys. 47, 292 (1976).
[7] M. Glicksman, Phys. Rev. 111, 125 (1958).
[8] M. Glicksman, Phys. Rev. 100, 1146 (1955).
[9] A.-B. Chen, G. Weisz, and A. Sher, Phys. Rev. B 5, 2897 (1972). See Eq. (134).
[10] D. S. Montgomery, J. Phys. C 16, 2923 (1983).
[11] S. Krishnamurthy, A.-B. Chen, and A. Sher (unpublished).
[12] S. M. Sze, *Physics of Semiconductors*, 2nd ed. (Wiley–Interscience, New York, 1981).
[13] V. A. Johnson and K. L. Horowitz, Phys. Rev. 79, 176 (1950); 79, 409 (1950).
[14] H. Jones, Phys. Rev. 81, 149 (1951).

Binding Energy and Spectral Width of Si $2p$ Core Excitons in Si_xGe_{1-x} Alloys

S. Krishnamurthy and A. Sher
Physical Electronics Laboratory, SRI International, Menlo Park, California 94025

and

A.-B. Chen[a]
Department of Physics, Auburn University, Auburn, Alabama 36849

(Received 24 January 1985)

A calculation is presented to explain the anomalous experimental behavior of the Si $2p$ core-exciton binding energy and linewidth in Si_xGe_{1-x} alloys. The observed minimum in the linewidth near $x \simeq 0.15$ can be explained as the result of a competition between intrinsic broadening due to screening and extrinsic alloy broadening. For pure Si, the binding energy is estimated to be 0.15 ± 0.05 eV and the width is shown to be smaller than that observed at $x \simeq 0.15$.

PACS numbers: 71.55.Fr, 71.35.+z, 78.70.Dm

Until 1984, the Si $2p$ core exciton was believed to have an anomalously large binding energy.[1-10] Later, Newman and Dow[11] proposed a radically different picture in which the Si $2p$ core exciton is in fact a resonance with a negative binding energy. They further predicted that the exciton binding energy remains negative throughout most of the Si_xGe_{1-x} alloy composition range, except near $x \simeq 0.20$ where it becomes positive. In a recent experiment,[12] Bunker *et al.* found an anomalous sharpening of the exciton spectra near $x = 0.15$; the data were interpreted to support the Newman-Dow point of view. Yet the most recent experiment[13] still suggests a positive value for the binding energy E_b^0 in silicon.

In this Letter, we present a calculation that offers a plausible resolution to the above problem. In our theory, the calculated Si $2p$ core-exciton binding energy $E_b(x)$ and the linewidth $\Delta(x)$ in Si_xGe_{1-x} alloys are sensitively dependent on the parameter E_b^0. A comparison of the calculated $\Delta(x)$ with the experiment[12] suggests a positive value 0.15 ± 0.05 for E_b^0. The anomalous experimental spectrum[12] near $x \simeq 0.15$ is explained as a result of a competition between an intrinsic broadening Δ_I due to screening and an extrinsic alloy broadening Δ_A. In the present theory, there is no need to suppose that the exciton suddenly changes its character from an extended effective-mass-like state to a deep localized state.

We need to calculate E_b and $\Delta = \Delta_I + \Delta_A$ as a function of alloy concentration x. The calculations are based on a quantitative coherent-potential-approximation (CPA) band structure. Details of the CPA calculations will be presented elsewhere. Below, we briefly discuss a Green's-function method for calculating E_b and Δ_A.

The one-particle effective Green's function in CPA takes the form

$$G(E) = [E - \bar{H} - \Sigma(E)]^{-1}, \tag{1}$$

where $\bar{H}$ is the virtual-crystal approximation Hamiltonian and $\Sigma(E)$ is the self-energy. The site-diagonal Green's function is denoted as

$$F_\alpha(E) = \langle \phi_\alpha | G(E) | \phi_\alpha \rangle, \tag{2}$$

where ϕ_α is a localized orbital of specified symmetry. Here we only need to consider $\alpha = s$ for A_1 symmetry. The corresponding function in pure Si is denoted as $F_s^0(E)$. Following the theoretical treatment of deep substitutional-impurity levels,[8] the core-exciton level for pure Si is determined by

$$F_s^0(E) = (V - E_s^{Si})^{-1}, \tag{3}$$

where E_s^{Si} is the site potential seen by an s electron in bulk silicon, and V is a central-potential parameter. For a chosen value of V_b, Eq. (3) can be solved for E, and vice versa. Then $E_b^0 = E_c^0 - E$, where E_c^0 is the conduction-band edge in pure silicon. Because of the uncertainties in the value of experimental E_b^0 and theoretical V, we treat E_b^0 (or V) as a parameter. The binding energy E_b in a Si_xGe_{1-x} alloy can be calculated by solving

$$F_s(E) = [V - \bar{E}_s - \Sigma_s(E)]^{-1}, \tag{4}$$

where

$$\bar{E}_s = xE_s^{Si} + (1-x)E_s^{Ge}. \tag{5}$$

Then E_b is given by

$$E_b = E_c - E. \tag{6}$$

The calculated values of the conduction-band edge and the exciton level measured relative to the top of the valence band are plotted in Fig. 1. The band gap increases with x with a slope discontinuity at ≈ 0.15. The dashed lines *a*, *b*, and *c* represent exciton levels obtained with $E_b^0 = 0.1$, 0.15, and 0.30 eV, respectively. The binding energy E_b is also an increasing function of x, with a slope discontinuity near $x \approx 0.15$. The CPA introduces a slight bowing in E_g and E_b.

Strinati[10] has calculated the variation of Δ_I with E_b

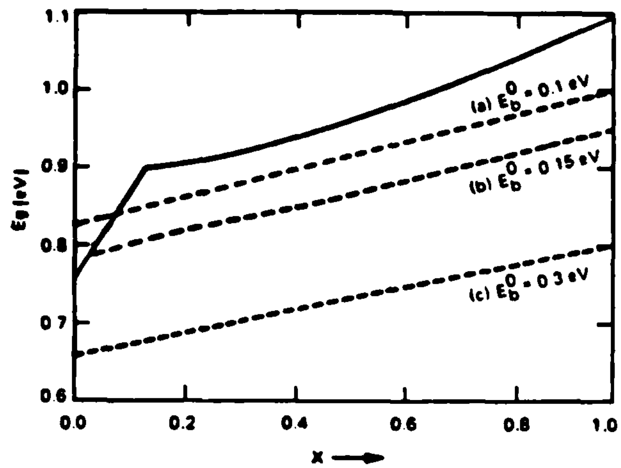


FIG. 1. Variation of the band gap (solid line) and the Si $2p$ core-exciton level with x in $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ alloys. The energy is measured from the top of the valence band. The dashed curves a, b, and c represent exciton levels calculated with $E_b^0 = 0.1$, 0.15, and 0.3 eV, respectively.

by replacing the short-range Coulomb potential with a spherical square well of variable depth and a screened Coulomb tail. Strinati's results can be used to estimate Δ_l corresponding to the calculated E_b. Δ_l decreases rapidly with E_b, then saturates for larger E_b.

The contribution to the natural linewidth from the alloy broadening is calculated by a consideration of the electron part of the exciton wave function, ψ_s. The ψ_s is expanded in a linear combination of the s part of the conduction-band wave functions ϕ_n^s:

$$\psi_s(\mathbf{k}) = \sum_{n,s} C_{ns}\phi_n^s(\mathbf{k}). \tag{7}$$

We found that alloy scattering is only moderate and s scattering is dominant; thus, the alloy broadening $\Delta_A(E)$ is well approximated by

$$\Delta_A(E) \approx x(1-x)\delta_s^2 \operatorname{Im} F_s(E), \tag{8}$$

where δ_s is the difference between E_s^{Si} and E_s^{Ge}. Hence, the alloy-broadening contribution to Δ is related to the alloy broadening of the band states, $\Delta_n(\mathbf{k},\mathbf{E})$:

$$\begin{aligned}\Delta_A &= \frac{1}{N}\sum_{\mathbf{k}} \langle \psi_s(\mathbf{k})|\Delta_A(E)|\psi_s(\mathbf{k})\rangle \\ &\approx \frac{1}{N}\sum_{\mathbf{k}}\sum_{n} C_{n,s}^2 \Delta_n(\mathbf{k},E) \\ &= \int \rho_s(E)\Delta_A(E)\,dE \\ &= x(1-x)\delta_s^2\pi \int \rho_s^2(E)\,dE. \end{aligned} \tag{9}$$

The integral in Eq. (9) is evaluated numerically.

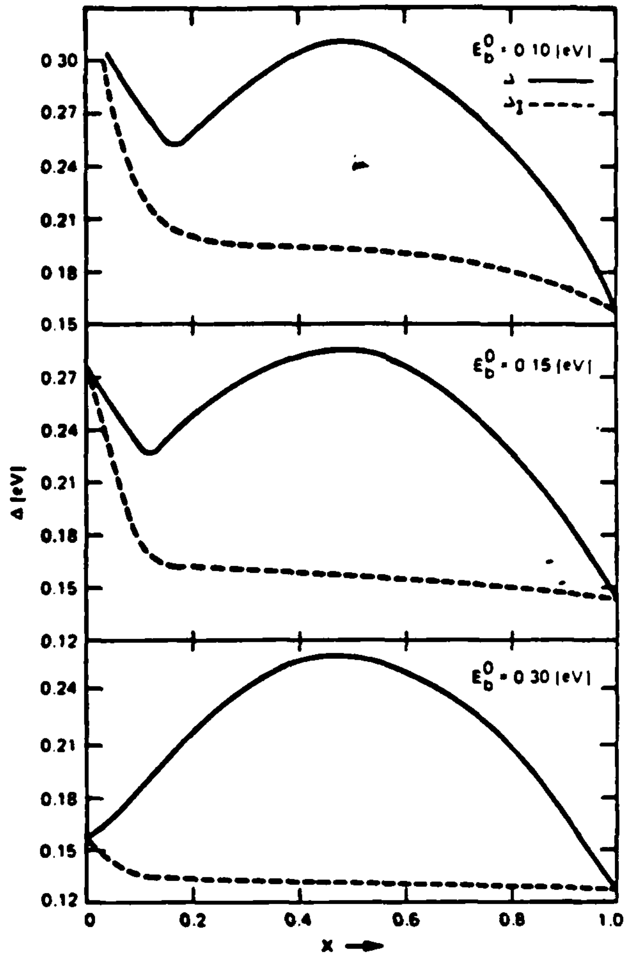


FIG. 2. Variation of Δ (solid lines) and Δ_l (dashed lines) with x for three E_b^0 values.

The calculated Δ, which is the sum of Δ_l and Δ_A, is plotted against x in Fig. 2 for three values of E_b^0. In all three panels, the dashed curve represents Δ_l and the solid line represents Δ. It is seen from Fig. 1 that the exciton level follows the X edge of the conduction band. Hence the binding energy E_b, relative to the conduction band edge, remains almost constant (for a given E_b^0) until the minimum switches from the X edge to the L edge. Because of the change in the slope of E_g, E_b decreases rapidly when L becomes the minimum. Correspondingly, Δ_l varies slowly until the X to L crossover and then increases rapidly. This feature is clearly seen in Fig. 2.

For $E_b^0 = 0.15$, the Δ_l and Δ_A are comparable near $x = 0.50$, and Δ_l dominates for all small x and large x. These two competing mechanisms give a relative minimum near $x \approx 0.15$, a broader maximum near $x \approx 0.50$, and a smaller minimum for pure silicon. As E_b^0 is decreased, the relative minimum is shifted to larger x, e.g., the minimum shifts to $x = 0.20$ for $E_b^0 = 0.10$ eV. For $E_b^0 = 0.15$ eV, the position of the relative minimum is in agreement with the experiment.[12] (By measuring the relative width at $x \simeq 0.15$ to that $x = 1$, one can make a better estimate of E_b^0.) To correlate the theory with experiment, the calculated $1/\Delta^2$ is compared with the measured[12] $(\Delta\mu)^{-1}(d\mu/dE)_{\max}$ in Fig. 3, where $\Delta\mu$ is the edge step and $(d\mu/dE)_{\max}$ are the maximum values of the derivative of absorption spectra with respect to photon energy. Because the experimental values are given in arbitrary units, the values are normalized to agree at $x = 0.5$. The observed anomalous behavior near $x = 0.15$ and the qualitative x dependence in that region is clearly replicated by the theory. However, the calculation predicts a larger maximum at $x = 1$. It would be interesting to have experiments that cover the entire range of x to further test this prediction.

For larger values of E_b^0, the calculated E_b is also large and hence Δ_l decreases slowly with x. Because the broadening is determined mainly by Δ_A, the linewidth is expected to be small for $x = 0$ and $x = 1$ only; this occurs for $E_b^0 = 0.30$ eV. For negative values of E_b^0, E_b remains negative for all values of x. Accordingly, the linewidth is broad for all x, and there would be no such anomaly as in Fig. 3.

The calculations presented in this Letter are slightly different from alchemy approximations.[8] We treat the central-cell potential V as a parameter and narrow its range from other considerations. We examine values of $V - E_s^{\text{Si}}$ of −8.49, −7.09, and −6.56 eV, corresponding to E_b^0 values of 0.30, 0.15, and 0.10 eV, respectively. If the strict alchemy approximation were taken, the value of $V - E_s^{\text{Si}}$ would be $E_s^P - E_s^{\text{Si}} = -4.59$ eV in the tight-binding approximation, and a negative E_b^0 (~ -0.10 eV) would be obtained. When long-range interactions are included, however, the above resonance state becomes a shallow donor level, which is the experimental situation for a P impurity in Si. Our results suggest that V for core excitons in Si is deeper than those implied by alchemy approximations. However, if we use the alchemy approximation as a means of scaling, the value of V for Ge $3p$ core excitons should be deeper than that for Si $2p$ core excitons. Hence, the curve corresponding to $E_b^0 = 0.30$ in Fig. 2 should be a reasonable estimate for Ge $3p$ core-exciton binding energy in alloys. Therefore, we do not expect to see an anomalous behavior of Δ in alloys for this case.

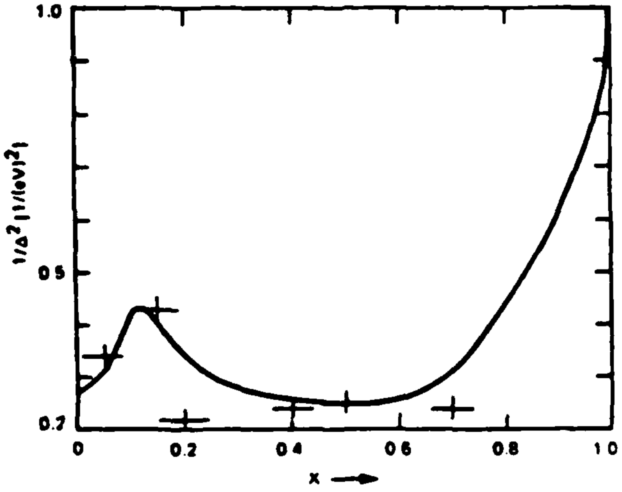


FIG. 3. The calculated $1/\Delta^2$ values (solid line) compared with the experimental results (marks) from Ref. 12. The value of calculated $\Delta(\text{Si})$ is 0.127 eV. The experiment is normalized to the theory at $x = 0.5$.

In summary, the present calculations of the Si $2p$ core-exciton binding energy and linewidth suggest that the exciton level is about 0.15 ± 0.05 eV below the conduction-band edge for pure Si. It follows the X edge for $x > 0.15$ in the $\text{Si}_x\text{Ge}_{1-x}$ alloys, and E_b may eventually reach zero in the dilute limit $x \to 0$. Our value for E_b^0 represents the lower end of the previous measurements,[1–6] but is in good agreement with a recent experimental[13] value of 0.120 ± 0.03 eV. By considering the intrinsic linewidth and the alloy broadening, we can explain the observed relative minimum in the linewidth near $x \simeq 0.15$, without requiring a sudden change of the exciton character. On the basis of this calculation, we expect the corresponding width in pure Si to be even smaller than that observed near $x = 0.15$. We further argue that the binding energy of Ge $3p$ core excitons should be larger than that of Si $2p$ core excitons and there should be no anomaly in the Ge $3p$ linewidth in alloys.

This work was supported in part by the U.S. Department of Defense Advanced Research Projects Agency, under Contract No. MDA 903-83-C-0108 and the Air Force Office of Science Research through Grant No. AFOSR-84-0282. One of us (A.-B.C.) would like to thank Professor W. E. Spicer for his hospitality at Stanford University.

---

[a] Visiting professor at Stanford University.

[1] W. Eberherdt, G. Kalkoffen, C. Kunz, D. E. Aspnes, and M. Cardona, Phys. Status Solidi (b) **88**, 135 (1978).

[2] M. Altarelli and D. L. Dexter, Phys. Rev. Lett. **29**, 1100 (1972).

[3] F. C. Brown, R. S. Bachrach, and M. Skibowski, Phys. Rev. B **15**, 4781 (1977).

[4] R. S. Bauer, R. S. Bachrach, J. C. McMenamin, and D. E. Aspnes, Nuovo Cimento **39B**, 409 (1977).

[5] G. Margaritondo and J. E. Rowe, Phys. Lett. **59A**, 464 (1977).

[6] G. Margaritondo, A. Franciosi, N. G. Stoffel, and H. S.

Edelman, Solid State Commun. 36, 298 (1980).

[7] C. Kunz, J. Phys. (Paris), Colloq 39, C4-119 (1978).

[8] H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B 24, 6010 (1981).

[9] M. Altarelli, Phys. Rev. Lett. 46, 205 (1981).

[10] G. Strinati, Phys. Rev. Lett. 49, 1519 (1982).

[11] K. E. Newman and J. D. Dow, Solid State Commun. 50, 587 (1984).

[12] B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. 53, 2157 (1984).

[13] S. E. Schnatterly and R. D. Carson, Bull. Am. Phys. Soc. 30, 416 (1985).

IV.12

Band structures of Si_xGe_{1-x} alloys

Srinivasan Krishnamurthy and A. Sher

SRI International, Menlo Park, California 94025

and

A.-B. Chen

Physics Department, Auburn University, Auburn, Alabama 36849

ABSTRACT

Starting from realistic band structures of the constituent materials, the electronic structure of Si_xGe_{1-x} alloys are obtained in the coherent potential approximation (CPA). Various quantities, including the bowing parameter of the fundamental gap and the energies of several optical gaps, the masses, and the linewidths of the E_0 and E_1 transitions, are calculated on the basis of both diagonal and off-diagonal CPA. All of the band-energy and line-width predictions are in good agreement with experiments. Furthermore, the theory yields alloy-scattering-limited electron-drift mobility in qualitative agreement with experimental results.

I. INTRODUCTION

Semiconductor alloys offer the freedom to design material properties by choosing appropriate alloy constituents. In some cases, the physical properties of the alloys can be quite different from those of the constituents.(1-4) In recent years, there has been a renewed interest in Si_xGe_{1-x} alloys(5) and superlattices.(5-10) Because silicon is the most technologically advanced semiconductor, the results on SiGe systems have many potential applications.

The lattice constants of silicon and germanium differ by ~4%. Hence, the strain introduced in the formation of Si_xGe_{1-x} alloys can affect the band structure(10) and the transport properties.(8) Prior authors used virtual crystal approximation (VCA)(11,12) and coherent potential approximation (CPA)(13) to study the band structure and related properties. Either because of less accurate band structures of the constituent materials, or because of the approximations involved in the alloy formalism, these calculations predicted only trends of specific quantities, not quantitatively accurate results. Because the s-state site potentials (ϵ_s) for silicon and germanium differ by approximately 1.5 eV, VCA cannot accurately describe effective masses and other finer details of the band structure. Because of the use of poor basis functions, earlier CPA work(13) predicted alloy broadening of conduction band states substantially differing from experiment. The purpose of this paper is to correct these flaws and treat transport phenomena.

Because of a substantial difference between the site potentials and lattice constants of silicon and germanium, we incorporated both chemical and structural disorder in the calculation of the electronic structure of Si_xGe_{1-x} alloys. Thus, both diagonal and off-diagonal CPA are included in the predicted band structure and related quantities. Parts of the band structure have been used to study the Si-2p core-exciton(14) and the alloy mobilities.(15) A comprehensive report of the calculations and results is presented here.

The rest of the paper is arranged as follows. The detailed procedure of fitting silicon and germanium band structures is given in Section II. The VCA, CPA, and off-diagonal CPA calculations are described in Section III. The results and interpretation of the alloy band structures and mobility are given in Section IV.

II. BAND STRUCTURE BASIS

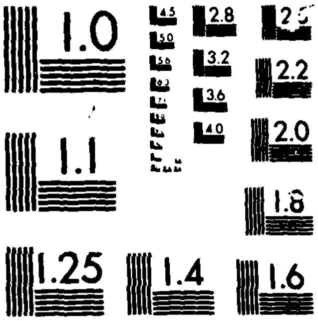
In order to derive an accurate alloy band structure, one must start from a realistic band structure of the constituent materials. Chen and Sher have developed a method(16) following a prescription of Kane(17) and Chadi(18) that includes all long-range interactions, and then fine tuned the band structure with an adjustable local Hamiltonian. Because the details have already been published,(16,19) the underlying method will be presented here in brief.

Gaussian orbitals of the type α (α can be s, p_x, p_y, or p_z) for each sublattice in a cell are used to construct the corresponding Bloch basis. In this basis set, the overlap matrix and the Hamiltonian derived from empirical pseudopotentials can be calculated.(17,18) It is possible to cast the problem in a basis set of Gaussian orbitals in which, in crystal units (cu), the same exponential factors apply for all III-V compounds.(19) In this universal basis, the overlap matrix and the kinetic energy matrix are same for all III-V compounds. Then, by a unitary transformation, the basis set is orthonormalized.(20) The Hamiltonian in this new basis set is denoted $H_0(\vec{k})$. The band structure resulting from this method reproduces the results of elaborate band structure calculations within a few percent throughout the Brillouin zone (BZ). To establish accurately certain important band structure features adjacent to the gap, an extra small 8x8

AD-A166 794 HGCDTE SURFACE AND DEFECT STUDY PROGRAM(U) SANTA BARBARA RESEARCH CENTER GOLETA CALIF J A WILSON ET AL. JUL 85 SBRC-60070 MDA903-83-C-0108 3/3

UNCLASSIFIED F/G 20/12 NL

END
FILMED



MICROCOP CHART

Hamiltonian matrix $H_1(\overline{k})$ is added to $H_0(\overline{k})$. This $H_1(\overline{k})$ has the form of a tight-binding (TB) *Hamiltonian*, in which only the nearest neighbor interactions are included, and stimulates the effect of nonlocal pseudopotentials and an expanded orbital set. The total Hamiltonian, $H(\overline{k})$, in this orthonormalized basis set is diagonalized to obtain the band energies and the corresponding wave functions.

Following this procedure with the same exponential factor $\beta = 0.26$ in the Gaussian orbitals for both silicon and germanium, the matrix $H_0(\overline{k})$ is obtained. For silicon and germanium, H_1 contains 6 adjustable parameters: namely, the corrections to the term values Δ_s and Δ_p and to the nearest-neighbor interactions V_{ss}, V_{sp}, V_{xx}, and V_{xy}. The values of Δ_s, Δ_p, V_{ss}, and V_{xx} are determined from fitting the three experimental energy gaps(21-28) at $\Gamma(\overline{k} = 0)$: $\Gamma_{2'} - \Gamma_1$, $\Gamma_{15} - \Gamma_{25'}$, and $\Gamma_{2'} - \Gamma_{25'}$, and the photoelectric threshold (PT) values -5.07 and -4.80 eV for silicon and germanium respectively.(29) The remaining parameters V_{sp} and V_{xy} are obtained from the experimental values(21-28,30) of the gaps $X_{1c} - X_{1v}$ and $L_{1c} - L_{3'v}$. Some adjustments in these input quantities are made to obtain an overall good band structure with more accurate effective masses. Table I lists the empirical pseudopotential form factors and the parameters used to obtain the band structure. The calculated band structures and experimental values are given in Table II. From Table II, one can see that an excellent fit to the silicon and germanium band structure is obtained: All the calculated values lie within the experimental uncertainties. The optical difference between L_{1c} and L_{3v}, $\Gamma_{25'}$ and Γ_{15} are in excellent agreement with the known optical transition values.

Although the calculated effective transverse masses agree very well with experiment, the effective longitudinal mass for germanium is less than the experimental value. This is due mainly to our attempt to have a common β and the choice of local pseudopotentials, causing $H_0(\overline{k})$ to be the same in crystal units (cu) for both germanium and silicon. Because of the common $H_0(\vec{k})$, the alloy disorder is contained in these adjusted parameters. This H_0 would also be useful for the interface and superlattice(30) problems. If we grant ourselves the freedom to adjust V_{xy}, longitudinal effective mass in germanium can be fitted to the experimental value. When V_{xy} is changed, the L_{1c}, $L_{3'v}$ will also change. We have chosen not to do this because little is gained for the extra complexity. For an indirect gap semiconductor, the important effective mass used in transport studies is the conductivity mass, $3\left(1/m^*_{el} + 2/m^*_{et}\right)^{-1}$. Because $m^*_{el} >> m^*_{et}$ in germanium, m^*_e will not be much different if a less accurate value of m^*_{el} is used. Moreover, the Si_xGe_{1-x} alloys which have potential device applications are in the silicon-rich region, where the effective mass at the L edge is not expected to affect the further studies.

It is important to note that an excellent fit to the experimental values can be obtained with only seven adjustable parameters (β, Δ_s, Δ_p, V_{ss}, V_{sp}, V_{xx}, V_{xy}). with β being universal in cu. The calculated band structure of silicon and germanium are shown in Fig. 1(a) and 1(b) respectively. The characteristic indirect gaps are clearly seen. These band structures compare favorably with the best results available, and, in contrast with those obtained in the usual empirical TB approaches, produce good conduction bands.

III. ALLOY CALCULATION

A. *VCA*

Because we have the same $H_0(\vec{k})$ matrix for both silicon and germanium, it is only the H_1 matrices of the constituents in scaled VCA which distinguishes them. In this approximation, the diagonal elements of the alloy Hamiltonian $\bar{H}(\vec{k})$ are simply the concentration-weighted average of the corresponding elements of the pure silicon and germanium Hamiltonians, whereas the off-diagonal elements of $\bar{H}(\vec{k})$ are obtained by assuming a $1/d^2$ dependence. $\bar{H}(\vec{k})$ can be diagonalized to obtain the VCA band structure for various concentrations, x. The VCA band structures for $x = 0.1$ and 0.5 are shown in Fig. 1(c) and 1(d) respectively.

B. *Diagonal CPA*

An earlier work on CPA band structure of SiGe alloy(13) is based on a local but energy-dependent pseudopotential approximation. While the value of the scattering potential parameter was 1.49 eV, close to our value, the calculation predicted too-large linewidth in in the E_0 spectrum and essentially no effect on the electron mobility. With the availability of a set of good basis functions and constituent band structures, more realistic band structures of the alloy can be obtained.

In the current model, we have a TB Hamiltonian, which contains matrix elements to all ranges. The simplest alloy model is to assume that the important disorder resides only in the diagonal matrix elements, ϵ_s and ϵ_p. In our model, the ϵ_s^{Si} and ϵ_s^{Ge} differ by

1.46 eV, whereas ϵ_p^{Si} and ϵ_p^{Ge} differ by 0.21 eV. For the present, we neglect the disorder in the off-diagonal element. Mathematically, we have

$$H_{alloy} = \bar{H} + \sum_{\vec{T}} V_{\vec{T}} \quad , \tag{1}$$

where $\vec{T}$ is a fcc lattice vector identifying a site, and $V_{\vec{T}}$ is the 8x8 diagonal matrix with elements $U_s = \epsilon_s - \bar{\epsilon}_s$, $U_p = \epsilon_p - \bar{\epsilon}_p$ in the orthonormal local orbitals $|\vec{T}j\alpha>$; j denotes the two atoms in the unit cell labeled $\vec{T}$, α represents s or p symmetry, and $\bar{\epsilon}_s$ and $\bar{\epsilon}_p$ are the concentration-weighted average values of s and p silicon and germanium term value energies.

The one particle alloy Green's function is defined as

$$G_{alloy}(Z) = \frac{1}{Z - H_{alloy}} \quad . \tag{2}$$

We are after the configurational average of this Green's function, which, in effective medium theory, is replaced by an effective Green's function G,

$$G(Z) = \frac{1}{Z - \bar{H} - \Sigma(Z)} \quad , \tag{3}$$

where Σ is the self energy. In CPA, we can now write $\Sigma = \sum\limits_{T} (\Sigma_l)$, with Σ_T being an 8x8 matrix in the basis $|Tj\alpha>$ having the form

$$\Sigma = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \quad , \tag{4}$$

where

$$A = \begin{bmatrix} \Sigma_s & 0 & 0 & 0 \\ 0 & \Sigma_p & 0 & 0 \\ 0 & 0 & \Sigma_p & 0 \\ 0 & 0 & 0 & \Sigma_p \end{bmatrix} \ .$$

Here Σ_s and Σ_p are the s and p part of the self energy. The Σ_s and Σ_p are determined from the conditions that the average atomic t-matrix with respect to the CPA Green's function G is zero. With our assumption of Σ, the matrix equation $<t>=0$ reduces to two coupled equations $<t_s> = 0$ and $<t_p> = 0$, where the average is the concentration-weighed average $<Q> = xQ^{Si}+yQ^{Ge}$, and the t is defined as

$$t_\alpha^\beta = (U_\alpha^\beta - \Sigma_\alpha)\,[1 - F_\alpha(U_\alpha^\beta - \Sigma_\alpha)]^{-1}$$

$$(\alpha = s \text{ or } p,\ \beta = \text{Si or Ge}) \quad . \tag{5}$$

In the above expression, F_α is the diagonal matrix element of G in the local basis $F_\alpha(Z) \equiv <Tj\alpha \mid G(Z) \mid Tj\alpha>$. The Σ_s and Σ_p are coupled because F_s and F_p contains both Σ_s and Σ_p.

An iterative average-t-matrix (IATA) procedure(19) is employed to solve the CPA equation. This procedure improves Σ_a upon a guessed solution Σ_a° through the following equations

$$\Sigma_s = \Sigma_s^\circ + <t_s^\circ> [1 + F_s^\circ <t_s^\circ>]^{-1} \quad ,$$

$$\Sigma_p = \Sigma_p^\circ + <t_p^\circ> [1 + F_p^\circ <t_p^\circ>]^{-1} \quad , \qquad (6)$$

where $<t_a^\circ>$ and F_a° are similar to those in Eq. (5) except that Σ_a° now replaces Σ_a. The most time-consuming calculation is then the computation of the local Green's functions F_s° and F_p°, given by the BZ summation; e.g.

$$F_s^\circ(Z) = \frac{1}{N} \sum_{\vec{k}} \left(\frac{1}{Z - \overleftrightarrow{H}(\vec{k}) - \Sigma_s^\circ} \right)_{11} \quad ,$$

where the inverse of an 8x8 matrix is involved for every $\vec{k}$. This can be simplified by observing that Σ_a° has the same form as Σ in Eqs. (4) and (5) and that the 4x4 A matrix can be written as $A = \Sigma_p \overleftrightarrow{1} + (\Sigma_s - \Sigma_p)J$, where $\overleftrightarrow{1}$ is the identity matrix and

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad .$$

Defining the matrix

$$\sigma = (\Sigma_s^\circ - \Sigma_p^\circ) \begin{pmatrix} J & O \\ O & J \end{pmatrix}$$

$F_s^\circ(Z)$ and $F_p^\circ(Z)$ can now be calculated from

$$F_s^\circ(Z) = \frac{1}{N} \sum_{\bar{k}} g_{11}(\bar{k},Z) \quad ,$$

$$F_p^\circ(Z) = \frac{1}{3N} \sum_{\bar{k}} [g_{22}(\bar{k},Z) + g_{33}(\bar{k},Z) + g_{44}(\bar{k},Z)] \quad , \tag{7}$$

where

$$g = g^\circ + g^\circ (1 - \sigma g^\circ)^{-1} \quad . \tag{8}$$

with

$$g_{\alpha\beta}^{(o)}(\bar{k},Z) = \sum_n \frac{U_{\alpha n}(\bar{k}) U_{\beta n}^*(\bar{k})}{Z - \epsilon_n(\bar{k}) - \Sigma_p^\circ} \quad . \tag{9}$$

In Eq. (9), $\epsilon_n(\vec{k})$ is the band energy in VCA and $\{U_{\alpha,n}(\vec{k})\}$ satisfy the following Eigen equation:

$$\sum_{\beta} \bar{H}_{\alpha,\beta}(\vec{k})U_{\beta,n}(\vec{k}) = \epsilon_n(\vec{k})U_{\alpha,n}(\vec{k}) \quad .$$

Because σ matrix has only two nonzero elements, the matrix inversion in Eq. (8) is obtained analytically.

A substantial reduction in computer time is made possible by using an analytical continuation method.(32) In this method, Σ_s and Σ_p are calculated as a function complex Z, and then, using the analytical properties of the self energy and Green's functions, they are interpolated for real Z. Because the functions Σ_s, Σ_p, and G are smooth for complex Z, the CPA iterations and BZ integrations can be carried out with substantially less computer time.

For the concentration x = 0.10 and x = 0.50, the L and X(Δ) gap respectively are preferred. The L to X(Δ) crossover takes place near x ≈ 0.15. The CPA correction to L and X edges at x = 0.10, 0.15, and 0.50 should be good enough to study the quantitative variation of band gap in Si_xGe_{1-x} alloys. Hence, the calculations are carried out for these three cases. In addition, because the experimental results are available for x = 0.109, CPA calculations are also done here for comparison. As expected, Σ_s is much larger than Σ_p for all the cases. The self energy as a function of energy is plotted in Fig. 2 for an x = 0.5^ alloy.

C. *Off-Diagonal CPA*

As mentioned earlier, silicon and germanium differ in their lattice constant by ~4%. In order to include the effect of the structural disorder, the CPA calculation is repeated next with off-diagonal (OD) disorder included. By an application of the molecular coherent potential approximation (MCPA)(33), Hass et al. included OD disorder in the CPA calculation of an $A'_x A''_{1-x}B$ semiconductor alloys.(34) Assuming that B atoms occupy the sites of an ordered zinc-blende virtual lattice, they modeled the dominant structural effect as the difference in A' –B and A'' –B hopping matrix elements. Hence, the chemical and structural disorder effects are treated as random variations of ϵ^A, V_1^A, V_2^{AB}, where the symbols have their usual meaning.(35)

The extension of the method to Si_xGe_{1-x} alloys is not straightforward, mainly because silicon and germanium can occupy both sublattices; hence, there can be no ordered virtual lattice in this case. If we choose the tetrahedral unit cell as the molecular unit for MCPA, we see that the disorder is not cell diagonal. However, by choosing an appropriate basis set, we can make the intercell interaction be the highest order effect. We start with a hybrid basis $|\bar{l}h\rangle_i$ obtained from the sp^3 hybrid orbitals.(35) The hybrids 1 through 4 ($i = 1$-4) are obtained from orbitals centered on a sublattice I-site, and the states 5 through 8 ($i = 5$-8) are those from the orbitals located at the four nearest neighbor sites on sublattice II. The Bloch basis states, corresponding to A_1, T_2 symmetries, $|k\bar{h}\rangle_i$ located on an I-site ($i = 1$-4) and II-site ($i = 5$-8) are obtained from the corresponding hybrid states given by the relation

$$|\bar{l}h\rangle_i = \sum_{j=1}^{8} C_{ij}\ |\bar{l}h\rangle_j \quad , \tag{10}$$

where

$$C = \begin{pmatrix} c_1 & 0 \\ 0 & c_1 \end{pmatrix}$$

and

$$c_1 = 1/2 \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} .$$

An explicit definition of these orbitals can be found in Ref. (20). In this new basis, the self energy Σ at the given site takes the form

$$\Sigma = \begin{pmatrix} \sigma_0 & \sigma_2 \\ \sigma_2 & \sigma_0' \end{pmatrix} , \tag{11}$$

where

$$\sigma_0 = \begin{pmatrix} \Sigma_s & 0 & 0 & 0 \\ 0 & \Sigma_p & 0 & 0 \\ 0 & 0 & \Sigma_p & 0 \\ 0 & 0 & 0 & \Sigma_p \end{pmatrix} , \tag{12}$$

$$\sigma_2 = \begin{pmatrix} \Sigma_{2s} & 0 & 0 & 0 \\ 0 & \Sigma_{2p} & 0 & 0 \\ 0 & 0 & \Sigma_{2p} & 0 \\ 0 & 0 & 0 & \Sigma_{2p} \end{pmatrix} ,$$

$$\sigma_0' = \begin{pmatrix} \Sigma_b' & 0 & 0 & 0 \\ 0 & \Sigma_b' & 0 & 0 \\ 0 & 0 & \Sigma_b' & 0 \\ 0 & 0 & 0 & \Sigma_b' \end{pmatrix} ,$$

and

$$\Sigma_b' = \frac{1}{4} (3\Sigma_p + \Sigma_s) . \qquad (13)$$

The self energies can be obtained again from IATA iteration procedure.

$$\Sigma = \Sigma_0 + <<T>> [1 + F<<T>>]^{-1} , \qquad (14)$$

where

$$F_{ij} = \frac{1}{N} \sum_{n} \sum_{\bar{k}} \frac{Q^{+}_{in} Q_{nj}}{Z - E_n(Z,\bar{k})} , \qquad (15)$$

with

$$(\bar{H}(\bar{k}) + \Sigma)Q = E(Z,\bar{k})Q ,$$

$$<<T>> = x <T_A> + y <T_B> ,$$

and

$$<T_A> = x^4\, t^A{}_{A_4} + 4x^3y\, t^A{}_{A_3B} + 6x^2y^2\, t^A{}_{A_2B_2} + 4y^3x\, t^A{}_{AB_3} + y^4\, t^A{}_{B_4}\ ,\ A \equiv \mathrm{Si}\ , \qquad (16)$$

with a similar expression for $<T_B>$. Physically, for a given A atom at the center, the other four atoms in the molecular unit cell can be all A atoms, 3 A atoms and 1 B atom, 2 of each 1 A atom and 3 B atoms, or all four can be B atoms. $<T_A>$ or $<T_B>$ represents the configuration-averaged t-matrices, and $<<T>>$ is the concentration-weighted average of the configuration. By exploiting the symmetry, as seen in Eq. (11), one can reduce this problem to solving two 2x2 coupled matrix equations. Eq. (14) can be iterated to obtain Σ_s, Σ_p, Σ_{2s}, and Σ_{2p}. After every iteration, we get a new set of Σ_s, Σ_p, Σ_{2s}, Σ_{2p}, and Σ_h' : The new set has not been tested to see if Σ_h' is still given by Eq. (13). In our calculation, we did not iterate to obtain a new Σ_h' ; instead we fixed it by the relation given in Eq. (13). The error introduced by this approximation is expected to be very small. As in the case of diagonal CPA, the computation can be substantially reduced by the method of analytical continuation.(32)

IV. DISCUSSION

A. E_0 *and* E_1 *Optical Transition*

The VCA values of E_0 ($\Gamma_{2'c} - \Gamma_{25v}$) and E_0' ($\Gamma_{15c} - \Gamma_{25v}$) and their measured values are plotted as a function of x in Fig. 1(e). Because the measurements(36) are made at room temperature, the experimental values are smaller than the values calculated from the zero temperature band structure. Inclusion of the relativistic effects, which are not present in our calculations, is expected to form a more accurate basis for comparison with the experiments. As seen from Fig. 1(e), the theoretical and the experimental values both have a linear variation with x. Similar calculations of E_1 $(L_{1c} - L_{3' v})$ also have a linear variation on x and are in qualitative agreement with experiments.(36)

From the CPA self energies Σ_s and Σ_p, it is straightforward to calculate the correction to the VCA bands. The calculated complex band structure is plotted for $x = 0.50$ in Fig. (3). The CPA corrections are shown only in the vicinity of the band gap. The shaded portion represents the half-width of that energy state. Because s-scattering is dominant in these alloys, we see that the major disorder lies in the conduction band. The topmost valence band, with its rich p-content is least affected. The CPA band structure is used to calculate the E_0 and E_1 peak positions for $x = 0.10$, 0.109, 0.15, and 0.50 concentrations. The calculations and the data from Reference (36) show a small bowing that is not seen on the scale of Fig. 1(e).

The self energies Σ_s, Σ_p, Σ_{2p}, and Σ_{2s} are calculated in MCPA for the $x = 0.50$ alloy. As in the case of CPA, the self energies associated with s-symmetry are much larger than the ones associated with the p-symmetry. While Σ_{2s} is found to be very small, Σ_{2p} is at least an order of magnitude smaller—almost zero. However, the $\mathrm{Im}\Sigma_s$ obtained by CPA and MCPA differ considerably. As seen from Fig. 4, the difference increases as one goes away from the band edge. Therefore, the lifetime associated with the alloy disorder is decreased by the inclusion of OD structural disorder. In addition, the OD disorder lowers the conduction band, introducing an extra bowing. The E_0 and E_1 values are reduced by 27 and 12 meV respectively. The VCA, CPA, and MCPA values of E_0 and E_1 are listed in Table III.

The half-width of the alloy states is calculated from the imaginary part of the CPA self energies. The half-width of the lowest-lying conduction band of $Si_{0.5}Ge_{0.5}$ alloy is plotted in Fig. 5 as a function of K_x in the [100] direction. The calculated half-width is 186 meV for the $\Gamma_{2'}$ state and decreases to zero at the band edge. Because of the negligible alloy broadening of the topmost valence band state, the half-width corresponding to the E_0 transition, $\Delta(E_0)$, is 186 meV, which is approximately one-half of the previously published CPA results.(13) The CPA value of the half-width corresponding to the E_1 transition, $\Delta(E_1)$, is 31 meV. Because of the increase in the imaginary part of the self energies, the MCPA values of the half-widths of the E_0 and E_1 transitions are 206 meV and 32 meV respectively. Because the complete E_0 peak is not shown in the published electroreflactance spectrum(36), it is difficult to estimate the corresponding half-width. However, one can conclude from the spectrum of the $x = 0.458$ alloy that the half-width of the E_1 transition is considerably smaller ($\simeq$ 50 meV) than that of the E_0 transition. The agreement between the experimental and the theoretical values can be regarded as good because there are errors in estimating the width from the published spectra, and we have neglected the extrinsic broadening due to the apparatus used in the experiments.

In order to make a more accurate comparison with the experiments, the CPA values of $\Delta(E_0)$ and $\Delta(E_1)$ are calculated for the $x = 0.109$ alloy. The calculated half-widths of the E_0 and E_1 transitions are 13 and 2 meV respectively. From the spectrum, we estimate the corresponding values to be 8 to 15 meV and 3 to 6 meV. We see that CPA values are in excellent agreement with these experiments. Because x is small, the inclusion of off-diagonal disorder is not expected to change the calculated values significantly.

B. *Energy Gap*

The fundamental gaps of these alloys are calculated as a function of concentration. The VCA gap is an increasing function of x with a slope discontinuity at $x \simeq 0.11$. The conduction band minimum changes from L-point to X(Δ)-point at this crossover. In addition to the band gap, the effective electron masses and the band edge K_0 are also calculated. When the X(Δ) gap is preferred, the band edge moves linearly from Γ at $(0.9,0,0)_{x=0.15}$ to $(0.8,0,0)_{x=1}$. The effective masses at a given minimum increase linearly from their pure germanium values to the corresponding pure silicon values.

Using CPA self energies, the band gap, band masses, and the band edge are also calculated. The position of the band minimum did not change by virtue of the inclusion of the diagonal disorder. While the effective transverse mass remains almost the same as the VCA value, the longitudinal mass has a maximum of 12% enhancement. Because the real part of CPA self energies is negative in the forbidden gap region, an extra bowing is introduced to the VCA energy gap. Because of this bowing, the L-X(Δ) crossover takes place near $x \simeq 0.13$. The VCA, CPA, and experimental(36) bowing parameters are 0.06, 0.18 and 0.24 respectively. The calculated energy gap is plotted as a function of x in Fig. 6.

Because of the negligible change in the effective masses, the corresponding values in the pure materials are used in the calculation of the alloy-scattering-limited electron mobility. The CPA X-gap E_g^X and L-gap E_g^L are fitted to a polynomial form. The generalized Brooks' formula that is applicable to the alloys with an indirect gap and multiple bands is used.(15) The calculated electron drift mobility and the experimental Hall mobility(1) are plotted in Fig. 7, where the theory explains the qualitative behavior of experimental results.(1) As observed,(5) even a few percent alloy concentration can reduce the drift mobility substantially. It can be seen that the rate of decrease near

$x = 0$ and $x = 1$ are quite different. This is because the L-edge has more s content than the X-edge. Because the s scattering is dominant in these alloys, the L electrons are scattered more than the X electrons. Precisely for this reason, one observes the dip in the mobility near the L to X(Δ) crossover. For $x < 0.13$, the minimum gap is the L gap. After the crossover, the minimum gap is the X(Δ) gap, and the reduced alloy scattering increases the average mobility. For still larger x, the mobility decreases because of the increased alloy disorder. All these features are clearly seen in Fig. 7. While our calculations include the inter-valley scattering mediated by alloy disorder, the effect of other scattering mechanisms is expected to increase the dip near the crossover.

The calculated alloy scattering rate for the holes is several orders larger than that for the electrons, because (1) the valence band edge has dominant p content, (2) the p-scattering parameter ($\Delta\epsilon_p = 0.21$) is only 1/7 of $\Delta\epsilon_s$, which alone increases the scattering rate for holes by a factor of 50, and (3) finally, the imaginary part of the self-energy is proportional to the density of states, which approaches zero at the band edge. Hence, the hole mobility in this system is insensitive to alloy disorder.

In MCPA, the conduction band is pushed down, because of an increase in the imaginary part of the self energy, giving rise to an additional bowing in the fundamental gap. For an $x = 0.50$ alloy, the gap is reduced by 7 meV. The bowing parameter, including the MCPA correction, is 0.21, which is in excellent agreement with experiment.(37)

It is interesting to compare the results of our calculations with those of Hass et al.(34) In their calculations on the $Ga_{1-x}In_xAs$ alloy, CPA introduced an extra bowing in the fundamental gap. However, after the MCPA corrections, the total scattering was diminished and the results were similar to VCA results. These results were explained in terms of the relative strength and sign of the atomic term values and V_2^{AB}. We extend their argument to Si_xGe_{1-x} alloys. The hybridization level of silicon is higher than that of germanium. Because of its shorter bond length, the V_2 of silicon is larger than that of germanium. Thus, in this case, both effects combine to give more disorder in the conduction and valence bands. Therefore, the scattering is enhanced in these alloys. This explains the increase in the imaginary part of the self energy due to inclusion of OD disorder in our calculation.

In conclusion, we have incorporated both chemical and structural disorder into the calculation of the CPA band structure of Si_xGe_{1-x} alloys. The calculation, based on a realistic band structure of silicon and germanium, suggests that the band gap is an increasing function of x with a slope discontinuity at $x \simeq 0.13$. The linewidths of the E_0 and E_1 transitions calculated by CPA and MCPA are in good agreement with experiments. Addition of the structural disorder to the diagonal CPA decreases the band gap slightly but increases the s-part of the self energy considerably over certain energy ranges. The calculated alloy-scattering-limited electron-drift mobility is in qualitative agreement with the observed Hall mobilities.

ACKNOWLEDGMENTS

This work was supported in part by Grant AFOSR-84-0282 and DARPA/AFOSR PR: FQ8671-851100. A.-B. Chen would like to thank Professor W.E. Spicer for his hospitality at Stanford University.

REFERENCES

[1] M. Glicksman, Phys. Rev. *111*, 125 (1958); Phys. Rev. *100*, 1146 (1955).

[2] B.A. Bunker, S.L. Hulbert, J.P. Stotl, and F.C. Brown, Phys. Rev. Lett. *53*, 2157 (1984).

[3] H.J. Lee, L.Y. Juravel, and J.C. Woolley, Phys. Rev. *B21*, 659 (1980).

[4] N. Lifshitz, A. Jayaraman, and R.A. Logan, Phys. Rev. *B21*, 670 (1980).

[5] H.M. Manasevit, I.S. Gergis, and A.B. Jones, Appl. Phys. Lett. *41*, 464 (1982); J. Elec. Materials *12*, 637 (1983).

[6] S. Krishnamurthy and J.A. Moriarty, Superlattices and Microstructures *1*, 209 (1985); Phys. Rev. *B* in press (1985).

[7] J.C. Bean, L.C. Feldman, A.T. Flory, S. Nakahara, and I.K. Robinson, J. Vac. Sci. Technol. *A2*, 436 (1984).

[8] R. People, J.C. Bean, D.V. Lang, A.M. Sergent, H.L. Stormer, K.W. Wecht, R.T. Lynch, and K. Baldwin, Appl. Phys. Lett. *45*, 1231 (1984).

[9] F. Cerdeira, A. Pinczuk, and J.C. Bean, Phys. Rev. *B31*, 1202 (1985).

[10] T.P. Pearsall, F.H. Pollak, and J.C. Bean, Bull. Am. Phys. Soc. *30*, 266 (1985).

[11] K.E. Newman and J.D. Dow, Phys. Rev. *B30*, 1929 (1984).

[12] M.Z. Huang and W.Y. Ching, Superlattices and Microstructures *1*, 137 (1985).

[13]D. Stroud and H. Ehrenreich, Phys. Rev. *B2*, 3197 (1970).

[14]S. Krishnamurthy, A. Sher, and A.-B. Chen, Phys. Rev. Lett., in press (July 1985).

[15]S. Krishnamurthy, A. Sher, and A.-B. Chen, Appl. Phys. Lett., in press (1985).

[16]A.-B. Chen and A. Sher, Phys. Rev. *B22*, 3886 (1980).

[17]E.O. Kane, Phys. Rev. *B13*, 3478 (1976).

[18]D.J. Chadi, Phys. Rev. *B16*, 3572 (1977).

[19]A.-B. Chen and A. Sher, Phys. Rev. *B23* 5360 (1981).

[20]A.-B. Chen and A. Sher, Phys. Rev. *B26*, 6603 (1982).

[21]R.R.L. Zucca, J.P. Walter, Y.R. Shen, and M.L. Cohen, Solid State Commun. *8*, 627 (1970).

[22]R.R.L. Zucca and Y.R. Shen, Phys. Rev. *B1*, 2668 (1970).

[23]M. Welkowsky and R. Braunstein, Phys. Rev. *B5*, 497 (1972).

[24]R.A. Pollak, L. Ley, S. Kowalcyzk, D.A. Shirley, J. Joannopolos, D.J. Chadi, and M.L. Cohen, Phys. Rev. Lett. *29*, 1103 (1973).

[25]L. Ley, R.A. Pollak, F.R. McFeely, S.P. Kowalcyzk, and D.A. Shirley, Phys. Rev. *B9*, 600 (1974).

[26]W.D. Grobman and D.E. Eastman, Phys. Rev. Lett. *29*, 1508 (1972).

[27]D.E. Eastman, W.D. Grobman, J.L. Freeouf, and M. Erbudak, Phys. Rev. *B9*, 3473 (1974).

[28] W.E. Spicer and R.C. Eden, *Proceedings of the Ninth International Conf. Phys. Semiconductors*, 1968 (Nauka, Leningrad, 1969), Vol. 1, p. 61.

[29] G. W. Gobeli and F.G. Allen, Phys. Rev. *A245*, 137 (1965).

[30] J.R. Cheliskowsky and M.L. Cohen, Phys. Rev. *B14*, 556 (1976); and the references cited therein.

[31] J.A. Moriarty and S. Krishnamurthy, J. Appl. Phys. *54*, 1892 (1983).

[32] K.C. Hass, B. Velicky, and H. Ehrenreich, Phys. Rev. *B29*, 3697 (1984).

[33] F. Ducastelle, J. Phys. *C7*, 1795 (1974).

[34] K.C. Hass, R.J. Lampert, and H. Ehrenreich, Phys. Rev. Lett. *52*, 77 (1984).

[35] W.A. Harrison, *Electronic Structure and the Properties of Solids*, (W.H. Freeman and Co., San Francisco, 1980) p. 65.

[36] J.S. Kline, F.H. Pollak, and M. Cardona, Helv. Phys. Acta. *41*, 968 (1968).

[37] R. Braunstein, A.R. Moore, and F. Herman, Phys. Rev. *109*, 695 (1958).

TABLE I. Pseudopotential form factors

and the band parameters (in eV).

Parameter	Silicon	Germanium
$V(\sqrt{3})$	− 2.872	− 2.872
$V(\sqrt{4})$	0.124	0.124
$V(\sqrt{8})$	0.638	0.638
$V(\sqrt{11})$	0.109	0.109
Δ_s	−16.175	−16.922
Δ_p	−16.109	−14.971
V_{ss}	− 0.111	0.131
V_{sp}	0.040	0.150
V_{xx}	0.025	0.030
V_{xy}	0.050	0.100

TABLE II. Band structure of silicon and germanium

(all energies are expressed in eV).

	Silicon		Germanium	
Bands	Calculated	EXPTL^a/EPM^b	Calculated	EXPTL^a/EPM^b
Γ_1	−12.60	−12.4 ± 0.6	−12.56	12.6 ± 0.3
L_{2v}	−10.26	− 9.3 ± 0.4	−10.74	10.6 ± 0.5
L_{1v}	− 6.99	6.8 ± 0.2	− 7.65	7.4 ± 0.3
X_{1v}	− 8.29		− 9.20	
X_{4v}	− 2.55		− 2.55	
$L_{3'v}$	− 1.11	1.2 ± 0.2	− 1.13	1.1 ± 0.2
Γ_{25v}	0.0	0.0	0.0	0.0
L_{1c}	2.24		0.76	0.76
$\Gamma_{2'c}$	4.10	4.00 ± 0.05	0.99	0.99
Γ_{15c}	3.43	3.40	3.24	
X_{1c}	1.34	1.17	0.95	
$L_{3'c}$	4.34		4.16	
E_g	1.11	1.11	0.76	0.76
K_0	(0.8,0,0)	(0.8,0,0)	(0.5,0.5,0.5)	(0.5,0.5,0.5)
m^*_{el}	0.89	0.91	1.09	1.59
m^*_{et}	0.16	0.19	0.077	0.082
m^*_v	0.35	0.50	0.28	0.34

(a) References 21–28

(b) Reference 30

TABLE III. Calculated values of E_0, E_1 and their respective half-widths $\Delta(E_0)$ and $\Delta(E_1)$ (all energies are in eV).

X	Quantity	VCA	CPA	MCPA
0.10	E_0	1.290	1.248	--
	$\Delta(E_0)$		0.011	--
	E_1	2.016	1.995	--
	$\Delta(E_0)$		0.001	--
0.109	E_0	2.028	2.009	--
	$\Delta(E_0)$		0.013	--
	E_1	1.318	1.275	--
	$\Delta(E1)$		0.002	--
0.15	E_0	1.442	1.382	--
	$\Delta(E_0)$		0.032	--
	E_1	2.083	2.051	--
	$\Delta(E_1)$		0.002	--
0.50	E_0	2.517	2.418	2.391
	$\Delta(E_0)$		0.186	0.206
	E_1	2.578	2.510	2.498
	$\Delta(E_1)$		0.0308	0.0319

FIGURE CAPTIONS

FIG. 1. Calculated VCA band structures of (a) silicon, (b) germanium, (c) $Si_{0.5}Ge_{0.5}$ alloys, and (d) $Si_{0.10}Ge_{0.90}$. (e) Calculated VCA values(solid) and the experimental values (dashed) of the E_0 and E_0' peak positions are plotted as a function of alloy concentration x.

FIG. 2. The variation of the imaginary part of the self energy Σ_s (CPA) as a function of energy for x = 0.50.

FIG. 3. Calculated CPA complex band structure of the $Si_{0.5}Ge_{0.5}$ alloy. Only the bands in the vicinity of the energy gap are shown. The shaded portion represents the alloy broadening.

FIG. 4. Imaginary part of Σ_s (CPA) and Σ_s (MCPA) as a function of energy for x = 0.50 alloy.

FIG. 5. The variation in the width of the lowest lying conduction band as a function of K_x in the [100] direction for the x = 0.50 alloy.

FIG. 6. The variation of the VCA energy gap (dash-dotted) and the CPA energy gap (solid) as a function of x.

FIG. 7. Calculated drift mobility (solid) and the experimental (dashed) Hall mobility (Reference 1) as a function of x.

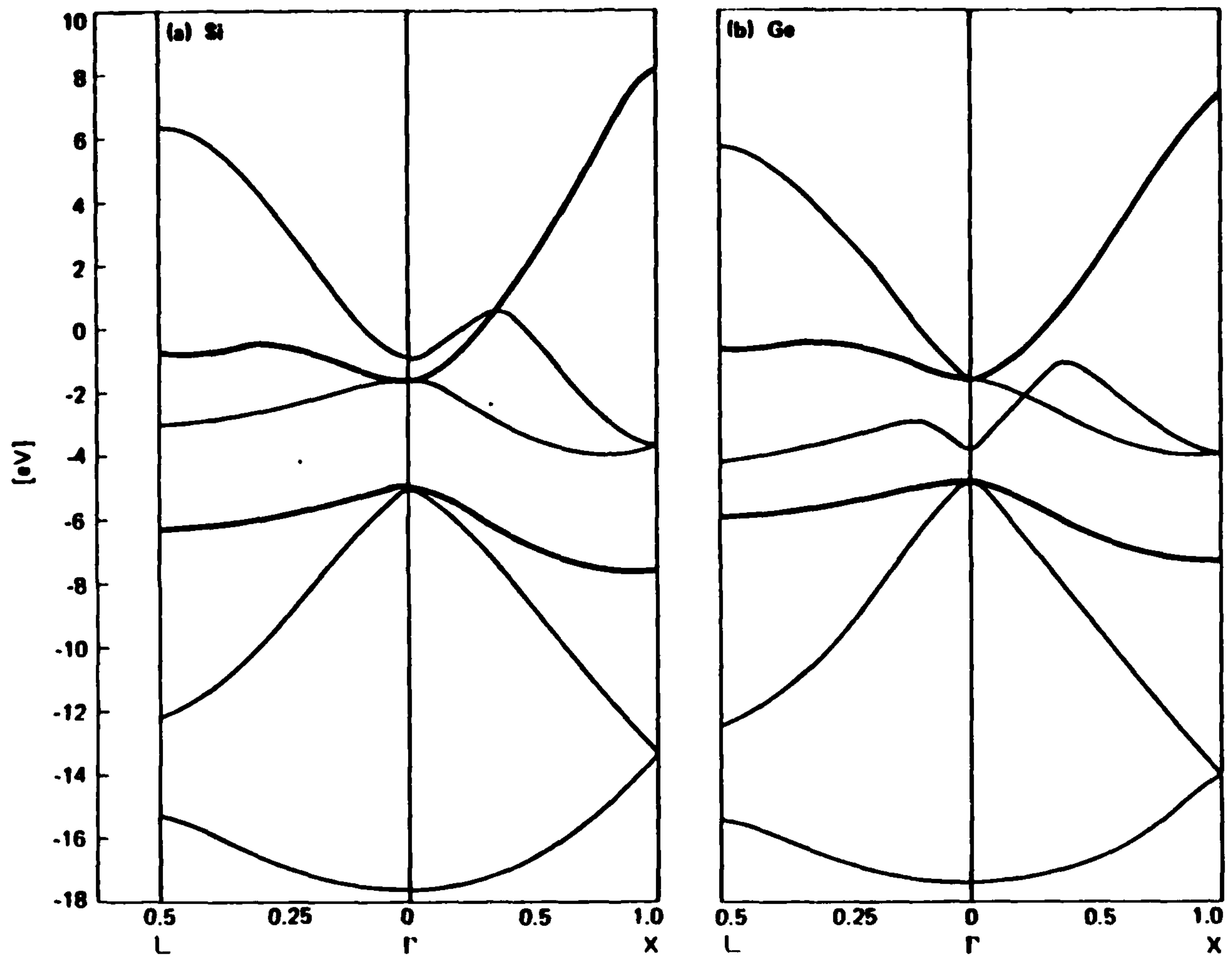


FIGURE 1

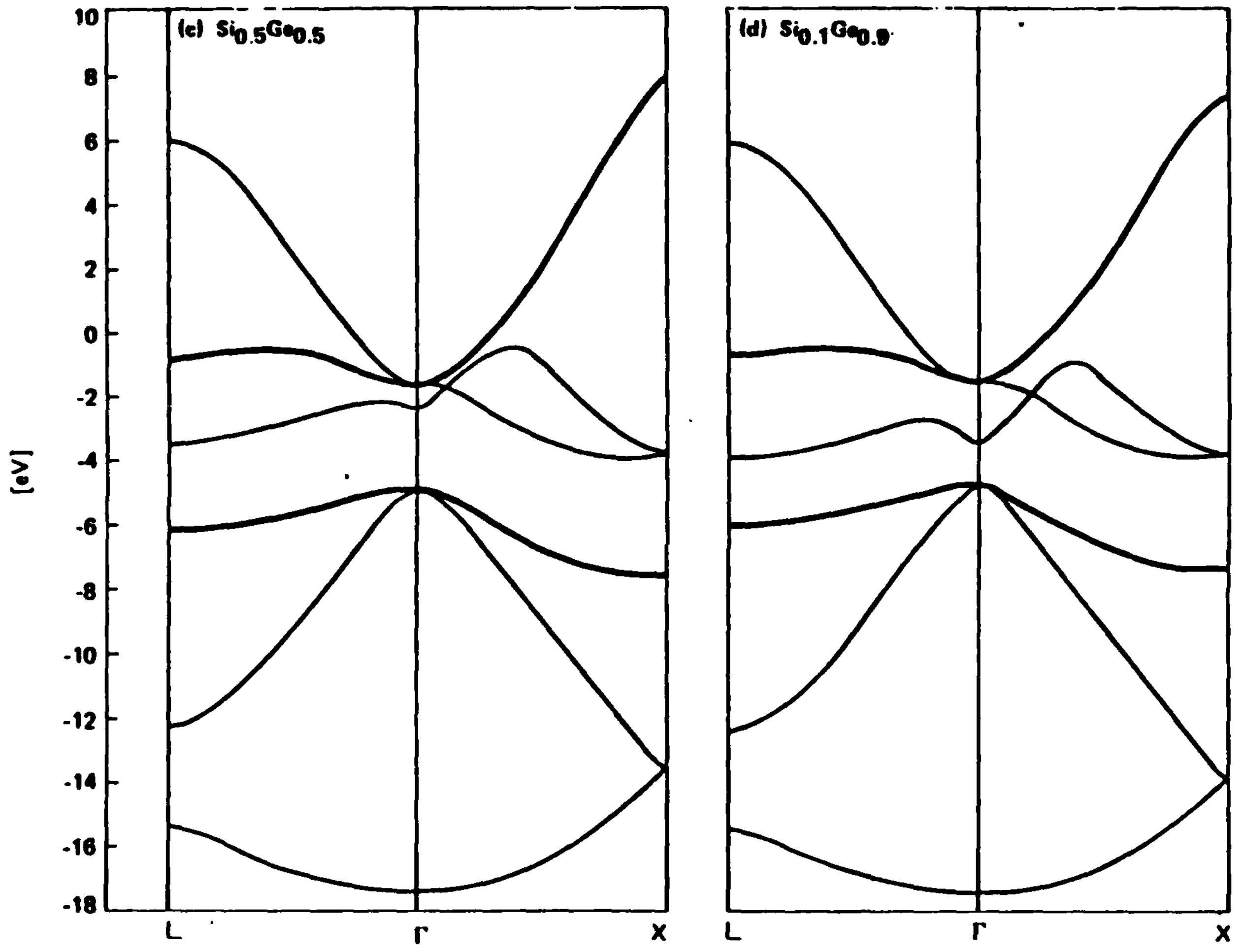


FIGURE 1 (cont'd)

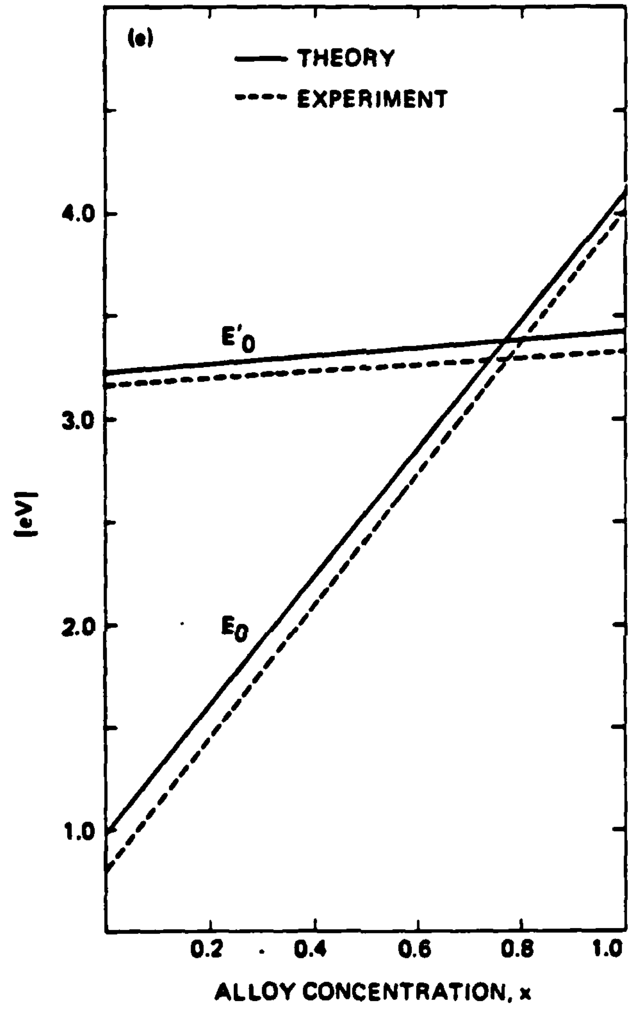


FIGURE 1 (concluded)

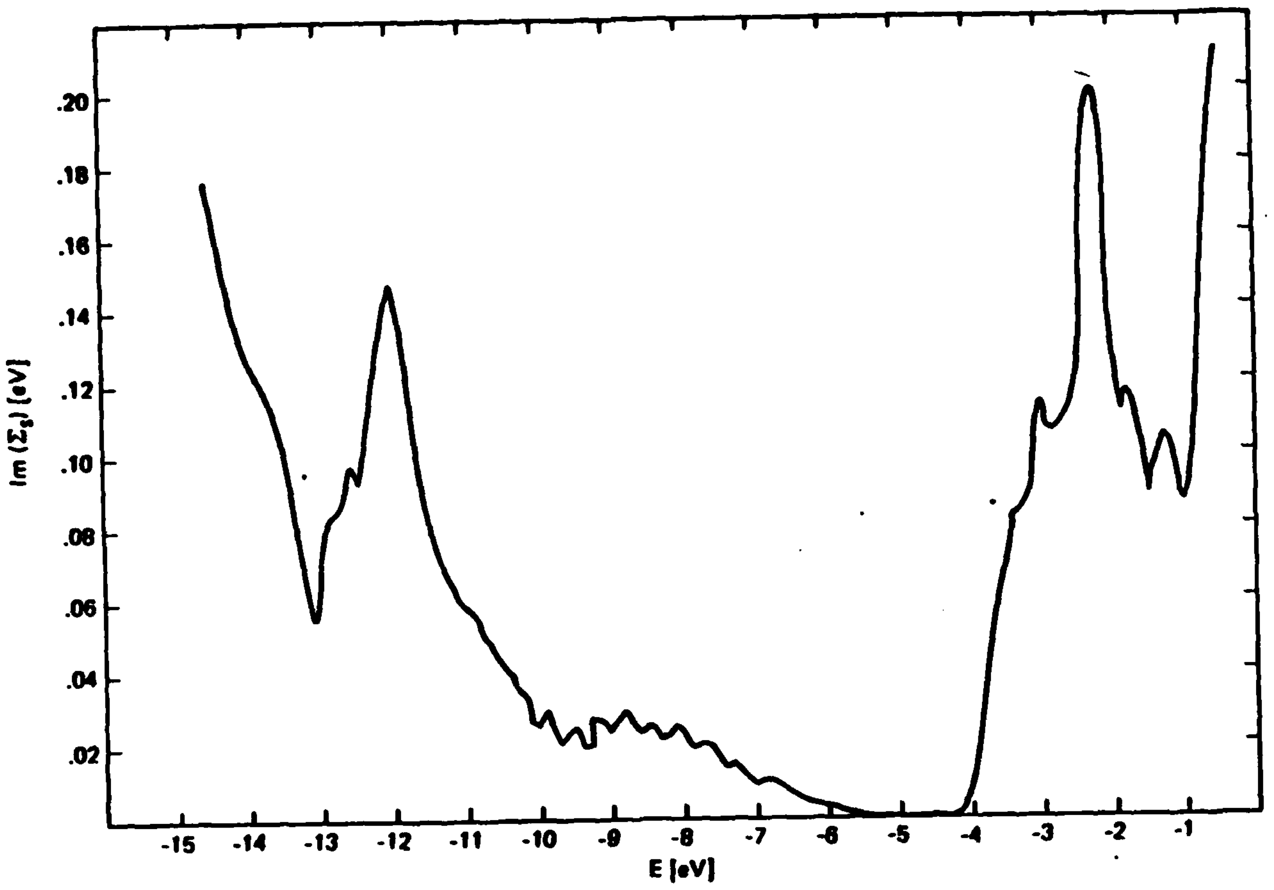

Im (Σ_s) [eV]
.20
.18
.16
.14
.12
.10
.08
.06
.04
.02
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
E [eV]

FIGURE 2

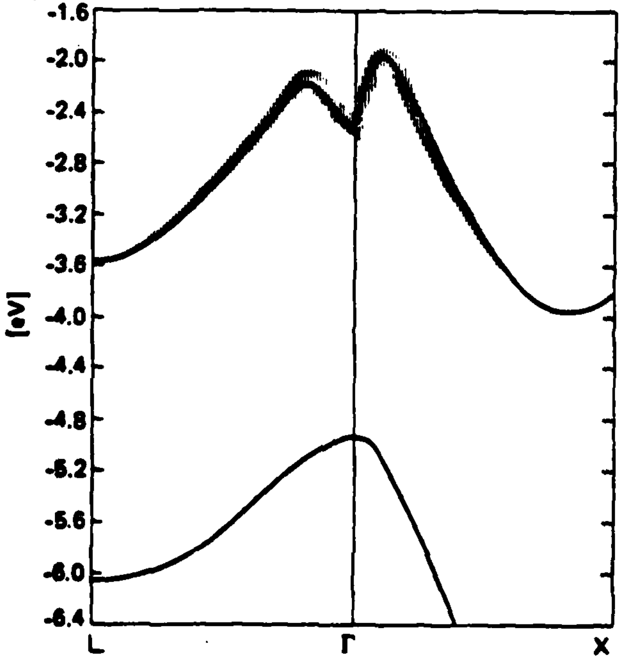
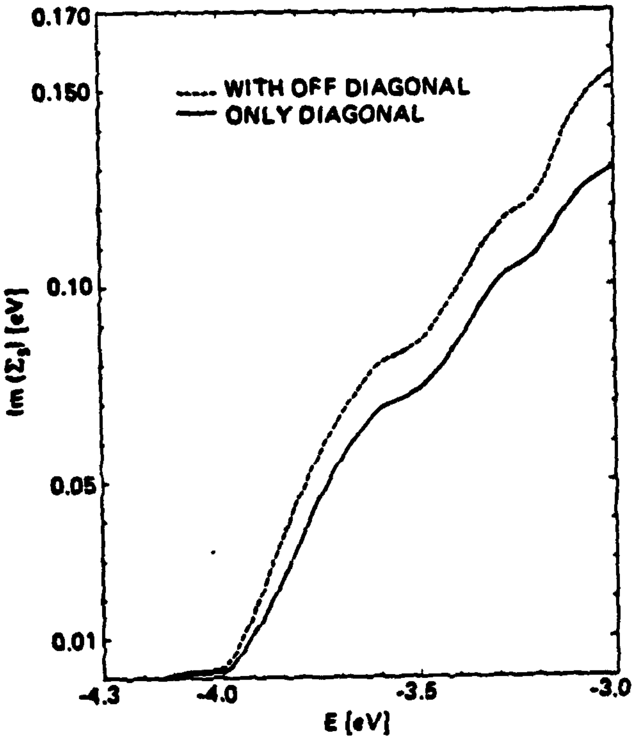


FIGURE 3



0.170
0.150
0.10
0.05
0.01
Im (Σ_s) [eV]
-4.3
-4.0
-3.5
-3.0
E [eV]
----- WITH OFF DIAGONAL
—— ONLY DIAGONAL

FIGURE 4

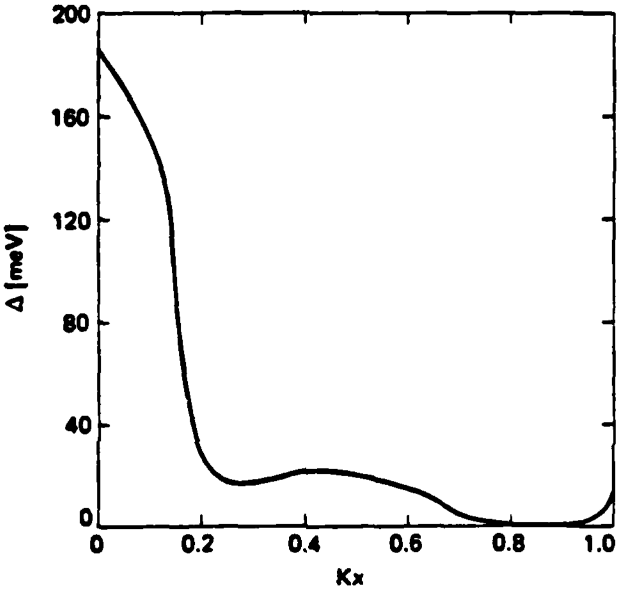


FIGURE 5