(12)

SRI International

Final Report • October 1993

THEORY OF ENERGY-ASSISTED EPITAXY; THEORY OF ORDERED SEMICONDUCTOR ALLOYS; $In_{1-x}Tl_xSb$ AS A LONG-WAVE INFRARED MATERIAL

M.A. Berding, Senior Research Physicist
A. Sher, Associate Laboratory Director
Physical Electronics Laboratory

SRI Project 6147

Prepared for:

Procuring Contracting Officer
Office of Naval Research
Department of the Navy
800 N. Quincy Street
Code 1512A: DLM
Arlington, VA 22217

Attn: Dr. George Wright
Scientific Officer

Contract N00014-88-C0096

333 Ravenswood Avenue • Menlo Park, CA 94025-3493 • (415) 326-6200 • FAX: (415) 326-5512 • Telex: 334486

Final Report • October 1993

THEORY OF ENERGY-ASSISTED EPITAXY; THEORY OF ORDERED SEMICONDUCTOR ALLOYS; $In_{1-x}Tl_xSb$ AS A LONG-WAVE INFRARED MATERIAL

M.A. Berding, Senior Research Physicist
A. Sher, Associate Laboratory Director
Physical Electronics Laboratory

SRI Project 6147

Prepared for:

Procuring Contracting Officer
Office of Naval Research
Department of the Navy
800 N. Quincy Street
Code 1512A: DLM
Arlington, VA 22217

Attn: Dr. George Wright
Scientific Officer

Contract N00014-88-C0096

Approved:
Eric Pearson, Director
Physical Electronics Laboratory

Donald L. Nielson, Vice President
Computing and Engineering Sciences Division

SRI International 333 Ravenswood Avenue • Menlo Park, CA 94025-3493 • (415) 326-6200 • FAX (415) 326-5512 • Telex 334486

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE — *Form Approved OMB No. 0704-0188*

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): SRI Project 6147	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: SRI International; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code): 333 Ravenswood Avenue, Menlo Park, CA 94025-3493	7b. ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Contract N00014-88-C0096
8c. ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Code 1512A; DLM, Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification): Theory of Energy-Assisted Epitaxy; Theory of Ordered Semiconductor Alloys; $In_{1-x}Tl_xSb$ as a Long-Wave Infrared Material

12. PERSONAL AUTHOR(S): M.A. Berding, A. Sher

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Report	From 4-27-88 To 7-1-93	October 1993	252

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Epitaxial growth, MBE, ALE, semiconductor surface, LWIR materials, InTlSb
20	02		
20	12		

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

In this report we summarize research by SRI International on the theory of energy-assisted epitaxy and the properties of ordered semiconductor compounds. We include our published papers, and summarize the major conclusions from our work. We also report on our first year of work on the properties of InTlSb, a proposed new long-wave infrared detector material. We include a discussion of major aspects of our work, including the advantages of this material over HgCdTe and III-V superlattice and quantum-well devices.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: [X] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. [ ] DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Ms. Gail Boeger	22b. TELEPHONE (Include Area Code): (703) 696-2594; 22c. OFFICE SYMBOL: Code 1513:GDB

DD Form 1473, JUN 86 *Previous editions are obsolete* SECURITY CLASSIFICATION OF THIS PAGE

(SRI on-line version)

CONTENTS

1. INTRODUCTION

This final report covers a five-year program conducted by SRI International (SRI) for the Office of Naval Research. The focus of this program has changed over its five-year duration. The original focus was on the theory of energy-assisted epitaxy, a continuation of the previous contract. Our focus then shifted to the properties of ordered semiconductors grown by epitaxial methods. Finally, in the last year of the program, our focus shifted toward the development of a new material we had predicted, $In_{1-x}Tl_xSb$.

Because of the volume of material covered in this program, we have organized this final report as a brief summary of the work, with the contract-supported published papers presented as an appendix. Because much of the work on the last topic studied, InTlSb, has not been published, we include a section on this topic.

2. THEORY OF ENERGY-ASSISTED EPITAXY AND THE THEORY OF ORDERED SEMICONDUCTOR ALLOYS

The primary goal of this portion of the program was to understand energy-assisted epitaxial growth of semiconductors and to predict and study the properties of ordered semiconductors that might result. Subtasks include the study of surface energetics including the effects of surface charge transfer; the development of an epitaxial growth model including the effects of energy assist; the prediction of the ordered alloys that may result from the nonequilibrium growth; and the study of the properties, including stability, of ordered materials grown by nonequilibrium methods. A very brief summary of this work and major conclusions drawn from it are given here.

The development of a growth model requires a reliable estimate of the interaction energy between constituent atoms with various local environments. Once these energies are known, epitaxial growth can be modeled using appropriate statistics. We have used the calculated pair interaction energies in a quasichemical approximation to study the qualitative nature of semiconductor surfaces. The effect of energy assist has been modeled through the consequent change in surface entropy.

In this program we have shown that charge transfer between surface cations and anions substantially affects the surface atom interaction energies and consequently the nature of the epitaxially grown surfaces. Relaxation of surface atoms is found to reduce the surface energies by of the order of an electron volt. We have also developed models of the surfaces during atomic layer epitaxy (ALE) and molecular beam epitaxy (MBE) growth, using the calculated pair energies in a quasi-chemical approximation.

An observed difference in the growth of HgCdTe on the (111) A and B surfaces, which cannot be explained using simple bond-breaking energy arguments, has been explained using a growth model based on our surface binding energies. A varient of the generalized perturbation method (GPM) was developed to obtain cluster energies in semiconductor alloys.

Under this program we have also developed a model of segregation at the free surfaces and interfaces of semiconductor alloys. This model has been applied to a number of systems, including HgCdTe and HgZnTe. We have also developed phase diagram models using our quasichemical formalism and applied it to several systems.

Several properties, such as the elastic constants and cohesive energies of the semiconductors, are necessary input for the growth models we have developed. Thus some effort under this program was expended on refining some tight-binding models of these properties to provide more accurate and predictive models on which the growth models were based. We have also used these models to study native point defects in the semiconductor alloys.

A linear combination of atomic orbitals (LCAO) Hamiltonian was developed to apply to the growth problem in the semiconctor alloys. While no papers were published on this work because our funding was depleted before the work could be completed, a preliminary paper* (see Appendix) on this work was written and applied to the ordering in GaAlAs, which we find to be driven by surface interactions.

The important conclusions we have drawn from our work are:

- Surface sublimation and binding energies needed in growth models are surface-orientation dependent
- Surface sublimation energies are not simply proportional to the number of first neighbor bonds made to the surface
- Charge transfer among surface atoms is a large component of surface sublimation energies and will be most important in wide bandgap materials
- Charge transfer among surface atoms can lead to ordering of the semiconductor surface below some critical temperature, T_c
- A simple growth model incorporating our surface-dependent binding and sublimation energies can explain the difference in the growth rate of HgCdTe on the (111)A and (111)B surfaces
- A multilayer growth model was developed and applied to HgTe, CdTe, Si, and GaAs to determine regimes in which layer-by-layer and three-dimensional growth occurs
- Ordering in semiconductor alloys at the growth temperature is found to be metastable in all systems studied
- Surface segregation is surface-orientation-dependent

---

* S. Krishnamurthy and A. Sher, "Ordering in GaAlAs."

- Long-range alloy compositional oscillation away from a surface can occur metastably below some critial temperature, and its formation is driven by the same mechanisms which drives surface segregation
- Vacancy formation energies in pseudobinary semiconductor alloys are very sensitive to the alloy environment
- Ordering in GaAlAs is driven by surface Coulomb interactions.

3. $In_{1-x}Tl_xSb$ AS A LONG-WAVE INFRARED MATERIAL

We demonstrated, based on full-potential linearized muffin tin orbital (FP-LMTO) based self-consistent solutions to the Schrodinger equation with the exchange term expressed in the local density approximation (LDA), that TlSb in the zincblende structure has roughly the same bond length (~ 2% smaller) as that of InSb, that it is a semimetal, and that it has a large cohesive energy comparable to that of InSb. We also demonstrated that TlSb minimum energy is attained in the CsCl structure. However, for $x < 0.15$, the zincblende structured alloy was predicted to be stable for attainable growth temperatures. We also predicted that at about $x \cong 0.09$ the band gap of the alloy would be ~ 0.1 eV, the value needed for long-wave infrared (LWIR) focal plane arrays (FPA). Because for the same band gap, 0.1 eV, the average bond energy of $In_{0.91}Tl_{0.9}Sb$ is ~ 1.43 eV, compared to $Hg_{0.78}Cd_{0.22}Te$ at ~ 0.88 eV, we suggested the InTlSb alloy may be a more robust material for FPAs.

Several experimental groups are trying to grow the material, and one at Northwestern claims to have done so.* There are aspects of the Northwestern work that are suspect, and while we would like to believe they have prepared the material, we recommend judgment be withheld until further tests are performed on their material.

4. COMMENTS ABOUT FPA SYSTEM ISSUES: WHAT ARE THE SYSTEM MERITS OF $In_{1-x}Tl_xSb$

There has been a heated debate about the relative importance of individual pixel detectivity (high quantum efficiency and low noise) and the spatial uniformity of this detectivity over an array. Some types of nonuniformity are correctable in the output circuitry; others are not. A theory says that if you are observing a high-background-temperature scene, then quantum efficiency is less important to the noise equivalent contrast temperature (NEΔT) (related to the minimum resolvable temperature in a scene) than the spatial nonuniformity. However, for low-

* Y.H. Choi, C. Besikci, R. Sudharsanan, and M. Razeghi, "Growth of $In_{1-x}Tl_xSb$, a new infrared material, by low-pressure metalorganic chemical vapor deposition," *Appl. Phys. Lett.* 63, 361 (1993).

background-temperature scenes, the reverse is true. The arguments center around the critical scene temperature T_{cs} that separates these regions and the kinds of nonuniformities it is practical to correct. Some people contend that T_{cs} lies below 0°C, so only Arctic and space observations require high quantum efficiencies. Others place T_{cs} well above all typical earth backgrounds, so a high quantum efficiency is always beneficial. From what we have seen from cameras made from different materials, we believe the high-quantum-efficiency supporters are right; besides that, it is logical.

One important source of incorrectable nonuniformity is a quantum efficiency that has a sharp spectral dependence beyond its cut-on. Thus, a flat quantum efficiency is a distinct advantage.

Being able to build the signal processing circuit on the same chip as the detector array is also highly desirable. This is not possible at present for HgCdTe, so the FPA is In bump bonded to a Si chip that contains the readout circuits. The thermal expansion mismatch between these dissimilar materials limits the size of practical FPAs made this way. If larger arrays are needed, they must be assembled in a mosaic pattern, which limits the picture quality and adds complexity to the system. This is why there is so much effort these days devoted to growing HgCdTe on Si substrates. The problem there is that the lattice constant difference is ~19%, so dislocations are plentiful, which again degrades device performance.

There are many other system issues, but this minimal set suffices to set the stage for the major arguments surrounding the relative merits of LWIR alloy detectors like HgCdTe and InTlSb, quantum well detectors, and strained layer superlattice detectors.

HgCdTe 256 × 256 pixel arrays have been built with $NE\Delta T \cong 6$ mK. This is more than adequate for many applications, but it is hard to do with high yield, so these arrays are expensive. Some strategic applications require much larger arrays. The quantum efficiency of HgCdTe is ~70%, and spectrally flat, so corrections to spatial nonuniformity are reliable over a large dynamic range. CdZnTe currently is the best substrate material. It is good, but not superb, since the sizes that can be prepared are limited, it has moderate dislocation densities, and some troublesome impurities tend to migrate into the active material. Some of these properties are being fixed.

The quantum well and strained layer superlattice detectors are basically built on GaAs substrates. Thus, the signal processing can be done on the chip that holds the array. However, both materials have low quantum efficiencies (< 10%), and they have sharp spectral variations. The supporters of these devices contend that because they are constructed from III-V rather than II-VI material, their relative robustness will lead to good spatial uniformity (limiting the need for corrections), few dislocations, and generally high-quality material. Reasonable performance has been demonstrated on 64 × 64 arrays, which is quite good for a newly evolved technology. However, they were not as good as current similar HgCdTe arrays. Thus, quantum well and superlattice supporters invoke expectations based on intuition and how well devices have worked in view of the immaturity of the technology. They are saying, "bet on what you don't know about, rather than spend your effort to fix the troubles you have encountered in HgCdTe." They have a firm, legitimate claim to better substrates compatible with readout circuits.

MWIR InSb detectors arrays are slightly better than MWIR $Hg_{0.7}Cd_{0.3}Te$ detectors, but the two are close competitors. LWIR $Hg_{0.7}Cd_{0.3}Te$ is a much more difficult material, for a wide variety of reasons, but much of the trouble can be traced back to the fact that the material is 70% HgTe (bond strength 0.81 eV/bond) and 30% CdTe (1.1 eV/bond).

$In_{0.91}Tl_{0.9}Sb$ should have optical and electronic properties similar to those of $Hg_{0.78}Cd_{0.22}Te$, but it is 91% InSb (1.43 eV/bond) and only 9% TlSb (1.34 eV/bond). The constituents' bond length difference in InTlSb is not as small as that for HgCdTe, but it is only ~2%—still a near-lattice-matched alloy. The most important feature is that the average bond strength is much larger, which will cause native-point defect populations and other imperfection densities to be much smaller and less troublesome.

The hope for a fast utilization of $In_{0.91}Tl_{0.09}Sb$, once it is grown, is that it will be essentially InSb-like in its defect properties. If this is the case, then it will be possible with only slight modifications to substitute InTlSb epitaxial layers on InSb substrates into an InSb FPA production line, use essentially the same readout circuitry, and end up with an LWIR response rather than MWIR. If the LWIR characteristics are as good as we project, then a superior (cheaper) system is the result. We will have to wait to see.

Appendix

PAPERS WRITTEN WITH FULL OR PARTIAL SUPPORT OF ONR UNDER CONTRACT N00014-88-C0096

The following papers, reproduced here in their entirety, document the work accomplished during the five-year program.

R.S. Patrick, A.-B. Chen, A. Sher, and M.A. Berding, "Phase diagrams and microscopic structures of (Hg,Cd)Te, (Hg,Zn)Te, and (Cd,Zn)Te Alloys," *J. Vac. Sci. Technol. A* **6**, 2643 (1988).

A.Sher, M.A.Berding, A.-B.Chen, and R.Patrick, "Correlations and alloy properties: growth, vacancies, surface segregation," *J. Cryst. Growth* **98**, 27 (1989).

J.T. Schick, S. Boss, and A.-B. Chen, "Electronic structure of HgTe and CdTe surfaces and HgTe/CdTe interfaces," *Phys. Rev. B* **40**, 7825 (1989).

R.S. Patrick, A.-B. Chen, A. Sher, and M.A. Berding, "Surface segregation in pseudobinary alloys," *Phys. Rev. B* **39**, 5980 (1989).

A. Sher, M.A. Berding, R.S. Patrich, and A.-B. Chen, "On passivation of $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$," *Proceedings of the 1990 Meeting of the IRIS Specialty Group on Infrared Materials*, p. 323 (1990).

A.-B. Chen, Y.-M. Lai-Hsu, S. Krishnamurthy, M.A. Berding, and A.Sher, "Band structures of HgCdTe and HgZnTe alloys and superlattices," *Semicond. Sci. Technol.* **5**, S100 (1990).

C.Y. Yeh, A.-B. Chen, D.M. Nicholsen, and W.H. Butler, "Full-potential Korringa-Kohn-Rostoker band theory applied to the Mathieu potential," *Phys. Rev. B* **42**, 10976 (1990).

M.A. Berding, A. Sher, A.-B. Chen, and R. Patrick, "Vacancies and surface segregation in HgCdTe and HgZnTe," *Semicond. Sci. Technol.* **5**, S86 (1990).

M.A. Berding, A. Sher, and A.-B. Chen, "Vacancy formation and extraction energies in semiconductor compounds and alloys," *J. Appl. Phys.* **68**, 5064 (1990).

M.A. Berding, M. van Schilfgaarde, A.T. Paxton, and A. Sher, "Defects in ZnTe, CdTe and HgTe: Total energy calculations," *J. Vac. Sci. Technol. A* **8**, 1103 (1990).

M.A. Berding, S. Krishnamurthy, and A. Sher, "Energetics for vapor phase growth models of HgCdTe," *Proceedings of the 1990 Meeting of the IRIS Specialty Group on Infrared Materials*, p. 257 (1990).

S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, "Epitaxially grown semiconductor surfaces," *J. Cryst. Growth* **109**, 88 (1990).

S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, "Energetics of molecular-beam epitaxy models," *J. Appl. Phys.* **68**, 4020 (1990).

S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, "Semiconductor surface sublimation energies and atom-atom interactions," *Phys. Rev. Lett.* **64**, 2531 (1990).

S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, "Surface roughness theory and low-temperature epitaxy," *Mat. Res. Soc. Symp. Proc.* **Vol. 161**, 291 (1990).

A. Sher, M. van Schilfgaarde, and M.A. Berding, "Review of the status of computational solid-state physics," *J. Vac. Sci. Technol. B* **9**, 1738 (1991).

A. Sher, M.A. Berding, M. van Schilfgaarde, and A.-B. Chen, "HgCdTe status review with emphasis on correlations, native defects and diffusion," *Semicond. Sci. Technol.* **6**, C59 (1991).

C.Y. Yeh, A.-B. Chen, and A. Sher, "Formation energies, bond lengths and bulk moduli of ordered semiconductor alloys from tight-binding calculations," *Phys. Rev. B* **43**, 9138 (1991).

M.A. Berding, A. Sher, and A.-B. Chen, "$Hg_{1-x}Cd_xTe$: Defect structure overview," *Mat. Res. Soc. Symp. Proc.* **Vol. 216**, 3 (1991).

M.A. Berding, S. Krishnamurthy, and A. Sher, "Surface energies for molecular beam epitaxy growth of HgTe and CdTe," *J. Vac. Sci. Technol. B* **9**, 1858 (1991).

S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, "Surface energies and order-state: Effects on semiconductor growth," in *Computer Aided Innovation of New Materials*, ed. M. Doyama, T. Suzuki, J. Kihara, and R. Yamamota (Elsevier Science Pub., North Holland, 1991).

A. Sher, M. van Schilfgaarde, M.A. Berding, "InTlSb as an LWIR Material," *Proceedings of the 1992 Meeting of the IRIS Specialty Group on Infrared Materials*.

A.-B. Chen, A. Sher, and T. Yost, "Elastic constants of semiconductors and alloys," in *Semiconductors and Semimetals*, Academic Press Inc., p. 1 (1992).

S. Krishnamurthy and A. Sher, "Ordering in GaAlAs," unpublished work.

A.-B. Chen, M. van Schilfgaarde, and A. Sher, "Comparison of $In_{1-x}Tl_xSb$ and $Hg_{1-x}Cd_xTe$ as LWIR materials," *J. Electr. Mater.* **22**, 8, 843–846 (1993).

M. van Schilfgaarde, A. Sher, and A.-B Chen, "InTlSb: An infrared detector material?," *Appl. Phys. Lett.* **62**, 1857 (1993).

Phase diagrams and microscopic structures of (Hg,Cd)Te, (Hg,Zn)Te, and (Cd,Zn)Te alloys

R. S. Patrick and A.-B. Chen
Auburn University, Auburn, Alabama 36849

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

(Received 3 November 1987; accepted 1 March 1988)

A cluster theory based on the quasichemical approximation has been applied to study the local correlation, bond-length distribution, and phase diagrams of the II–VI pseudobinary alloys $Hg_{1-x}Cd_xTe$, $Hg_{1-x}Zn_xTe$, and $Cd_{1-x}Zn_xTe$. The cluster energy is calculated by letting it relax in some effective alloy medium and then considering the contributions from the strain and chemical energies. Two different models are presented to simulate the alloy medium. While both models show that all three alloys have nearly random distributions, the signs of the local correlation prove to be sensitive to the alloy medium chosen for the energy calculation. Good agreement is found between experiment and the bond lengths and phase diagrams in both models.

I. INTRODUCTION

This paper is a summary of our theoretical investigations[1–4] into the statistical properties of $Hg_{1-x}Cd_xTe$, $Hg_{1-x}Zn_xTe$, and $Cd_{1-x}Zn_xTe$ alloys. A primary objective of these studies is to correlate microscopic quantities such as bond lengths and atomic distributions to macroscopic quantities such as internal strain energy, phase diagrams, and mechanical properties of the alloy (e.g., bulk modulus).

The theory is based on a generalized quasichemical approximation (GQCA),[1] where the usual pair interaction is replaced by a microcluster unit of arbitrary size. This entails the calculation of cluster energies in some effective medium along with a statistical formulation. Once the cluster energies are available, the cluster populations can be computed, which in turn yields the mixing free energy and other quantities mentioned above.

The next section will present a summary of the statistical theory. (Full details are provided in Ref. 1.) Section III describes the calculation of cluster energies along with a discussion of effective alloy media in which the clusters are placed. Section IV will present results of some quantities calculated with the theory. Section V contains conclusions and discussions.

II. GQCA STATISTICAL THEORY

For a zinc-blende alloy of the form $A_{1-x}B_xC$ at a given temperature T, the mixing free energy ΔF in the GQCA is given as

$$\Delta F = \Delta E - T\Delta S = \sum M_j \epsilon_j - T\Delta S , \tag{1}$$

where M_j is the number of clusters in the alloy with excess cluster energy ϵ_j and S is the configurational entropy,

$$\Delta S = k \ln g , \tag{2}$$

where g is given as

$$g = \frac{N!}{N_A! N_B!}\left[\frac{M!}{\prod_j M_j!}\prod_j (x_j^0)^{M_j}\right] . \tag{3}$$

In Eq. (3), $N_{A(B)}$ are the number of A (B) atoms in the alloy with $N = N_A + N_B$, x_j^0 is the *a priori* fraction of clusters of the j type (random distribution), and M is the total number of clusters in the alloy (e.g., $M = \sum M_j$). Using Eqs. (2) and (3), Eq. (1) can be rewritten as

$$\Delta F = M\left[\sum \epsilon_j x_j + kT\sum x_j(\ln x_j/x_j^0)\right] + NkT[x\ln x + (1-x)\ln(1-x)] , \tag{4}$$

where $x_j = M_j/M$ is the fraction of clusters with energy ϵ_j.

The equilibrium values of the probabilities $\{x_j\}$ are obtained by minimizing ΔF with respect to x_j while obeying the following constraints on the alloy composition

$$\sum x_j = 1 , \tag{5}$$

$$\sum n_j(B)x_j = nx ,$$

where n is the number of alloying atoms (A and B) per cluster. The result of a variational calculation (equivalent to a steepest descent argument given in Ref. 1) of ΔF in Eq. (4) with respect to x_j subject to the constraints of Eq. (5) leads to the following jth order equation for an effective activity coefficient (η)

$$nx = \sum \eta^j e^{-\epsilon_j/kT} g_j/q , \tag{6}$$

where q is the single-cluster grand partition function,

$$q = \sum \eta^j g_j e^{-\epsilon_j/kT} , \tag{7}$$

and g_j is the number of ways of arranging the alloying atoms in a cluster with energy ϵ_j,

$$g_j = n!/j!(n-j)! . \tag{8}$$

Once η is found, the cluster population probabilities are obtained via the equation

$$x_j = g_j \eta^j e^{-\epsilon_j/kT}/q , \tag{9}$$

 0734-2101/88/042643-07$01.00

which should be compared with the random distribution

$$x_j^0 = g_j x^n (1-x)^{n-n_j}. \quad (10)$$

A variational calculation with respect to alloy volume is not performed because it has been shown experimentally[5,6] that the corresponding change in alloy volume with concentration follows Vegard's law. Calculations done[4] show that the deviation of average bond lengths from the simple weighted average is <0.01 Å for all III–V and II–VI alloys studied. The free energy can now be calculated from Eq. (4), once the cluster energies $\{\epsilon_j\}$ are known.

III. CLUSTER ENERGIES AND BULK MODULUS

The alloy interactions in a real pseudobinary cannot be accurately described by the simple fixed-value pair energies often assumed in statistical models. This is particularly true for an alloy with sizable bond-length difference (e.g., $Hg_{1-x}Zn_xTe$), where the strain energy dominates the mixing enthalpy. To write the ΔE as a sum of cluster energies $\{\epsilon_j\}$ in the GQCA, each effective ϵ_j must include all the interactions inside the cluster plus one-half of all the interactions across the cluster. This can be achieved by embedding a cluster in an effective alloy medium.[4] The smallest cluster to use is a four-bond cluster centered at the alloying A and B atoms. This choice of cluster, however, does not yield information about the local correlation. The next smallest reasonable cluster that considers local correlation is the 16-bond cluster with a central C atom and four alloying atoms bound to the environment by 12 second-neighbor C atoms (Fig. 1). Note that this 16-bond cluster differs from the five-atom, four-bond cluster used by other workers.[7–9] The 16-bond cluster choice means that the counting scheme for the configurational entropy given by Eq. (3) is exact, since no two clusters share the same alloying atoms. Approximations in the theory occur in the cluster energy calculation, where the cluster–cluster interactions are treated as a mean-field theory.

For the cluster energy calculation, $\{\epsilon_j\}$, we place the 16-bond cluster in an alloy environment and then allow the cluster to relax. The resulting strain energy associated with bond length lattice mismatches is treated with the valence force field[10] (VFF) as in Ref. 11 and other cluster model calculations. The strain energy in the VFF is expressed as

$$\epsilon_{\text{strain}} = \frac{3}{8d^2}\sum_i \alpha_i (\Delta d_i \cdot d_i)^2 + \frac{3}{8d^2}\sum_i \sum_j \beta_{ij} (\Delta d_i \cdot d_j)^2, \quad (11)$$

where d is the average alloy bond length, $\Delta d_i \cdot d_j$ is the change of the scalar product of two bond vectors due to distortion, and α and β are the two-body radial and three-body angular force constants, respectively.

The excess energies due to the change in chemical environment in an alloy are treated using Harrison's model,[12] where it is known as the metallicity. For a pair of AC and BC bonds, this chemical energy is

$$\epsilon_{\text{chem}}(\text{A,B}) = \Delta\epsilon_m(\text{A}) + \Delta\epsilon_m(\text{B}), \quad (12)$$

where

$$\Delta\epsilon_m(\text{A}) = 2\left(\frac{|\langle a_B | H | b_A\rangle|^2}{\epsilon_b(\text{A}) - \epsilon_a(\text{B})} - \frac{|\langle a_A | H | b_A\rangle|^2}{\epsilon_b(\text{A}) - \epsilon_a(\text{A})}\right) \quad (13)$$

and a similar expression for $\Delta\epsilon_m(\text{B})$. In Eq. (13), A and B represent AC and BC bonds, respectively, and ϵ_b and ϵ_a are the energies of the bonding and antibonding states denoted by $|b\rangle$ and $|a\rangle$, respectively.

To obtain the cluster energies, we are required to solve the problem of finding the equilibrium configurations of the atoms in the 16-bond cluster embedded in an infinite alloy medium with correct bond lengths and elastic constants.[4] While we have yet to carry out the numerical problem to its desired accuracy, we have investigated two different approximations which simulate the alloy environment. For the first medium, hereafter referred to as the *discrete model*, the cluster is placed in a medium in which all atoms on the third shell and beyond are fixed at their virtual-crystal approximation (VCA) positions. It was shown in Ref. 11 that the truncation of the relaxation medium overestimates the actual strain. To counteract this overestimation, the bond-angle terms in the VFF across the cluster are discarded, while those within the cluster are retained. This same procedure was done for the case of impurities in alloys with good results for the bond lengths and mixing enthalpies.[11] The second medium, or *elastic continuum* model, replaces the atoms outside the second shell by an elastic medium. We further assume that all atoms in the elastic medium (i.e., second-shell atoms and beyond the second-shell radius R) only have radial displacements which are inversely proportional to the square of the radius. The strain energy for the elastic continuum is now given as[11]

$$\epsilon_{\text{strain}} = RCu^2, \quad (14)$$

where u is the displacement at R and C is an effective shear stiffness coefficient given by $C = 1.6\pi(C_{11} - C_{12} + 3C_{44})$. Table I lists the values of the bond lengths d, the force constants α and β, and the coefficients C for the three II–VI systems studied.

In the calculation of a particular cluster energy, the sum of the chemical energies within the cluster plus the strain ener-

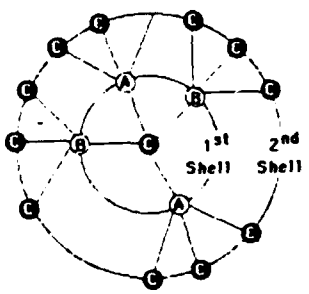


FIG. 1. The basic microcluster unit used for $A_{1-x}B_xC$ zinc-blende compounds. It includes a central C atom, four neighboring A and B alloying atoms in the first shell, and 12 second neighbors to the central atom in the second shell (shared with other clusters).

TABLE I. The bond lengths d, force constants α and β, and effective shear coefficients C for the three II–VI compounds. (For a detailed description of these values, see Ref. 11.)

	ZnTe	CdTe	HgTe
d(Å)	2.637	2.806	2.789
α(N/m)	29.45	26.57	26.40
β(N/m)	4.66	2.72	2.75
C ($\times 10^{11}$ ergs/cm^3)	62.4	38.4	40.4

gies of th cluster and the medium is varied with respect to the posit is of the cluster atoms until a minimum value ϵ_{min} is reach The cluster energy ϵ_n for the $A(4-n)B(n)C$ cluster becomes the sum of ϵ_{min} plus one-half the averaged chemical energies of this cluster with the cluster medium:

$$\epsilon_n = \epsilon_{min} + 9/2[n(1-x) + (4-n)x]\epsilon_{chem}(A,B) . \quad (15)$$

The long-range Coulomb interactions[2] were also estimated and found to be very small compared to the strain and chemical energies. Energy corrections that occur due to shifts in d-level energies via first-neighbor p–d interactions with antibonding states were also estimated and found to be small (<0.1 kcal/mol). In this p–d interaction, it is surmised that the shifts due to the cation d levels counterbalance the shifts due to the anion occupied p orbitals, leaving the average energy relatively intact.

For the calculation of the bulk modulus[4] B, one calculates the change of the average energy response of the clusters in a compressed medium (with external pressure δP)

$$\delta E = M(\delta\epsilon_n) , \quad (16)$$

which in turn yields the bulk modulus from the relationship

$$\delta E = (\delta P)^2 V/(2B) , \quad (17)$$

where V is the volume of the alloy. Results[4] show the bulk modulus of the pseudobinary alloy as a function of concentration bows only slightly from a straight-line average ($\overline{B}$),

$$B = \overline{B}[1 - 3x(1-x)(\Delta B/\overline{B})^2] , \quad (18)$$

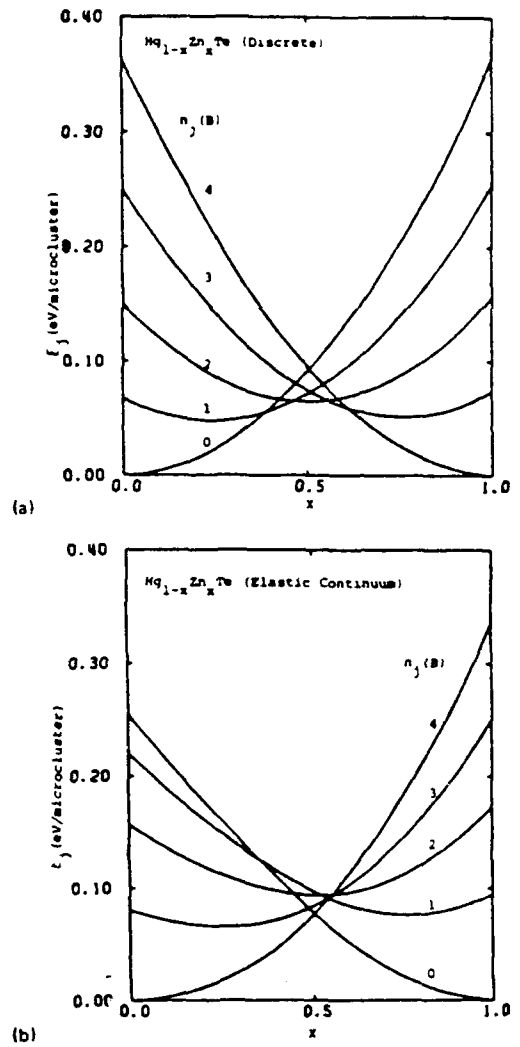


FIG. 2. $Hg(4-j)Zn(j)Te$ cluster energies per 16-bond microcluster as a function of ZnTe concentration for clusters with differing numbers of Zn atoms. (a) discrete model and (b) elastic continuum model

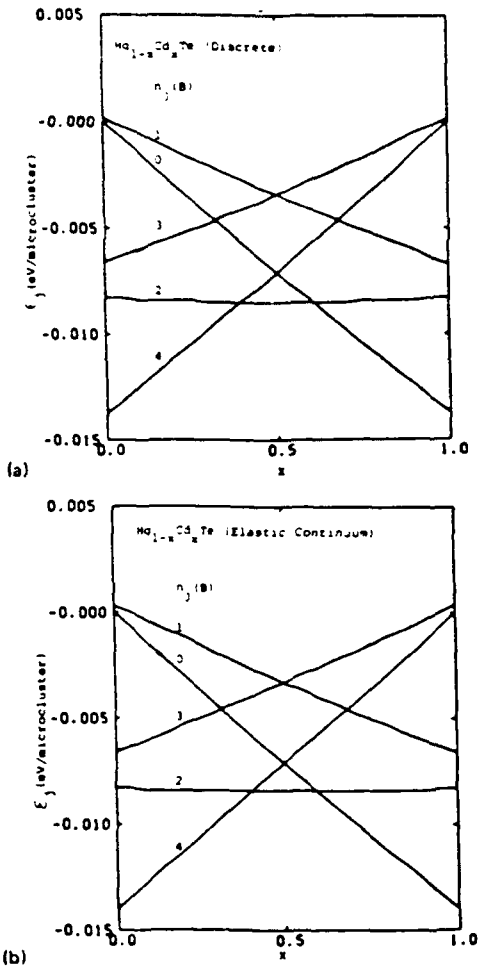


FIG. 3. $Hg(4-j)Cd(j)Te$ cluster energies per 16-bond microcluster as a function of CdTe concentration for clusters with differing numbers of Cd atoms. (a) discrete model and (b) elastic continuum model

where the fractional change in the bulk modulus is $< 1\%$ for all cases.

IV. RESULTS

Figures 2(a) and 2(b) show the cluster energies per microcluster of $Hg_{1-x}Zn_xTe$ for microclusters with differing numbers of Zn atoms. Figure 2(a) uses the discrete model while Fig. 2(b) uses the elastic continuum model for the alloy medium. Note that while the energies for the discrete model are almost symmetric, those from the elastic continuum model are rather asymmetric about the stoichiometric composition $x = 0.5$. This is due to the fact that the averaged α value varies slowly while the averaged C value changes considerably from HgTe to ZnTe (see Table I). Comparison of the cluster energies between the two models for $Cd_{1-x}Zn_xTe$ is similar to that for $Hg_{1-x}Zn_xTe$. The results for $Hg_{1-x}Cd_xTe$, however, are different[13] [Figs. 3(a) and 3(b)] from the other two alloys in that the strain energy is not dominant. Since there is very little variation of the β values across the alloy concentration, the cluster energies are more or less symmetric about $x = 0.5$ for both models.

Figures 4(a) and 4(b) show $Hg_{1-x}Zn_xTe$ deviation from randomness $\Delta x_j = \bar{x}_j - x_j^0$ for both alloy models at 1000 K. Similar results were obtained for $Cd_{1-x}Zn_xTe$. Note that while both discrete and elastic continuum models predict very small deviations from a random alloy, the tendency of a particular cluster to be greater or less than a random distribution differs for each model. Although the signs of Δx_j for $Cd_{1-x}Zn_xTe$ in the elastic continuum model agree with those obtained from a measurement of the chemical shifts,[14] the magnitudes of the experimental Δx_j appear to be much larger than the calculated values. On the other hand, the discrete model predicts a small reverse trend. If the experiments are correct, then there is some major interaction missing from all models published to date. The cluster populations for $Hg_{1-x}Cd_xTe$ (discrete model) are given in Fig. 6

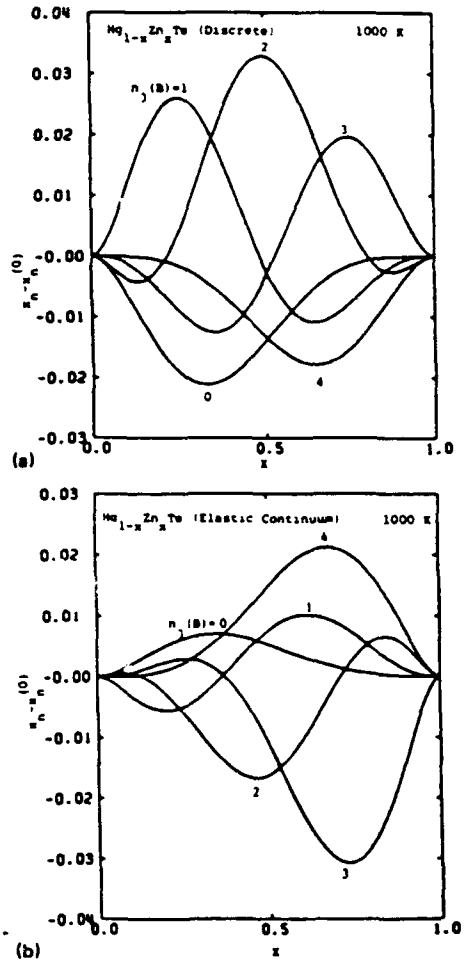


FIG. 4. Deviations of the $Hg(4-j)Zn(j)Te$ cluster population from those of a random alloy as a function of ZnTe concentration for clusters with differing numbers of Zn atoms: (a) discrete model and (b) elastic continuum model.

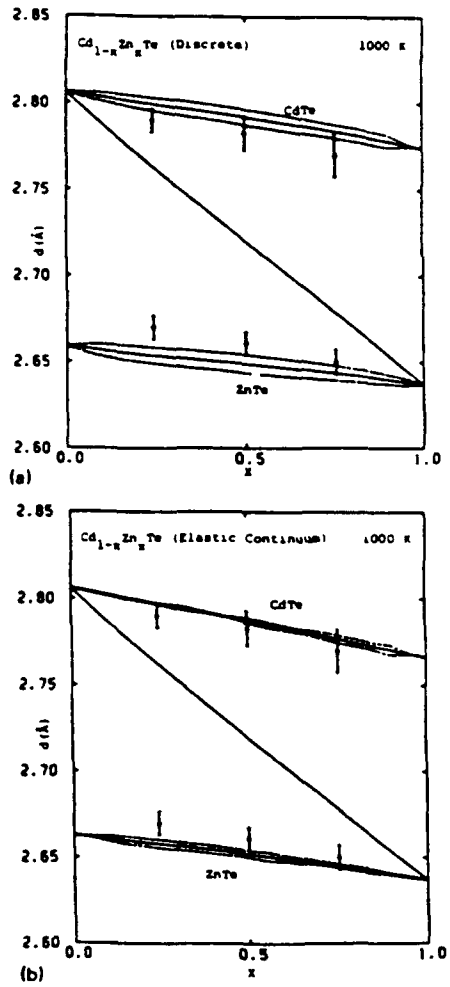


FIG. 5. $Cd_{1-x}Zn_xTe$ alloy average CdTe and ZnTe bond lengths as a function of ZnTe concentration. The solid labeled curves are the average values, the dashed curves are the extent of the rms variations, and the heavy solid line is the average bond length corresponding to Vegard's rule: (a) discrete model and (b) elastic continuum model. Experimental data from Ref. 15

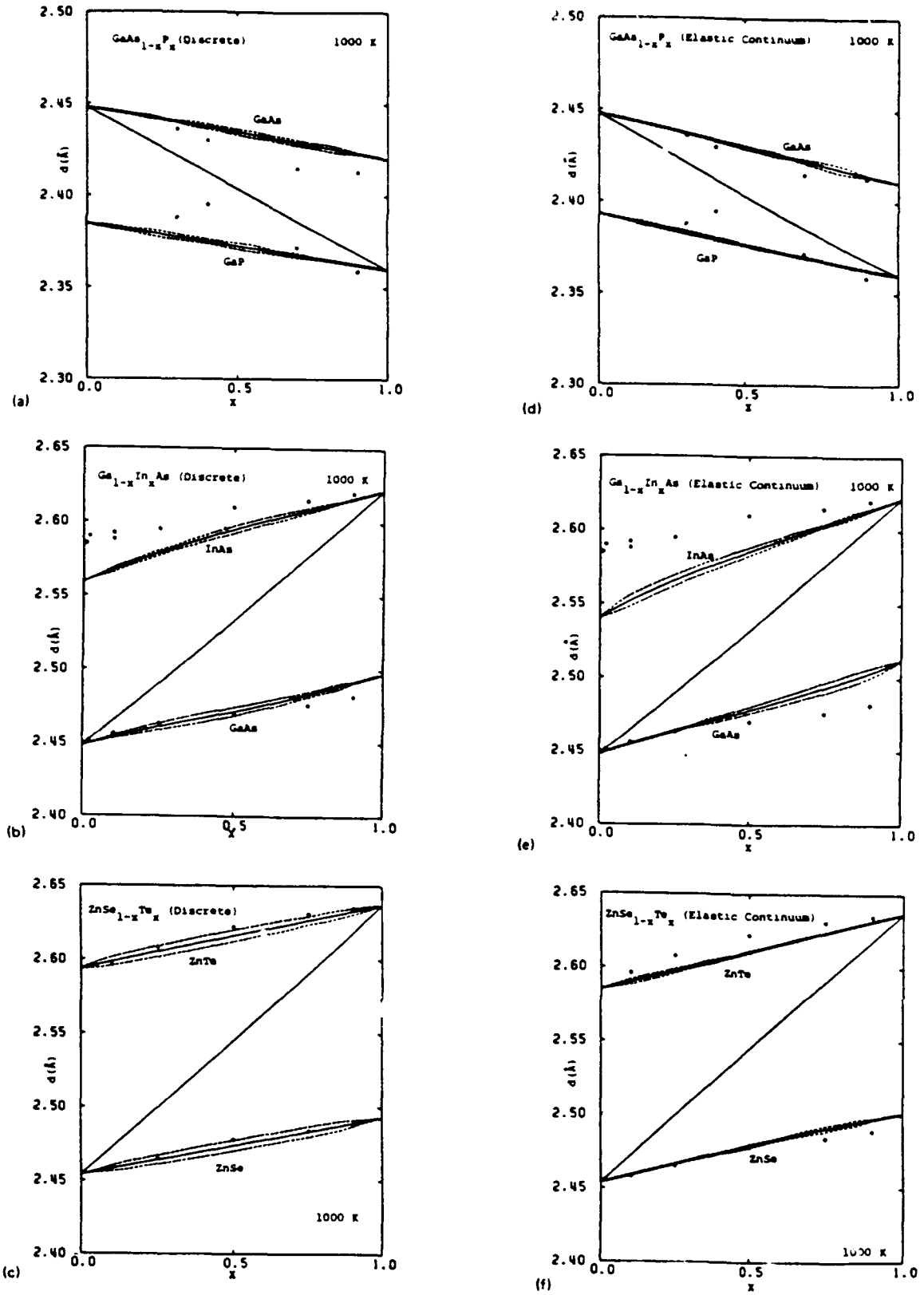


FIG. 6. A, , B, C alloy average constituent bond lengths as a function of B concentration. The solid labeled curves are the average values, the dashed curves are the extent of the rms variations, and the heavy solid line is the average bond ': ', corresponding to Vegard's rule. (a) GaAs, , P, , discrete model, data from Ref. 16; (b) GaAs, , P, , elastic continuum model, data from Ref. 16; (c) In, , Ga, As, discrete model, data from Ref. 5; (d) In, , Ga, As, elastic continuum model, data from Ref. 5; (e) Zn, , Se, Te, discrete model, data from Ref. 6; and (f) Zn, , Se, Te, elastic continuum model, data from Ref. 6.

of Ref. 13. [Please note a correction in Figs. 5 and 6 of Ref. 13 in that the cluster labels $n_j(\mathrm{Cd}) = 4$ and $n_j(\mathrm{Cd}) = 0$ in both figures are reversed.] While there is still a small deviation from randomness, the overall trend of the cluster populations is different than the discrete results presented here for $Hg_{1-x}Zn_xTe$ and the experimental results of Ref. 14. Since the magnitudes of the calculated cluster populations are very small and the experimental results for $Hg_{1-x}Cd_xTe$ are not as simple as for $Hg_{1-x}Zn_xTe$, we have not attempted to make a comparison.

The $Cd_{1-x}Zn_xTe$ bimodal bond-length distributions at 1000 K along with the rms deviations and average values are shown in Figs. 5(a) and 5(b) for both alloy models. Comparison of the two shows that the elastic continuum model is closer to the experimental values[15] than the discrete model. Similar bimodal results were obtained for $Hg_{1-x}Zn_xTe$, but different results for $Hg_{1-x}Cd_xTe$, in which the shorter HgTe bond gets shorter and the larger CdTe bond gets longer in the alloy. The average bond lengths for some III–V alloy and II–VI alloys for which experimental data are available[5,6,15,16] are shown in Figs. 6(a)–6(f). The only conclusion that we can draw from these comparisons is that both models gave a semiquantitative prediction of the bond lengths. The data cannot be used to disqualify either model.

The phase diagrams for the three II–VI alloys using the discrete model are shown in Figs. 7–9. The solid free energies were calculated from the GQCA while for the liquid free energies we have adopted a simple regular solution with a single adjustable mixing enthalpy parameter [e.g., $\Delta E = \Omega x(1-x)$]. The solidus/liquidus lines were then obtained by using the common tangent method[3] along with Vieland's formula but with no further adjustable parameters.[17] The parameters used for each phase diagram calculation are given in the figure captions. In the $Cd_{1-x}Zn_xTe$ and $Hg_{1-x}Zn_xTe$ cases, the discrete model gave slightly better agreement with experimental data[18–21] than the elastic continuum model. For $Hg_{1-x}Cd_xTe$, where the strain energy is small, the two models give the same curves.

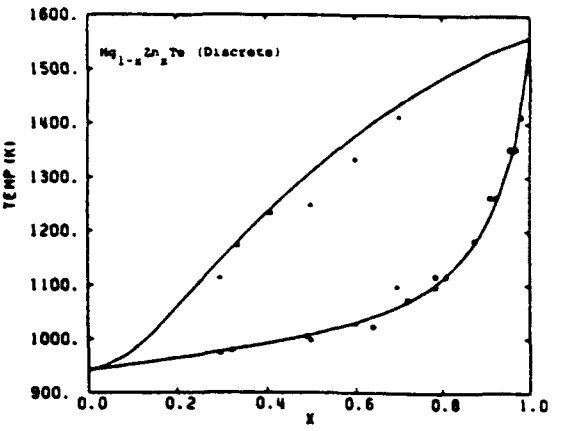


FIG. 8. $Hg_{1-x}Zn_xTe$ phase diagram (discrete model). Values for the liquid interaction parameter, melting temperature, and heats of fusion are $\Omega = -3.50$ kcal/mol, $T^m_{HgTe} = 943$ K, $T^m_{ZnTe} = 1560$ K, $\Delta H^m_{HgTe} = 8.727$ kcal/mol, and $\Delta H^m_{ZnTe} = 15.600$ kcal/mol. Experimental data. Δ Ref. 19. ● Ref. 20.

V. SUMMARY AND DISCUSSION

A generalized quasichemical approximation has been developed to treat the local correlation and free energies for semiconductor alloys. The two approximate alloy models considered produce semiquantitative results for bond lengths, mixing energies, and phase diagrams. While both models predict small deviation from randomness for all three II–VI alloys studied here, the population is sensitive to the model choice. However, the procedure has not been carried out for the infinite medium. Future work will incorporate more exact models for the alloy medium. In addition, better liquid free energy models should be used for the phase diagram calculation. However, more systematic, detailed, and accurate experimental measurements on these quantities are needed to provide checks on the theory. Further-

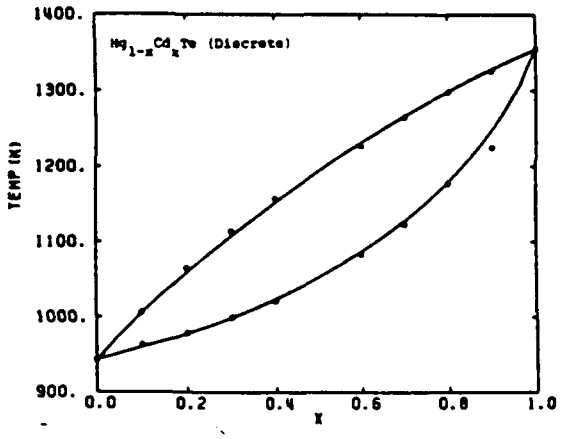


FIG. 7. $Hg_{1-x}Cd_xTe$ phase diagram (discrete model). Values for the liquid interaction parameter, melting temperature, and heats of fusion are $\Omega = -1.00$ kcal/mol, $T^m_{HgTe} = 943$ K, $T^m_{CdTe} = 1365$ K, $\Delta H^m_{HgTe} = 8.727$ kcal/mol, and $\Delta H^m_{CdTe} = 12.012$ kcal/mol. Experimental data from Ref. 18.

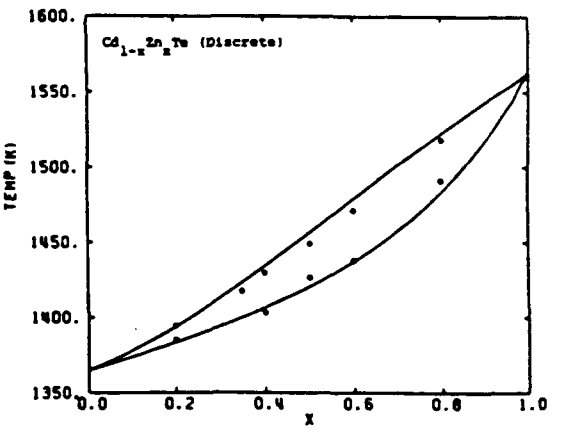


FIG. 9. $Cd_{1-x}Zn_xTe$ phase diagram (discrete model). Values for the liquid interaction parameter, melting temperature, and heats of fusion are $\Omega = +0.160$ kcal/mol, $T^m_{CdTe} = 1365$ K, $T^m_{ZnTe} = 1560$ K, $\Delta H^m_{CdTe} = 12.012$ kcal/mol, and $\Delta H^m_{ZnTe} = 15.600$ kcal/mol. Experimental data from Ref. 21.

more, the theoretical results (particularly the atomic distributions and the bond lengths) should be used to study more properties, e.g., surface ordering and segregation, phonon spectra, deep levels, transport, nuclear magnetic resonance, etc., to have a broader check and application of the theory.

ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with Dr. Tse Tung of SBRC. This work was supported in part by the U.S. Air Force Office of Scientific Research under Contract No. F49620-85-C-0023, the Office of Naval Research under Contract No. N00014-85-K-0448, and by NASA Grants No. NAG-1-708 and No. NAS1-18232.

[1] A. Sher, M. van Schilfgaarde, A.-B. Chen, and W. Chen, Phys. Rev. B **36**, 4279 (1987).
[2] M. van Schilfgaarde, A.-B. Chen, and A. Sher, Phys. Rev. Lett. **57**, 1149 (1986). In this manuscript, a factor of $\frac{1}{2}$ was missing in Eq. (6) for the nearest-neighbor Madelung energy. The correct formula for this term should be $6e^2/(16\sqrt{8/3}d)$. The correct values for U_{eff} for the II–VI systems listed in Table I now read: ZnCdTe (0.04), ZnHgTe (0.00), and CdHgTe (-0.01). S.-H. Wei, Phys. Rev. Lett. **59**, 2613 (1987); M. van Schilfgaarde, A.-B. Chen, and A. Sher, *ibid.* **59**, 2614 (1987).
[3] R. Patrick, A.-B. Chen, and A. Sher, Phys. Rev. B **36**, 6585 (1987).
[4] A.-B. Chen, A. Sher, and M. A. Berding, Phys. Rev. B **37**, 6285 (1988).
[5] J. C. Mikkelsen, Jr. and J. B. Boyce, Phys. Rev. B **28**, 7130 (1983).
[6] J. B. Boyce and J. C. Mikkelsen, Jr., in *Proceedings of the 7th International Conference on Ternary and Multinary Compounds* (Materials Research Society, Pittsburgh, 1986), p. 353.
[7] M. T. Czyzk, M. Podgorny, A. Balzarotti, P. Lelardi, N. Motta, A. Kisiel, and M. Zemnal-Slamawska, Z. Phys. B **62**, 153 (1986).
[8] A. A. Mbaye, L. G. Ferreira, and A. Zunger, Phys. Rev. Lett. **58**, 49 (1987).
[9] M. Ichimura and A. Sasaki, J. Appl. Phys. **60**, 3850 (1986).
[10] R. M. Martin, Phys. Rev. B **1**, 4005 (1970).
[11] A.-B. Chen and A. Sher, Phys. Rev. B **32**, 3695 (1985).
[12] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).
[13] Note that the energies in Fig. 3(a) are different from those shown in a previous publication [A. Sher, M. A. Berding, A.-B. Chen, and W. Chen, in *Proceedings of the Materials Research Society Symposium on R-Materials for Infrared Detectors and Sources* (Materials Research Society, Pittsburgh, 1987), Vol. 90, p. 91], because the effective chemical energies of the clusters with the environment [2nd term of Eq. (15)] have been included in the present case. These energies will affect the magnitude of ΔE but will not affect the cluster populations.
[14] D. Zamir, K. Beshah, P. Becla, P. A. Wolff, R. G. Griffin, D. Zax, S. Vega, and N. Yellin, J. Vac. Sci. Technol. A **6**, 2612 (1988) (these proceedings); D. Zax (private communication).
[15] A. Balzarotti, in Ref. 6.
[16] T. Sakai, T. Onda, R. Ito, and N. Ogasawara, Jpn. J. Appl. Phys. **25**, 231 (1986).
[17] L. J. Vieland, Acta. Metall. **11**, 137 (1963).
[18] F. R. Szofran and S. L. Lehoczky, J. Electron. Mater. **10**, 1131 (1981).
[19] E. Cruceanu and N. Nistor, J. Electrochem. Soc. **113**, 955 (1966).
[20] J. C. Wooley and B. J. Ray, J. Phys. Chem. Solids **13**, 151 (1960).
[21] J. Steininger, A. J. Strauss, and R. F. Brebrick, J. Electrochem. Soc. **117**, 1305 (1970).

Journal of Crystal Growth 98 (1989) 27–36
North-Holland, Amsterdam

CORRELATIONS AND ALLOY PROPERTIES: GROWTH, VACANCIES, SURFACE SEGREGATION *

A. SHER and M.A. BERDING

Physical Electronics Laboratory, SRI International, Menlo Park, California 94025, USA

and

A.-B. CHEN and R.S. PATRICK

Physics Department, Auburn University, Auburn, Alabama 36849, USA

The properties of alloy semiconductors are influenced by the state of order of the materials. Some properties, such as the liquidus and solidus curves, are relative insensitive but others are strongly affected. We first demonstrate which aspects of bonding interaction mechanisms impact the correlation state. We then deduce the effect of the correlation state on liquidus and solidus curves, on surface segregation, and on vacancy densities. Surface segregation is a concentration variation of the alloy constituents away from surfaces, and is driven by Schottky vacancy-free energy differences between the constituents in the bulk and on a surface. The segregation profile for alloys equilibrated above a critical spinodal transition temperature exhibits a monatomic variation decaying over a few atomic layers from the surface to the bulk value. However, long-range concentration oscillations are predicted below the critical temperature. These oscillations may account for the long-range order observed in some MBE- and MOCVD-grown alloys (e.g., InPSb). The total free energy of this new ordered phase lies between that of a homogeneous but highly correlated alloy and that of a spinodally decomposed alloy. The phase, if it exists, would be locked into its metastable state by the presence of the surface. Boundary conditions at the interface between the epitaxial layer and the substrate would also affect the results.

1. Introduction

Many groups have been studying the correlation state of tetrahedrally coordinated pseudo-binary semiconductor alloys [1–4] of the form $A_{1-x}B_xC$. They all begin by decomposing the solid into microclusters, usually five atoms consisting of a central C atom and its four surrounding neighbors. The excess energy of the cluster relative to the concentrated weighted average of similar clusters of the AC and BC compounds is then calculated in a variety of formalisms and, while there are similarities in the gross trends, the groups generate rather different detailed excess-energy-versus-concentration curves. These excess energies are then inserted into statistical mechanics expressions for the populations of the various microclusters, and the excess free energy is then calculated. While all the groups use grand canonical ensembles for the microclusters, they differ in their treatment of the entropy that appears in the free energy. A number of alloy properties are calculated, e.g., the alloy bond-length variation [1–8], the mixing enthalpy [1–8], the critical order–disordered transition temperature and low-temperature phases [6], the cluster populations as a function of temperature [1–8], and liquidus and solidus phase diagrams [9]. Despite the differences in calculation methods, the bond-length variations and mixing enthalpies predicted by the different groups are nearly the same and agree reasonably well with experiments [5,10]. There are larger differences among the predictions of other properties: unfortunately, they are also less susceptible to direct experimental measurement.

* This work was supported by AFOSR Contract F49620-88-K-0009, NASA Contrast NAS1-18226, ONR Contract N00014-88-C-0096, and NASA Grants NAG-1-708 and NAS1-18232.

0022-0248/89/$03.50

In this paper we first identify the reasons for a lack of sensitivity to some details of the theory, and which aspects of the theory must be correct to make accurate correlation-state predictions. We then display some solidus and liquidus curves and examine the sensitivity of distribution coefficients to features of the theory. Next, we present vacancy energies for several alloys and draw tentative conclusions about their effects on diffusion, doping, and ion beam processing. Finally, we present features of the theory of surface segregation in semiconductor alloys.

2. Correlations

2.1. General theory

We and others [2–4] have demonstrated that there are always correlations in alloy semiconductors. Here we summarize our previous work, focusing attention on those aspects that may help us to identify what may be missing from the interactions driving correlations and those features to which correlations are insensitive. We consider an average population $\bar{x}_j$ of five-atom clusters including sixteen bonds (most others treat four bonds only). Five-atom clusters of the form $A_{4-n_j}B_{n_j}C$ can be shown to be distributed in grand canonical ensembles of the form

$$\bar{x}_j = g_j\, e^{(\mu_B n_j - \epsilon_j)/kT}/q(\{\epsilon_j\}, \mu_B), \tag{1}$$

where the partition function is

$$q(\{\epsilon_j\}, \mu_B) \equiv \sum_{j=0}^{J} g_j\, e^{(\mu_B n_j - \epsilon_j)/kT}, \tag{2}$$

and g_j is the degeneracy of cluster $j = 0, 1, \ldots, J (4 \le J \le 2^4)$, n_j is the number of B atoms on cluster j, ϵ_j is the excess energy of cluster j relative to the average $\bar{\epsilon} = (1-x)\epsilon_A + x\epsilon_B$ (called Δ_j by Sher et al. [1]), ϵ_A and ϵ_B are cluster energies of the pure AC and BC compounds, k is the Boltzmann constant, and T is the absolute temperature.

If the clusters do not have their normal counting degeneracies $g_j^0 = \binom{4}{n_j}$ split, then we have $g_j = g_j^0$, the binomial coefficient, and $J = 4$. The chemical potential is set by insisting that the composition x is correct.

$$4x = \sum_j n_j \bar{x}_j. \tag{3}$$

If we make a transformation to the reduced excess-energies representation,

$$\Delta_j = \epsilon_j - (1 - \tfrac{1}{4}n_j)\epsilon_0 - \tfrac{1}{4}n_j\epsilon_4 \tag{4}$$

(called Δ_j' by Sher et al. [1]), then $\bar{x}_j$ becomes

$$\bar{x}_j = g_j\, e^{n_j(\mu_B' - \Delta_j)}/q(\{\Delta_j\}, \mu_B') \tag{5}$$

where

$$\mu_B' = \mu_B - \tfrac{1}{4}(\epsilon_4 - \epsilon_0). \tag{6}$$

Note that we have $\Delta_0 = \Delta_4 = 0$, so at least two members of the set $\{\Delta_j\}$ vanish. More important, only the reduced excess energies actually drive the cluster populations. We shall see shortly that the energies of the $\{\Delta_j\}$ set are often five to ten times smaller than those of the $\{\epsilon_j\}$ set. Most of the strain contribution to the $\{\epsilon_j\}$ set subtracts out and contributes little to the $\bar{x}_j$ values.

There is an additional effect. It can best be appreciated when stated as a theorem. If (a) $\Delta_j = \Delta_0 + n_j\delta$, and if (b) $g_j = \binom{4}{n_j}$, then $J = 4$ and

$$\bar{x}_j = x_j^0 \equiv g_j^0(1-x)^{4-n_j}x^{n_j}, \tag{7}$$

the populations x_j^0 of a random alloy. Thus, no matter how large the constant δ may be, as long as the energies Δ_j vary linearly with the number of B atoms on the cluster n_j and the counting degeneracy is not split, the alloy is random. In fact $\delta = 0$, since we have $\Delta_0 = \Delta_4 = 0$. Strain energies from bond-length differences between the constituents, chemical energies, and electron-electron interactions all have some nonlinear variations with n_j. The degeneracy g_j^0 is split by coherent strains or temperature gradients, anything that establishes preferred directions for the locations of A and B atoms on a cluster. These splittings always establish a preference for one type of cluster in a particular orientation (e.g., for a strain in the (111) direction, A_3B with the B oriented along the (111) axis), and therefore drive deviations of

the $\{\bar{x}_j\}$ set from $\{x_j^0\}$ toward compound formation, and if the splittings are large enough, even toward long-range order. The effect of the energies $\{\Delta_j\}$ can drive the $\{\bar{x}_j\}$ set toward either compound formation or spinodal decomposition, depending on the relative bond length and bond energy.

The excess free energy of an $A_{1-x}B_xC$ alloy can then be written as

$$\Delta F = \Delta E - T\Delta S, \tag{8}$$

where the excess energy is

$$\Delta E = M\sum_j \epsilon_j \bar{x}_j, \tag{9}$$

with M being the number of clusters. The excess entropy is

$$\Delta S = k \ln\left\{ \frac{N!}{N_A!\,N_B!} \prod_{j=0}^{J} \frac{M!}{M_j!} \times \left[g_j (1-x)^{4-n_j} x^{n_j} \right]^{M_j} \right\}, \tag{10}$$

with $M_j = x_j M$ (see Sher et al. [1] for the justification). The total number of clusters, M, is related to the number of Bravais lattice sites, N, in the crystal by $M = \frac{1}{4}N$ for a 16-bond cluster and $M = N$ for a 4-bond cluster. Eqs. (9) and (10) can be rewritten for the 16-bond cluster as

$$\Delta E = \tfrac{1}{4}N\left[(1-x)\epsilon_0 + x\epsilon_J + \sum_j \Delta_j \bar{x}_j\right]. \tag{11}$$

and in general the excess entropy is

$$\Delta S = -kN\left[(1-x)\ln(1-x) + x\ln x + \frac{M}{N}\sum_j \left(\bar{x}_j \ln \bar{x}_j - \bar{x}_j \ln x_j^0\right)\right] = \frac{N}{4T}\sum_j \Delta_j \bar{x}_j - N\frac{\mu_B'}{T}x + \frac{N}{4}\ln q\left(\{\Delta_j\}, \mu_B'\right). \tag{12}$$

Note in the first equality in eq. (12) that M/N is unity for a 4-bond cluster and 1/4 for a cluster with 16 bonds. In the 16-bond cluster, after substitutions from eqs. (5) and (7), the $(1-x)\ln(1-x) + x\ln x$ term is exactly canceled by a portion of the third term in the brackets. This occurs physically because, in the 16-bond cluster, all the bonds associated with a given substituted atoms are in the cluster; in the 4-bond case, however, bonds from each substituted atom contribute to 4 different clusters. Several features of eqs. (11) and (12) are noteworthy. First, ϵ_0 and ϵ_J are functions of x but in lowest order are temperature independent. For bond-length-mismatched alloys, they vary with x roughly as $\epsilon_0 \approx \epsilon_0^0 x^2$ and $\epsilon_J \approx \epsilon_J^0(1-x)^2$, so that we have

$$\Delta E \approx N\left[\Omega x(1-x) + \sum_j \Delta_j \bar{x}_j\right], \tag{13}$$

where the mixing enthalpy parameter, Ω, is approximately

$$\Omega \approx \tfrac{1}{4}\left[\epsilon_J^0 + (\epsilon_0^0 - \epsilon_J^0)x\right], \tag{14}$$

and nearly independent of x if $\epsilon_0^0 \approx \epsilon_J^0$, as is often the case. The second term in eq. (13) is usually small compared with the first and contains all the temperature dependence. However, when eqs. (5), (12), and (13) are inserted into eq. (8) to obtain ΔF, a term in $T\Delta S$ exactly cancels the second term in eq. (13). Thus, it is impossible to determine the temperature variation of ΔE from a measurement of ΔF. The expression for ΔF becomes

$$\Delta F = N\left[\Omega x(1-x) + x\mu_B' - \tfrac{1}{4}kT \ln q\left(\{\Delta_j\}, \mu_B'\right)\right]. \tag{15}$$

The chemical potential μ_B' is determined from eq. (3) and both μ_B' and q are generally temperature dependent.

In a much different formalism than ours [1] and under more highly constrained conditions, Ferreira et al. [4] later reached some of the same conclusions presented in this section. The result that the $\{\bar{x}_j\}$ set depends only on $\{\Delta_j\}$, and that $\{\Delta_j\} \ll \epsilon_j$ because most of the strain contributions cancel in the $\{\Delta_j\}$ set, is a direct consequence in our approach of the concentration conservation constraint, eq. (3). No assumptions restricting Δ_j to be x independent are made in our

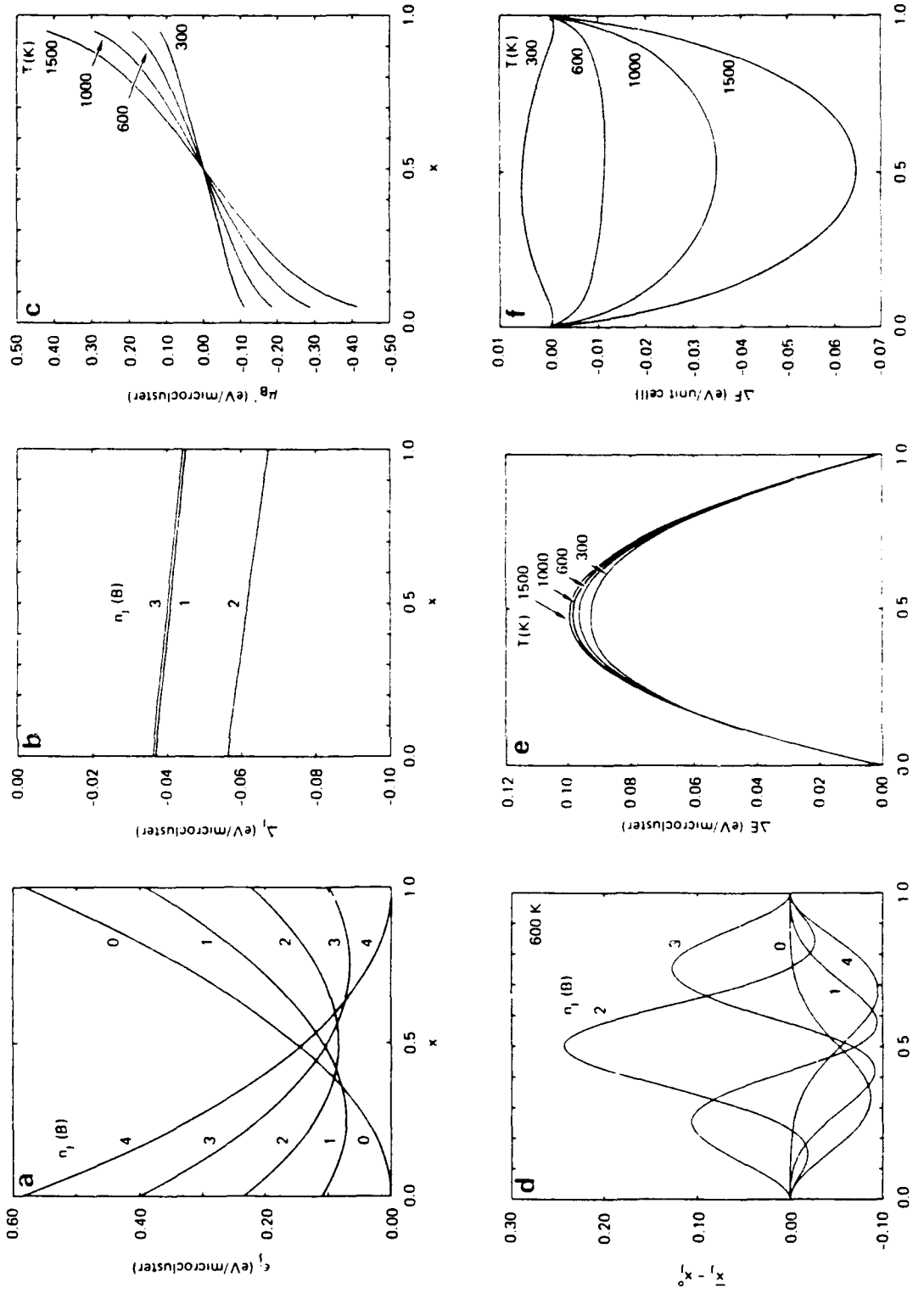


Fig. 1. Composition variation of correlation-state-dependent quantities for $Ga_{1-x}In_xAs$: (a) cluster excess energies, (b) cluster reduced excess energies, (c) reduced chemical potential at different temperatures, (d) deviation of cluster populations from random values at $T = 600$ K, (e) excess enthalpy at different temperatures, (f) excess free energy at different temperatures.

development, and in fact we find Δ_i does have a slow variation with x in the examples presented in the next section.

2.2. Numerical results

We have calculated $\{\epsilon_i\}$, $\{\Delta_i\}$, μ_B', $\{\bar{x}_i - x_i^0\}$, $\Delta E(T)$, and $\Delta F(T)$ for a number of alloys; fig. 1 shows those for $Ga_{1-x}In_xAs$. The energies ϵ_i were calculated as discussed in detail in ref. [1]. The sixteen-bond cluster was attached to rigid medium at the virtual crystal positions in the third shell from the cluster center. The atom positions inside the third shell were adjusted to minimize the strain and chemical energies. The energy ϵ_i is the total energy of the sixteen bonds with the atoms in their minimum energy positions. The variations of atom bond lengths in different alloys are well predicted by this method.

The variation of $\{\epsilon_i\}$ with x, and that of $\{\Delta_i\}$ are quite different. As is evident from the curves, the $\{\Delta_i\}$ set an order of magnitude smaller than the $\{\epsilon_i\}$ values, and most of the curvature identified with the strain contribution has subtracted out. The excess enthalpy ΔE has only a small temperature dependence, as we anticipated. The excess free energy ΔF at 600 K has a negative curvature at $x = \frac{1}{2}$, which indicates that the critical transition temperature is below 600 K.

The reasons why bond lengths and mixing enthalpies agree with experiments in a variety of quite different cluster model approximations are discussed extensively in refs. [2,8,9]. The underlying cause for the insensitivity of bond lengths is that if the region outside the cluster is allowed to relax in the energy-minimization procedure used to identify the bond lengths, the strain from the angle-restoring forces tends to be canceled. Thus, models that include angle-restoring forces and some outside relaxation or that ignore angle-restoring forces and outside relaxation both predict the bond lengths correctly. However, the two classes of models predict different $\{\epsilon_i\}$ sets. The set based on a stiff exterior (no relaxation) is larger with sharper x variations. However, as we have seen, most of this strain contribution subtracts from the $\{\Delta_i\}$ set. Thus the populations $\{\bar{x}_i\}$ are not as different in the two types of approach as might have been expected. The excess entropy ΔE should differ. However, the various groups also include chemical energies, and those using stiff-exterior approximations often report larger negative chemical terms so mixing enthalpy parameters agree with experiment. There is no general theory in disordered alloys of how to calculate cluster excess energies, for which the small parameter of the theory is defined. While a start was made on this theory [11], a more complete version is needed to make convincing progress.

We have also shown that only the small $\{\Delta_i\}$ set drives deviations in the populations $\bar{x}_i$ away from randomness. Moreover, if $\Delta_i = \delta n_i$ is linear in n_i, the number of B atoms, then despite the size of δ, the populations are still random. Thus, effects that would otherwise be considered small may compete with the larger energies retained if they have the proper nonlinearity. We are examining several possibilities not incorporated in our present results. These include effects caused by screening in the composition range where narrow-gap alloys are semimetal, direct second-neighbor chemical interactions treated in the context of Harrison's model, electron–electron Coulomb interactions driven by polarity differences between the constituents, and composition variations of the elastic constants.

3. Solidus and liquidus phase diagrams

We have reported on high-temperature phase diagrams of a number of alloys [11]; two examples are presented in figs. 2 and 3. The solidus and liquidus curves, figs. 2a and 2b, were calculated using our solid free energy, following the logic of Stringfellow and Green's [12] modification of Vieland's [13] theory and their liquid free energy. As can be seen, the agreement with experiment is excellent for GaInP, but that with GaInAs is poorer. To demonstrate the sensitivity of these curves to aspects of the theory, figs. 2b and 3b show the distribution coefficients (the ratio of the solidus to liquidus concentration at a given temperature) plotted against the resulting solid concentration for the correlated model and for an

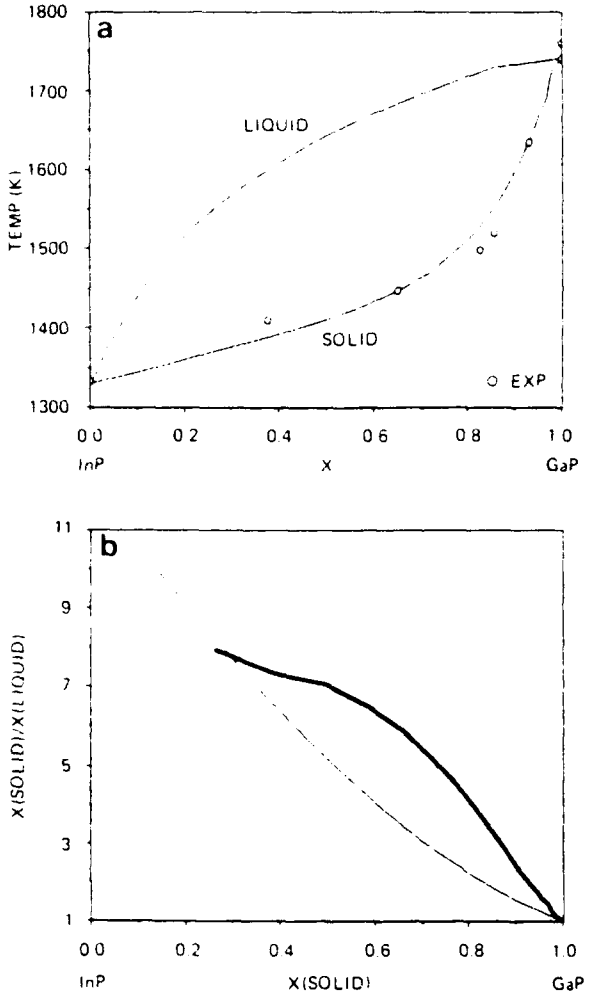


Fig. 2. $In_{1-x}Ga_xP$: (a) solidus and liquidus phase diagrams; (b) distribution coefficient x(solid)/x(liquid) as a function of x. Heavy curve, full cluster theory; light curve, ideal regular solution model.

ideal random alloy using the same enthalpy parameters. As can be seen, a regular-solution random model differs substantially from the correlated-alloy case.

4. Vacancies

Vacancy formation energies have been calculated for semiconductors and their alloys that take explicit account of the final state to which the removed atom goes as well as bond-lengths relaxation, charge redistribution, and dangling-hybrid interaction modifications to the lattice surrounding the vacated site [14,15]. The model is based on a tight-binding Hamiltonian [16] with universal scaling that yields correct cohesive energies. The extraction energies $E_v(\infty)$ where the final state is an atom at infinity, vacancy energies $E_v(111)$ and $E_v(100)$, where the atom final state is a (111) or (100) surface site, and sublimation energies $E_v(111)$ from a (111) surface are in table 1. Notice that $E_v(111)$ is not much different for the anions and cations of a given compound. To reach the ap-

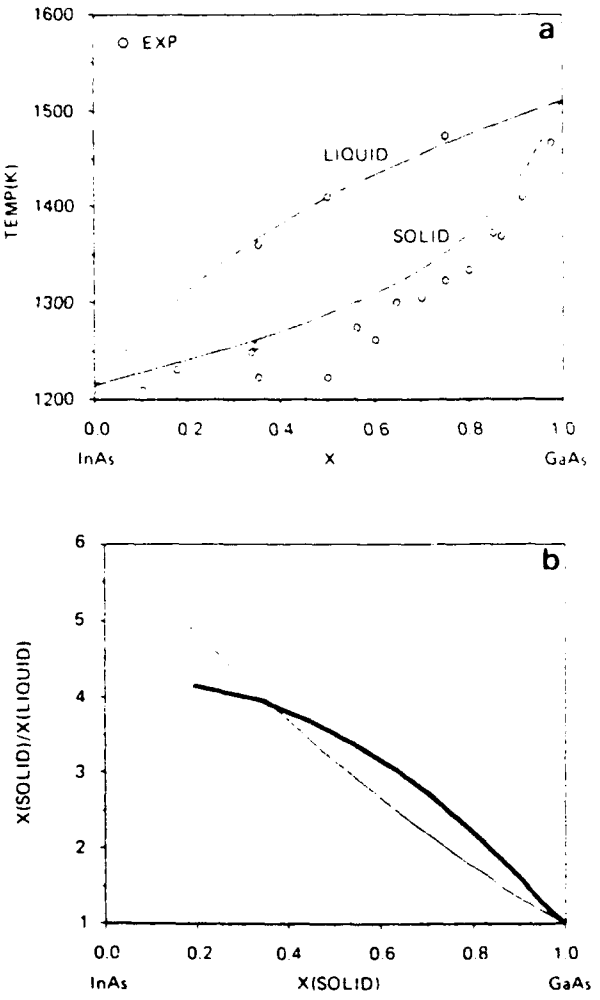


Fig. 3. $In_{1-x}Ga_xAs$: (a) solidus and liquidus phase diagrams; (b) distribution coefficient x(solid)/x(liquid) as a function of x. Heavy curve, full cluster theory; light curve, ideal regular solution model.

Table 1
Vacancy formation energies for removal of an atom from bulk to infinity, $E_v(\infty)$, to a triply bonding site on a (111) surface, $E_v(111)$, and to a doubly bonding site on a (100) surface, $E_v(100)$, and the sublimation energy from a triply bonding site on the (111) surface, $E_s(111)$; all energies are in eV

Element or compound	$E_v(\infty)$		$E_v(111)$		$E_s(111)$		$E_v(100)$ [a]	
	Cation	Anion	Cation	Anion	Cation	Anion	Cation	Anion
Carbon	19.25	19.25	5.93	5.93	13.32	13.32	11.89	11.89
Silicon	10.32	10.32	2.82	2.82	7.50	7.50	5.68	5.68
Germanium	8.11	8.11	2.10	2.10	6.01	6.01	4.23	4.23
Tin	6.30	6.30	1.59	1.59	4.71	4.71	3.18	3.18
AlP	10.65	11.12	3.55	4.11	7.10	7.01	6.62	6.62
GaP	8.30	8.76	2.62	2.83	5.68	5.93	5.92	4.02
InP	8.23	8.83	2.60	3.52	5.63	5.31	5.40	4.70
AlAs	8.79	9.65	4.74	5.10	4.05	4.55	4.85	6.03
GaAs	6.66	7.81	3.66	3.76	3.00	4.05	4.20	3.75
InAs	6.78	7.48	3.56	4.14	3.22	3.34	3.94	4.12
AlSb	4.70	6.06	2.72	2.73	1.98	3.33	2.67	3.29
GaSb	5.95	6.82	3.19	3.13	2.76	3.69	4.32	2.53
InSb	5.97	6.28	3.10	3.16	2.87	3.1	3.96	2.69
ZnS	6.90	9.99	4.21	4.19	2.69	5.80	5.00	5.53
CdS	6.09	8.78	3.72	3.72	2.37	5.06	4.17	5.02
HgS	4.53	7.78	3.08	3.08	1.45	4.70	4.35	3.88
ZnSe	5.66	8.66	3.57	3.55	2.09	5.11	4.18	4.98
CdSe	5.03	7.76	3.20	3.20	1.83	4.56	3.35	4.60
HgSe	3.39	6.77	2.49	2.52	0.90	4.25	3.37	3.39
ZnTe	5.10	8.08	2.09	2.11	3.01	5.97	4.23	4.15
CdTe	4.68	7.36	1.90	2.05	2.78	5.31	3.65	3.99
HgTe	3.10	6.36	1.50	1.51	1.60	4.85	3.57	2.61

[a] $E_v(100)$ is an estimate only and does not include back-bond relaxation at the surface.

propriate (111) surface, an atom breaks four bonds in the bulk and remakes three bonds on the surface. The differences between anions and cations arise from differences they cause on back-bonding states adjacent to the vacant site. The sublimation energies are another matter, since the promotion energies are quite different for anions and cations.

Following the systematic variations of extraction energies is a convenient way to examine trends in the alloys. The variations of $E_v(\infty)$ for different clusters are shown for the common arsenic and common gallium alloys, respectively, in figs. 4 and 5, for $x = 0.5$. The variation of the numbers with x is fairly small. There are a several noteworthy features to these curves that will significantly affect diffusion, doping, and other device processing steps. First, as would be expected, variation with cluster type is more prominent for the unalloyed atom type, which is affected by the occupation of the four near-neighbor sites; the alloyed atom character depends on the occupancy of the twelve second-neighbor sites. The variations all tend to bow down, some by large amounts. This means that vacancy concentrations in alloys will differ greatly from those that would be calculated from an average activation energy.

The hop process will also be affected. For example, in $GaP_{1-x}As_x$, it will cost more energy for a Ga on a PAs_3 cluster to hop to either a P_2As_2 or an As_4 cluster than to hop to another PAs_3 cluster. Thus, for high x (the arsenic-rich side) where the PAs_3 cluster population x_3 is large, we might expect diffusion profiles characteristic of parallel paths with different activation

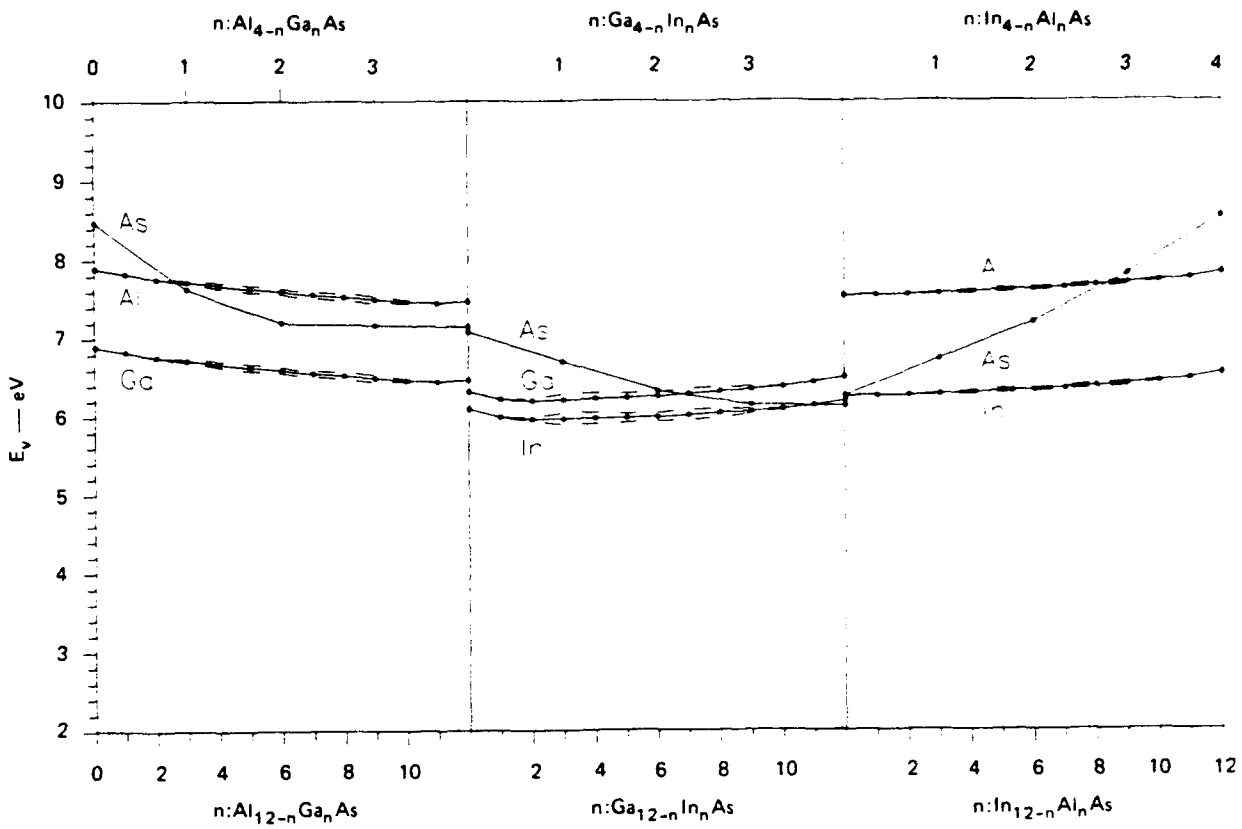


Fig. 4. Extraction energies of the arsenic common anion alloys as a function of neighboring configurations for a concentration $x = 0.5$.

energies. This will occur if $\bar{x}_3$ exceeds the percolation threshold [17]. Clearly these extraction energy curves also suggest that nonisoelectronic impurity substitution energies will vary for the different clusters. Moreover, ion milling and reactive ion etching may be quite different. For example, it has been observed [18] that in the presence of a gallium ion beam, InAs etches an order of magnitude more rapidly in a chlorine gas environment. This might be anticipated from fig. 3, since the indium extraction energy from a $GaIn_{11}$ cluster is smaller than that from an In_{12} cluster. It remains to be demonstrated that this effect, rather than some damage mechanism, is responsible for the increased etch rate. Many other effects are implied by figs. 4 and 5 as well as similar results for all the other III–V and II–VI alloys. These kinds of considerations obviously will play a role in any device processing design rules applied to alloys.

5. Surface segregation

It is well established in metal alloys that differences between the vacancy formation energies of the constituents drive concentration variations near the surface [19], but this effect has not been treated previously in semiconductor alloys. We have begun this theory by treating the solid as a succession of layers parallel to the surface [19]. The interaction energies of atoms within a layer and between layers are all included in the average enthalpy. The configuration entropy of each layer is calculated from the appropriate version of eq.

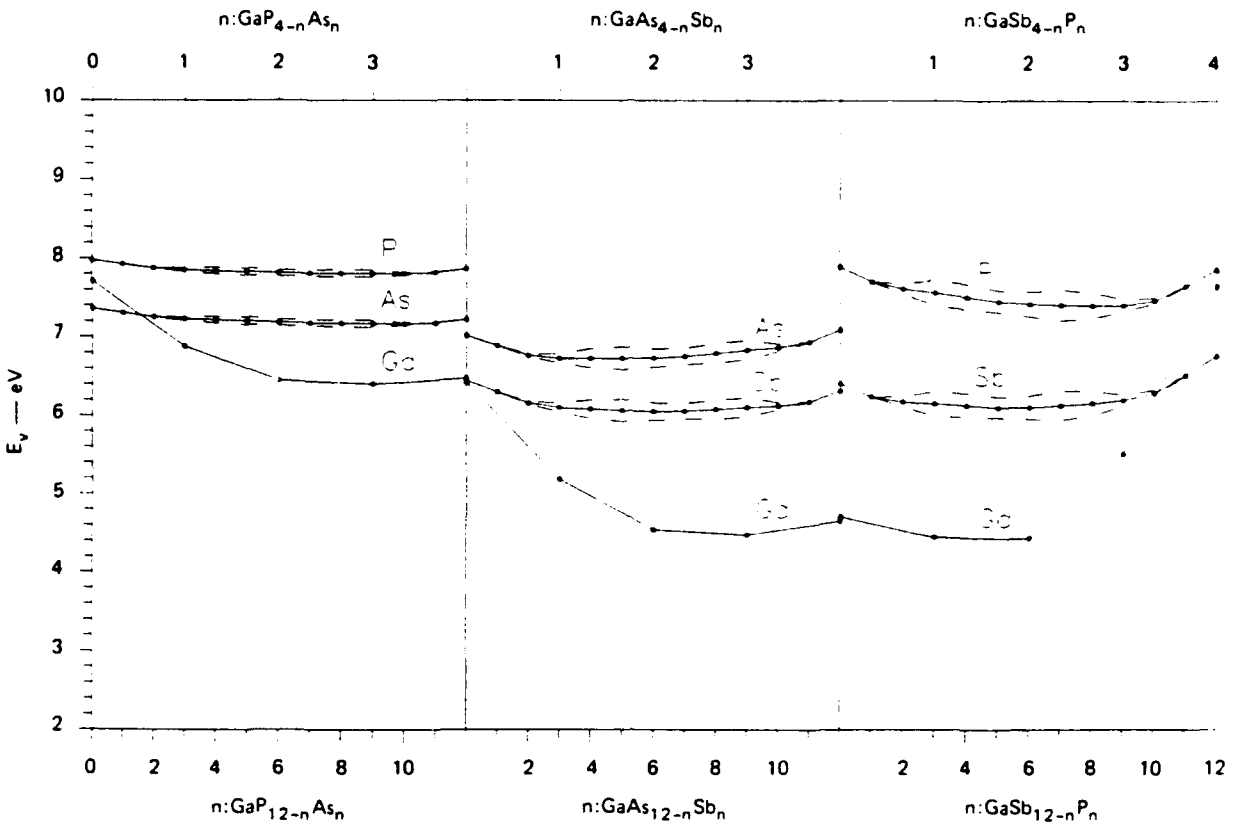


Fig. 5. Extraction energies of the gallium common cation alloys as a function of neighboring configurations for a concentration $x = 0.5$.

(10) for pair interactions and for each layer i, x is replaced by the layer concentration x_i. The concentration of each layer is then set by the condition that the chemical potential be constant.

We find several startling results. If the solid has equilibrated at a temperature $T > T_c$, the critical temperature for spinodal decomposition, then the surface concentration returns to the bulk value in a few atomic layers. However, if the equilibration temperature is below T_c, then the range of the segregation profile is much longer and the concentration will oscillate with a period and amplitude that depend on the equilibration temperature and the average composition. The oscillations can have a large modulation index and periods of 4 and 20 lattice spacings have been found. This is a superlattice phase that is stabilized by surface segregation and an order–disorder transition. We are currently examining the influence of a boundary condition corresponding to an epitaxial layer on a substrate.

It is too early in our calculations to determine whether surface segregation is responsible for any of the long-range order observed in several alloys. However, the effect is strong, being driven by the same mechanisms responsible for high-temperature order-disorder transitions, and requires relatively little diffusion to establish the relevant concentration patterns, so we anticipate that it will result in observable phenomena.

6. Concluding remarks

There are two separate aspects to the correlation state of an alloy. The first is specified by the population distribution of the different types of clusters, taking the alloy to be spatially homoge-

neous. The second is the way the alloy separates below the critical order–disorder transition temperature into a spatially inhomogeneous arrangement. The cluster population distributions are driven only by small subtle aspects of the lattice-mismatch-induced strain and chemical terms, the reduced excess energies,while the mixing enthalpy parameter that determines the order–disorder transition temperature depends on the total strain and chemical contributions to the cluster excess energies. Because only the small reduced excess energies $\{\Delta_j\}$ enter into the populations $\{\bar{x}_j\}$, the theory may have to take account of small effects not previously thought to be significant if accurate predictions are to be made.

Vacancy formation energies are sensitive to the local surroundings from which an atom is extracted, as well as the final state of the atom, e.g., the particular surface orientation and its local arrangement. From the systematic variation of the extraction energies alone, it is evident that multiple activation energy, and parallel path diffusion should be expected in alloys. Moreover, nonisoelectronic impurities are likely to have preferred cluster sites as well, although we have done no explicit calculations of dopant site preferences.

The vacancy formation energy differences between alloy-substituted atoms were also shown to drive surface segregation. For examples whose effective equilibration temperature is above the critical order-disorder transition temperature, the surface will be rich in whichever component prefers to go to the surface and the composition returns to the bulk value in a few lattice spacings. However, for samples equilibrated below the critical temperature, our preliminary results indicate that compositional oscillations may occur, producing a superlattice. Details remain to be established for various materials systems. This arrangement corresponds to a new type of metastable phase that is driven by a surface boundary condition. The free energy of this phase lies between that of a homogeneous alloy and its spinodal decomposed arrangement. Whether or not this new phase corresponds to any of the correlated arrangements [20,21] that have been observed in alloys grown at low temperature remains to be established.

Acknowledgment

We wish to thank W.E. Spicer for several informative conversations.

References

[1] A. Sher, M. van Schilfgaarde, A.-B. Chen and W. Chen, Phys. Rev. B36 (1987) 4279.
[2] M.T. Czyzyk, M. Podgorny, A. Balzarotti, P. Letardi, N. Motta, A. Kisiel and M. Zemmal-Slarnavska, Z. Physik B62 (1986) 153.
[3] M. Ichimura and A. Sasaki, J. Appl. Phys. 60 (1986) 3850.
[4] G.P. Srivastava, J.L. Martins and A. Zunger, Phys. Rev. B31 (1985) 2561;
A. Zunger and S.-H. Wei, US Workshop on the Physics and Chemistry of Mercury Cadmium Telluride, New Orleans, LA, 1987;
L.G. Ferreira, A.A. Mbaye and A. Zunger, Phys. Rev. B37 (1988) 10547.
[5] P. Letardi, N. Motta and A. Balzarotti, J. Phys. C (Solid State Phys.) 20 (1987) 2853.
[6] A.A. Mbaye, L.G. Ferreira and A. Zunger, Phys. Rev. Letters 58 (1987) 49.
[7] M. Ichimura and A. Sasaki, Japan. J. Appl. Phys. 26 (1987) 246.
[8] A.-B. Chen and A. Sher, Mater. Res. Soc. Symp. Proc. 46 (1985) 137, Phys. Rev. B32 (1985) 3695.
[9] R.S. Patrick, A.-B. Chen and A. Sher, Phys. Rev. B36 (1987) 6585.
[10] J.C. Mikkelsen, Jr. and J.B. Boyce, Phys. Rev. Letters 49 (1982) 1412.
[11] A.-B. Chen, A. Sher and M.A. Berding, Phys. Rev. B37 (1988) 6285.
[12] G.B. Stringfellow and P.E. Greene, J. Phys. Chem. Solids 30 (1969) 1779.
[13] L.J. Vieland, Acta Met. 11 (1963) 137.
[14] M.A. Berding, A. Sher and A.-B. Chen, J. Vacuum Sci. Technol. 5 (1987) 3009.
[15] M.A. Berding, A. Sher and A.-B. Chen, Vacancy formation energy in semiconductors, in preparation;
M.A. Berding, A. Sher and A.-B. Chen, Vacancy formation energies in semiconductor alloys, in preparation.
[16] W.A. Harrison, Electronic Structure and the Properties of Solids (Freeman, San Francisco, 1980).
[17] M.F.S. Tang and D.A. Stevenson, private communication.
[18] W.L. Brown, Microcircuit Engineering 88, Vienna, September 1988, Paper IP-05.
[19] V. Kumar, D. Kumar and S.K. Joshi, Phys. Rev. B19 (1979) 1954.
[20] J. Bevk, J.P. Mannaerts, L.C. Feldman and B.A. Davidson, Appl. Phys. Letters 49 (1986) 286.
[21] T.S. Kuan, T.F. Kuech, W.I. Wang and E.L. Wilkie, Phys. Rev. Letters 54 (1985) 201.
[22] H.R. Jen, M.J. Cherng and G.B. Stringfellow, in: Proc. 7th Intern. Conf. on Ternary and Multinary Compounds, Snowmass, CO (Mater. Res. Soc., Pittsburgh, PA, 1986).

Electronic structures of HgTe and CdTe surfaces and HgTe/CdTe interfaces

J. T. Schick* and S. M. Bose
Department of Physics and Atmospheric Science, Drexel University, Philadelphia, Pennsylvania 19104

A.-B. Chen
Department of Physics, University of Auburn, Auburn, Alabama 36849
(Received 12 September 1988; revised manuscript received 26 June 1989)

A Green's-function method has been used to study the surface and interface electronic structures of the II-VI compounds HgTe and CdTe. Localized surface and resonance states near the cation-terminated (100) surface of CdTe and the anion-terminated surface of HgTe have been found for the ideal surfaces. The energies and strengths of these surface states are altered by surface perturbations. The bulk states near the surface are drastically modified by the creation of the surface, but the band gaps remain unchanged. Numerical evaluation of the local densities of states at the Γ and J points shows that, at the (100) interface of HgTe/CdTe, the previously observed surface states are no longer present. However, in the interface region, bulk states of one material penetrate some distance into the other material.

I. INTRODUCTION

One of the earliest formulations used in the investigation of surfaces and interfaces was based on the linear combination of atomic orbitals (LCAO) method. Goodwin[1] first applied the LCAO model to study the existence conditions for localized Tamm[2] states in a crystal. He found that these states occur when the diagonal Coulomb integrals and the off-diagonal resonance integrals of the surface are allowed to be different from those of the bulk, for systems of single-level or of sp-hybridized atoms. At the same time, Shockley[3] investigated a periodic potential that is terminated at its maximum and found that, under the condition that the bulk bands crossed, surface states exist in the middle of the band gap. Shockley states, as they have come to be known, come into being when the perturbations at the surface are small in comparison to the widths of the allowed energy bands. Koutecky[4] and others[5] generalized Goodwin's model to study the energy and existence conditions of surface states. Electronic interface states were studied within the LCAO method by Davison and Cheng.[6] An exact tight-binding solution for the surface and interface problems of a one-dimensional semiconductor was obtained by Dy and co-workers.[7,8]

Since these early investigations, many other papers have appeared in the literature for the study of surface and interface properties of solids. Kalkstein and Soven (KS) (Ref. 9) introduced a Green's-function (GF) formalism to study the surface electronic properties of semi-infinite crystals. This is a relatively simple but powerful method by which both the bulk and surface properties of a semi-infinite crystal can be studied. The method can be generalized in a straightforward manner to study the electronic properties of an interface formed by joining two semi-infinite crystals. Because of its simplicity and power the method of KS was widely applied in many calculations during the decade following its development.[10-14] For the same reasons we apply this method in this work to investigate surface and interface electronic structures of the pure II-VI compounds CdTe and HgTe.[15,16]

Despite the intense interest in $Hg_{1-x}Cd_xTe$ over the past three decades due to its application in the preparation of infrared detectors, there have been relatively few experimental and theoretical investigations of the surface and interface properties of this system.[17-19] Since the KS method is well suited for the study of the surface properties of a system described by a tight-binding Hamiltonian, we use here a Hamiltonian closely related to that obtained by Hass, Ehrenreich, and Velicky (HEV) (Ref. 20) for HgTe and CdTe in the empirical tight-binding approximation (ETBA) based on the LCAO interpolation scheme of Slater and Koster.[21] In the ETBA the predicted band structures of the pure compounds HgTe and CdTe are matched to experimentally determined band energies.[22]

In Sec. II, we introduce the formalism of KS for the description of the surface properties of II-VI materials along with the extension of the technique to the problem of interfaces of these materials. Section III contains the results of our calculation and a discussion.

II. FORMALISM

In the formalism of Kalkstein and Soven,[9] a pair of semi-infinite crystals is formed by introducing a cleavage plane into an infinite crystal in one crystallographic direction. A GF describing the electronic properties of the semi-infinite systems is derived from the GF of the infinite crystal and a scattering potential representing the cleavage. When combined with a tight-binding formalism in which nearest- and next-nearest-neighbor interactions are included, the scattering potential is relatively easy to calculate making application to realistic systems simple. The Hamiltonian is constructed, following HEV,

for HgTe and CdTe using sp^3 basis states and including spin.[20-23] The parameters used are those of Slater and Koster[21] as determined by HEV (Ref. 20) except that, in our case, the values of $Esx(110)$ and $Esx(011)$ are interchanged. This produces band structures[15,16] that are qualitatively similar to those of Bryant which offer improved conduction-band mass.[15] We calculate the GF, G associated with this Hamiltonian by using the defining equation

$$G=(E+i\lambda-H)^{-1}, \tag{1}$$

where H is the Hamiltonian of the infinite unperturbed system. It is understood that G is to be calculated in the limit as the positive imaginary part λ approaches zero. These calculations are performed in k space, utilizing fully the periodicity of the crystal.

For the surface (interface) calculations, since we no longer have translational periodicity along the direction perpendicular to the surface (interface), we cannot use an ordinary k-space representation. Kalkstein and Soven[9] assume periodic structure parallel to the surface and use a representation consisting of states which are localized on planes of atomic sites parallel to the surface and represented by the index n, and of Bloch states reflecting the periodic symmetry within the planes which are represented by the index $\mathbf{k}_\parallel$. This is the mixed or Bloch-Wannier representation. We assume the same type of symmetry in the interface system. The Hamiltonian and the GF of the bulk crystal as well as the Hamiltonians of the semi-infinite and interface systems must all be expressed in this mixed representation. The formulas for the semi-infinite system were derived by KS and are the same for us if we reinterpret the algebraic expressions as matrix equations for the sp^3-spin basis states. Note that in this paper we examine the surfaces and interfaces perpendicular to the [100] direction for pure CdTe and HgTe samples. For these compounds, this structure corresponds to an arrangement in which the anions and cations are placed in alternating planes parallel to the surface or interface. For notational simplicity, in the following, a cation-anion pair of planes is given a single layer index n, with the species index ν left to distinguish between the two species (layers).[15,16]

For basis states of the infinite system, KS used states analogous to $|\alpha\sigma\nu\mathbf{k}\rangle$, where $\alpha=s,x,y,z$ indicates the atomic orbital, the spin $(\uparrow,\downarrow)$ is represented by σ, $\nu=a$ or c stands for anions or cations, respectively, and $\mathbf{k}$ is the wave vector. Following KS, we write the Bloch-Wannier states for our system as

$$|\alpha\sigma\nu n\mathbf{k}_\parallel\rangle=\frac{1}{N_\parallel}\sum_{\mathbf{k}_\parallel} e^{-i\mathbf{k}_\parallel\cdot\mathbf{R}_{n\nu}}|\alpha\sigma\nu\mathbf{k}\rangle, \tag{2}$$

where $\mathbf{R}_{n\nu}$ is the position of the ion sublattice represented by ν on the plane labeled by n, and $\mathbf{k}_\parallel$ and $\mathbf{k}_\perp$ are the components of $\mathbf{k}$, parallel and perpendicular to the surface (interface) plane, respectively. This basis set reflects the symmetry of the semi-infinite system and is therefore well suited for our purpose. In the following we suppress the $\mathbf{k}_\parallel$ index for compactness of notation, as in

$$G(n,n')=\sum_{\alpha\sigma\nu}\sum_{\alpha'\sigma'\nu'}|\alpha\sigma\nu n\mathbf{k}_\parallel\rangle\, G\,\langle\alpha'\sigma'\nu'n'\mathbf{k}_\parallel|\times|\alpha\sigma\nu n\mathbf{k}_\parallel\rangle\langle\alpha'\sigma'\nu'n'\mathbf{k}_\parallel|, \tag{3}$$

which is the GF submatrix between layer n and layer n'.

From the site-diagonal GF, $G(n,n)$, we may calculate the local density of states

$$\rho_{n\nu}(E)=-(1/\pi)\,\mathrm{Im}\,\mathrm{Tr}_\nu G(n,n), \tag{4}$$

where Tr_ν indicates that the trace is to be carried out only for the species (layer) given by ν.

The GF of the semi-infinite crystal, G' is found through the application of the Dyson equation,

$$G'=G+GVG'. \tag{5}$$

Besides calculating the matrix elements of the infinite crystal GF, G, we must also find the matrix elements of the scattering potential, V, introduced by the creation of the surface. We label the double layers in the crystal with integers such that the cleavage plane passes between the double layers -1 and 0, and the semi-infinite crystal of interest starts at the zeroth layer and has layer indices $n\geq 0$. Because $G'(n,n')$ is zero for all n and n' less than zero, the only matrix elements of the scattering potential that enter into Eq. (5) in the present model are $V(-1,0)$ and $V(0,-1)$ which express the severing of the interactions across the cleavage plane. We may also include a diagonal term $V(0,0)$, if we wish, to allow for surface perturbations such as relaxation and environmental shifts, into our calculation. These scattering potentials have explicit dependence on $\mathbf{k}_\parallel$ and the pair of indices associated with V refer to the double layers involved in the interaction. In terms of the Hamiltonian matrices, the scattering potential matrix is given by

$$V=H'-H, \tag{6}$$

where H' is the Hamiltonian of the semi-infinite crystal. The scattering potential matrix elements describing the breaking of interactions across the cleavage plane are[9]

$$\begin{aligned}V(0,-1)&=-H(0,-1),\\ V(-1,0)&=-H(-1,0),\end{aligned} \tag{7}$$

where, as before, the $\mathbf{k}_\parallel$ index has been suppressed for compactness of notation. The matrix element $V(0,0)$ is introduced parametrically to account for the shifts in the atomic levels and hopping interaction at the surface layer.

Explicitly, a general matrix of G' is found from Eq. (5) as

$$\begin{aligned}G'(n,n')=G(n,n')&\\ +[G(n,0)V(0,0)&\\ +G(n,-1)V(-1,0)]G'(0,n')&,\end{aligned} \tag{8}$$

with n and $n'\geq 0$. To solve this equation, it is first necessary to solve for $G'(0,n')$ which is done by setting n equal to zero in Eq. (8) and solving for the matrix element

$$G'(0,n')=[1-G(0,0)V(0,0)-G(0,-1)V(-1,0)]^{-1}G(0,n') . \quad (9)$$

To find the general matrix element of G' we need only to substitute Eq. (9) into Eq. (8). To include environmental effects on the surface atoms, we introduce a parameter δ which measures a shift of the surface layer on-site energies. Also, we introduce a parameter d which represents the strength of the interactions between the first two planes of atoms in the semi-infinite crystal relative to the same interactions in the bulk. The parameters δ and d are included in the matrices $V(0,0)$ and $V(-1,-1)$.[15,16]

This model was first extended to the interface between two model semi-infinite one-band crystals by Yaniv.[13] We further extend the technique to real crystals using the sp^3 with spin basis states. The interface is formed between crystal A ($n \le -1$) and crystal B ($n \ge 0$). In joining the crystals, the interactions between the Te and Cd planes, as well as those between the Te planes, across the interface, are taken to be the same as in the bulk. Since there are no data available for the hopping integrals between Hg and Cd planes we assume them to be equal to the average of the interactions between bulk-crystal (100) Cd planes in CdTe and interactions between bulk-crystal (100) Hg planes in HgTe. The interactions between the A and B sides of the interface are included in the scattering potentials $V'(0,-1)$ and $V'(-1,0)$. To our Green's function we add subscripts α and β which take on the values A and B to indicate the side of the interface involved in the calculations. Once again solving the Dyson equation for the interface GF, g (Ref. 16),

$$g = G' + G'V'g$$

we find the interface GF matrix elements as

$$g_{\alpha\beta}(m,n)=G'_{\alpha}(m,n)\delta_{\alpha\beta} + G'_{B}(m,0)V'(0,-1)g_{A\beta}(-1,n)\delta_{\alpha\beta} + G'_{A}(m,-1)V'(-1,0)g_{B\beta}(0,n)\delta_{\alpha A} , \quad (10)$$

where

$$g_{AA}(-1,n)=[1-G'_{A}(-1,-1)V(-1,0)G'_{B}(0,0)\times V(0,-1)]^{-1}G'_{A}(-1,n) , \quad (11a)$$

$$g_{BB}(0,n)=[1-G'_{B}(0,0)V(0,-1)G'_{A}(-1,-1)\times V(-1,0)]^{-1}G'_{B}(0,n) , \quad (11b)$$

$$g_{BA}(0,n)=G'_{B}(0,0)V(0,-1)g_{AA}(-1,n) , \quad (11c)$$

$$g_{AB}(-1,n)=G'_{A}(-1,-1)V(-1,0)g_{BB}(0,n) , \quad (11d)$$

and $\delta_{\alpha\beta}$ is the Kronecker delta. We have assumed a perfect match in the lattice spacing across the interface which is nearly true for HgTe and CdTe with a difference in spacing of only 0.3%.

In the evaluation of the GF of the bulk system a great deal of numerical integration is required. The integrals are evaluated through the use of a Lagrange interpolation scheme in which the time to calculate the band structures is reduced by evaluating them exactly at only a few points in k space and interpolating for the intermediate values.[24] Accuracy is ensured through sampling the functions at a sufficient number of points. All calculations are performed with a small positive imaginary component in the energy and the final results at the real energy axis are evaluated through the use of the analytic continuation procedure of Hass, Velicky, and Ehrenreich.[25]

One quite useful feature of our calculation is that the evaluation of the GF's of the semi-infinite crystal and the interface requires much less computational effort than the evaluation of the infinite crystal GF.[16] As a result of this speed, we are able to consider several values of the environmental shift and surface–nearest-neighbor interaction. In order to observe the effects of the variation of these parameters on the localized states we examine the local densities of states (LDOS) at various values of the parameters for a few CdTe and HgTe surfaces and interfaces. These LDOS are evaluated at fixed values of $\mathbf{k}_{\parallel}$ so that we may find the positions of the localized state bands in the surface Brillouin zone. For details of the evaluation of the matrix elements of the infinite crystal GF and those of the scattering potentials required for our calculation, one is referred to Refs. 15 and 16.

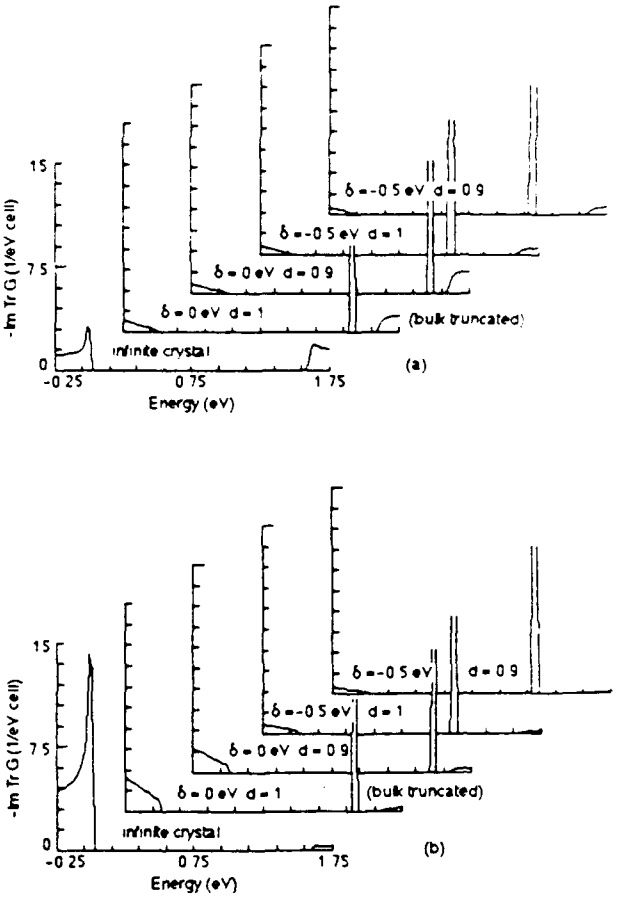


FIG. 1. Dependence of the LDOS on the parameters δ and d as compared to the infinite crystal for (a) the cation surface layer and (b) the first interior anion layer at Γ ($\mathbf{k}_{\parallel}=0$) for CdTe cation-terminated (100) surface.

III. RESULTS AND DISCUSSION

Once the matrix elements of the infinite crystal GF and the relevant scattering potentials are specified, one can calculate the matrix elements of the GF of the semi-infinite sample and that of the sample with the interface using Eqs. (8) and (10), respectively. One can further calculate the LDOS at various layers using an equation of the type of Eq. (4). In this paper we present the LDOS calculated at the symmetry points Γ and J and at energies near the valence- and the conduction-band edges of the sample where the principal band gaps occur.

In Figs. 1 and 2 we plot the surface LDOS at the Γ point for various combinations of the environmental shift parameter δ and the geometric shift parameter (relaxation parameter) d for CdTe cation-terminated and HgTe anion-terminated crystals, respectively. We begin by noting, in Fig. 1, the existence of a localized surface state (bound state) which is the sharp structure that appears in the LDOS within the band gap of the CdTe, whether or not there is a surface perturbation. The ideal cation-truncated (100) surface of CdTe has a surface state at Γ with energy lying at $E = 0.6$ eV above the top of the valence band. We also see that the bulk densities of states are modified as a result of the introduction of a surface but the band gap of CdTe remains unchanged. We note that the position of the bound state is sensitive to

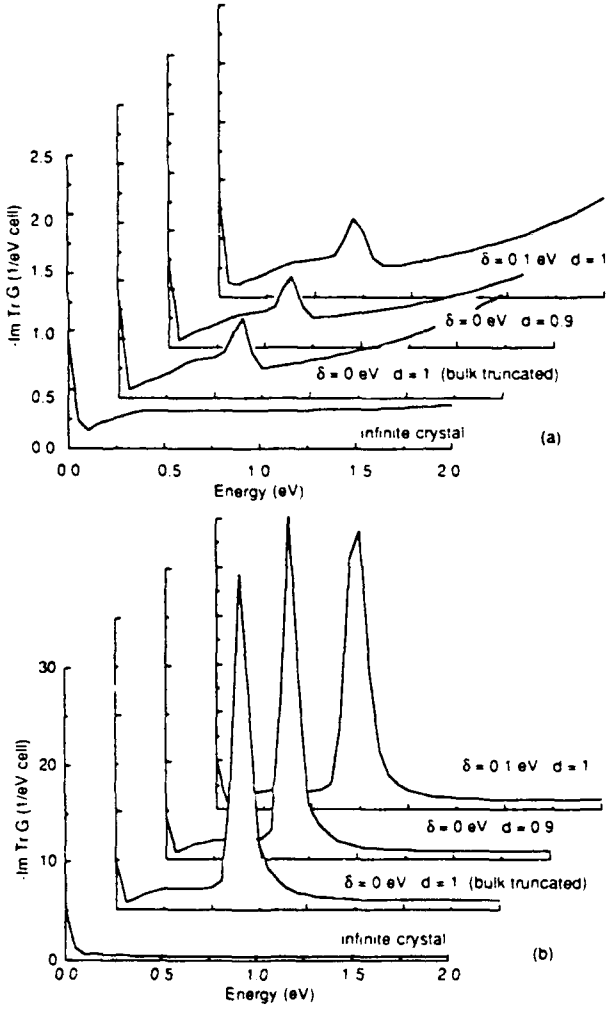


FIG. 2. Dependence of the LDOS on the parameters δ and d as compared to the infinite crystal for (a) the anion surface layer and (b) the first interior cation layer at Γ ($\mathbf{k} = 0$) for HgTe cation-terminated (100) surface.

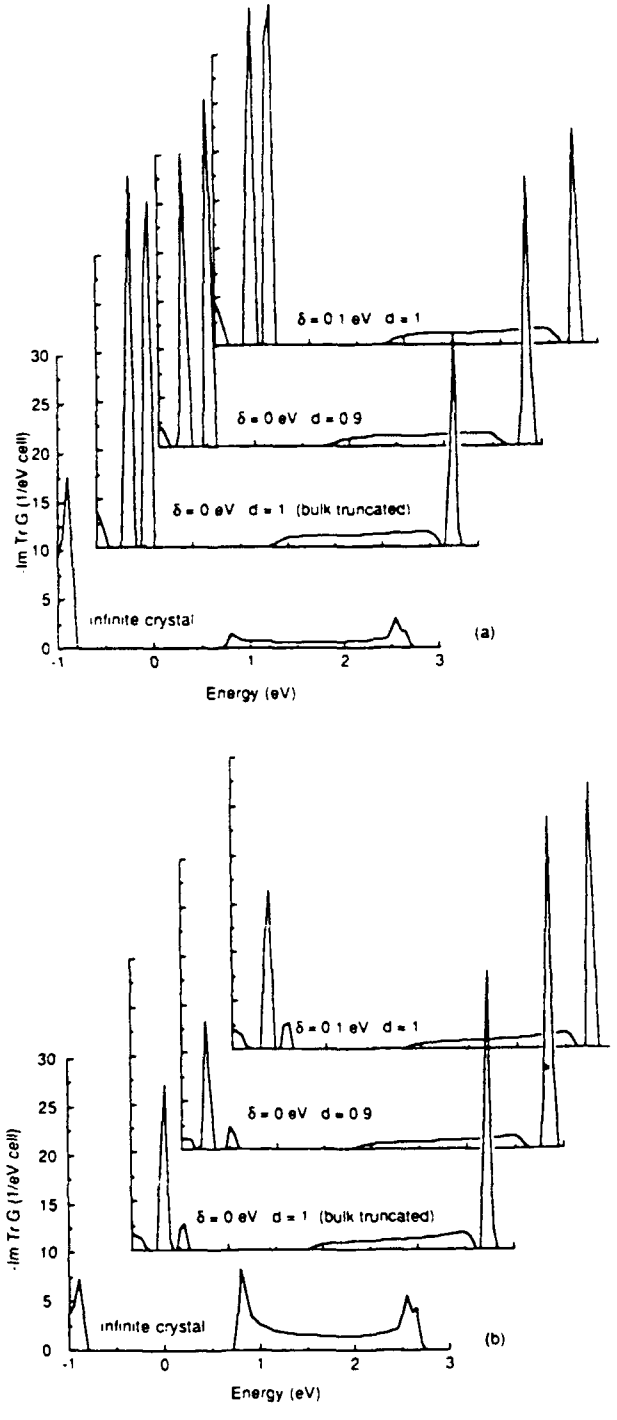


FIG. 3. Dependence of the LDOS on the parameters δ and d as compared to the infinite crystal for (a) the anion surface layer and (b) the first interior cation layer at J [$\mathbf{k} = \pi/a(\sqrt{2}, \sqrt{2})$] for HgTe anion-terminated (100) surface.

variations in both d and δ. In Fig. 2, we notice that the surface has again introduced a marked change in the LDOS of HgTe. A resonance state appears in the form of a spike inside the conduction band of semimetal HgTe. The contribution of this resonance to the LDOS is strongest for anion layers indicating that anion states are the most likely constituents of this state. Similar surface states have been calculated by Bryant[18] for the special case where no surface perturbations exist.

The surface states calculated for the HgTe at the J point where there exists a gap are shown in Fig. 3. These states exhibit the same dependence on the interactions that was seen at the Γ point. Here instead of a single peak we see a pair of bound states just above the valence band within the band gap. Also above the conduction band there is a bound state that was not observed for the Γ point. Anions contribute more strongly to the bound states inside the band gap while the cation contribution is stronger to the bound state above the conduction band. Also notice that the bulk states are strongly modified by the introduction of the surface and that the band gaps are not altered by the surface.

Our calculation of the layer dependence of the surface states shows that these states become progressively less significant as we examine deeper inside the crystal indicating these states are indeed localized bound states. The bulk LDOS in the deeper layers, on the other hand, start resembling the infinite crystal LDOS.

As mentioned before, the interface calculations have been performed assuming no lattice mismatch in the crystal structures of the HgTe and CdTe compounds, thus neglecting all strains that may be present at the interface. The parameters chosen for our calculation also do not allow for valence-band offset. With these assumptions, dramatic effects are seen at the interface of HgTe and CdTe as shown in Figs. 4 and 5, where the LDOS in the CdTe side of the interface are presented at Γ and J points, respectively. The most obvious effect is the lack of localized states that were previously seen at the free surfaces of these materials. Instead, we find that, close to the interface, the bulk states of HgTe that lie in the region between 0 and 1.6 eV appear in the band gap of

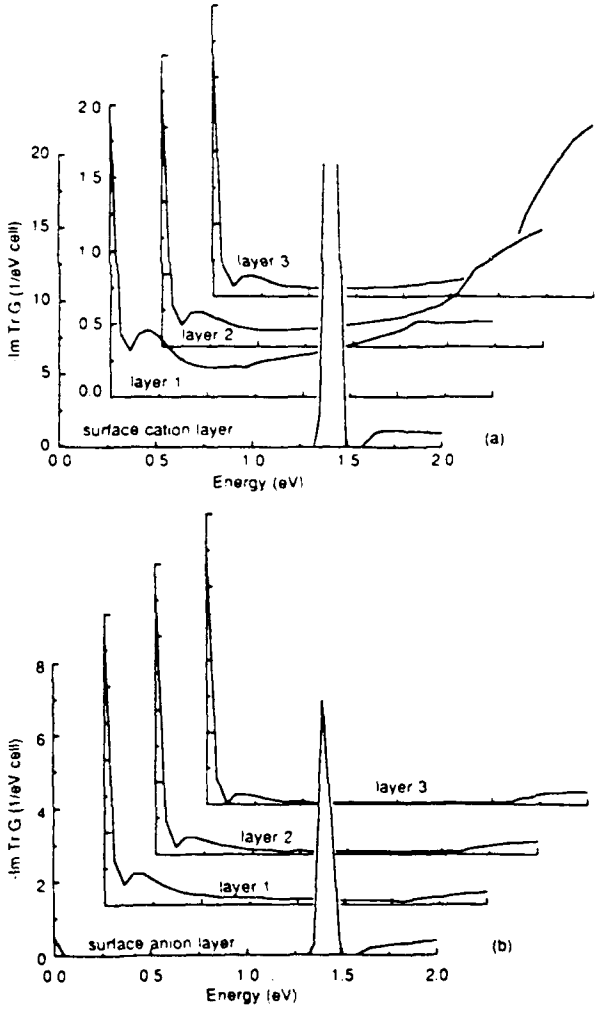


FIG. 4. Comparison of the surface LDOS of the cation-terminated (100) surface of CdTe with the LDOS plotted (a) at successive cation layers and (b) anion layers for the CdTe side of the (100) interface CdTe/HgTe at Γ.

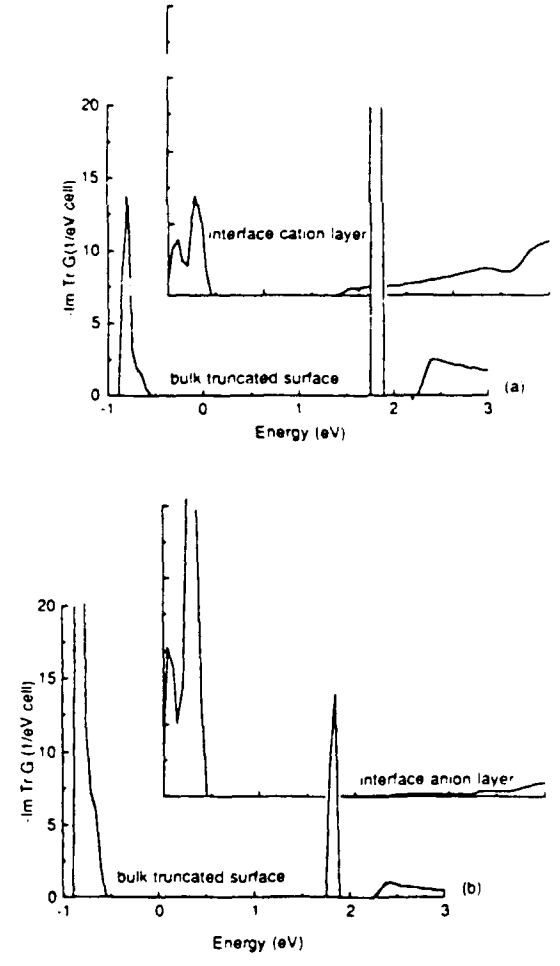


FIG. 5. Comparison of the surface LDOS of the cation-terminated (100) surface of CdTe with the LDOS plotted (a) at first cation layer at the interface and (b) the first interior anion layer for the CdTe side of the (100) interface CdTe/HgTe at J.

CdTe at the Γ point. These states in the CdTe gap become less significant at layers further from the interface while at the same time the LDOS in the energy ranges of the conduction and valence bands of CdTe increase in strength. Thus we have a narrowing of the gap of CdTe in a limited region near the interface which allows conduction electrons to penetrate into the CdTe from the HgTe over a short distance. In Fig. 5, we explicitly see how the gap at the CdTe layers is reduced with respect to pure CdTe at J. Calculation of the LDOS in the HgTe side of the interface shows similar effects.

The above results describe how the LDOS change from the bulk to the surface and then from the surface to the interface. While the bulk and the interface results can be considered realistic, the surface results may not be, since the surface reconstruction has not been included in the calculation. Recent experiments[17,26–28] and a structural theory[29] indicate that the surfaces of CdTe and some other II-VI compounds undergo reconstructions similar to those on the surfaces of III-V compounds. However, similar measurements are yet to be made on HgTe surfaces and CdTe/HgTe interfaces. Our calculations indicate that changes in the electronic structure in the form of environmental shifts at the surface lead to only minor changes in the localized surface and the bulk LDOS. Whether a surface reconstruction will result in substantial modification of these states is yet to be determined.

In conclusion, we have described in this paper how one can incorporate the GF method of KS to calculate the surface and interface structures of II-VI compounds without having to deal with the artificial super-cell approximation. Our calculation shows that drastic changes in the LDOS can occur because of creation of surfaces and interfaces. Localized surface or resonance states appear above the top of the valence band, and the effects of the surface on the LDOS persist at several layers inside the bulk. At the interface, there are no localized states but the bulklike states related to one material penetrate into several layers inside the other material resulting in narrowing of band gaps close to the interface for the wide band-gap material. At present there are no systematic experimental data available for comparison with our conclusions.

Finally, even though we have not included all aspects necessary for a complete description of the surfaces and interfaces, our work can be considered to be the first step toward the understanding of the surface and interface electronic structures of the II-VI compounds. To our knowledge, the results presented here are new for the interface and more general than any previous surface calculations on the II-VI compounds. Since the method is flexible, it will be possible to incorporate realistic structural models dealing with the surface reconstructions, when they become available, to calculate more realistic electronic properties of such surfaces and interfaces.

ACKNOWLEDGMENTS

This work was supported by National Aeronautics and Space Administration Grant No. NAG 1708.

*Present address: Department of Physics and Astronomy, Wayne State University, Detroit, MI 48202.

[1]E. T. Goodwin, Proc. Cambridge Philos. Soc. **35**, 232 (1939).
[2]I. Tamm, Z. Phys. **76**, 849 (1932); Phys. Z. Sowjet. **1**, 733 (1932).
[3]W. Shockley, Phys. Rev. **56**, 317 (1939).
[4]J. Koutecky, Phys. Rev. **108**, 13 (1957).
[5]J. Alstrup, Surf. Sci. **20**, 325 (1970).
[6]S. G. Davison and Y. C. Cheng, Int. J. Quantum Chem. Symp. **2**, 303 (1968).
[7]K. S. Dy, S.-Y. Wu, and T. L. Spratlin, Phys. Rev. B **20**, 4237 (1979).
[8]J. D. Brasher and K. S. Dy, Phys. Rev. B **22**, 4868 (1980).
[9]D. Kalkstein and P. Soven, Surf. Sci. **26**, 85 (1971).
[10]E. N. Foo and H. S. Wong, Phys. Rev. B **9**, 1857 (1974); **10**, 4819 (1974).
[11]S. M. Bose and E. N. Foo, Phys. Rev. B **10**, 3534 (1974).
[12]S. M. Bose and J. M. Yuan, Phys. Rev. B **24**, 1934 (1981).
[13]A. Yaniv, Phys. Rev. B **17**, 3904 (1979); **22**, 4776 (1980).
[14]N. F. Berk, Surf. Sci. **48**, 289 (1975).
[15]J. T. Schick and S. M. Bose, Prog. Surf. Sci. **25**, 107 (1987).
[16]J. T. Schick, Ph. D. thesis, Drexel University, 1987 (unpublished).
[17]W. E. Spicer, J. A. Silverman, P. Morgan, I. Lindau, and J. A. Wilson, J. Vac. Sci. Technol. **21**, 149 (1982).
[18]G. W. Bryant, Phys. Rev. Lett. **55**, 1786 (1985); Phys. Rev. B **35**, 5547 (1987); **36**, 6465 (1987).
[19]J. N. Schulman and T. C. McGill, Appl. Phys. Lett. **34**, 663 (1979).
[20]K. C. Hass, E. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 1088 (1983).
[21]J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[22]D. J. Chadi, J. P. Walter, M. L. Cohen, V. Petroff, and M. Balkanski, Phys. Rev. B **5**, 3058 (1972); A. Moritani, K. Taniguchi, C. Hamaguchi, and J. Nakai, J. Phys. Soc. Jpn. **34**, 79 (1973); N. J. Schevchik, J. Tejeda, M. Cardona, and D. W. Langer, Phys. Status Solidi B **59**, 87 (1973); L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalczyk, and D. A. Shirley, Phys. Rev. B **9**, 600 (1974).
[23]D. J. Chadi, Phys. Rev. B **16**, 790 (1977).
[24]A.-B. Chen, Phys. Rev. B **16**, 3291 (1977).
[25]K. C. Hass, B. Velicky, and H. Ehrenreich, Phys. Rev. B **29**, 3697 (1984).
[26]P. John, T. Miller, T. C. Hsieh, A. P. Shapiro, A. L. Wachs, and T.-C. Chiang, Phys. Rev. B **34**, 6706 (1986).
[27]K. O. Magnusson, S. A. Foldstrom, and P. E. S. Person, Phys. Rev. B **38**, 5384 (1988).
[28]P. Lu and D. J. Smith, Phys. Rev. Lett. **59**, 2177 (1987).
[29]Y. R. Wang and C. B. Duke, Phys. Rev. B **36**, 2763 (1987).

Surface segregation in pseudobinary alloys

R. S. Patrick and A.- B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849

A. Sher and M. A. Berding
Physical Electronics Laboratory, SRI International, Menlo Park, California 94025
(Received 12 September 1988)

Alloys of the form $A_xB_{1-x}C$ almost always have a different surface concentration from the bulk in order to maintain a constant chemical potential for each layer in the alloy. We have calculated the degree of surface segregation for the pseudobinary alloys $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$. The enthalpy responsible for segregation is the difference in the energies for moving an A or B atom from the bulk alloy to the surface. There are two major contributions to this energy process: (1) a bond-breaking mechanism, whereby the element with the lowest surface energy segregates to the top, and (2) strain release, where the dilute element in the compound segregates to the surface to alleviate the strain energy due to mismatch of the AC and BC bond lengths. In our segregation model, the free energy of each layer is calculated in the regular and quasichemical approximations. By equating the chemical potentials of each successive layer to the bulk, the composition of each layer is obtained. Our results indicate that there is strong surface enrichment of Hg in $Hg_{1-x}Cd_xTe$ while there is less surface segregation of Hg in $Hg_{1-x}Zn_xTe$ at the low x values appropriate for infrared application. Mercury segregation to the surface will lower the band gap or may turn the surface into a semimetal, thereby affecting the passivation of the surface.

I. INTRODUCTION

Applying such probe techniques as Auger electron spectroscopy (AES), low-energy electron diffraction (LEED), and ion-scattering spectroscopy (ISS) allows one to determine the chemical compositions of surfaces with great accuracy. These measurements[1] have shown that the chemical composition of the surface is different from the bulk composition in most metal alloys. This phenomenon of surface segregation has also been studied theoretically in great detail starting with the work of Gibbs,[2] who noted that the surface composition will be enriched with the alloy component that minimizes the free energy associated with the creation of the surface (i.e., bond breaking at the surface). From Gibbs's adsorption equation for a two component system (A_xB_{1-x}) it has been shown[3] that the surface concentration of A atoms (x_S) differs from that of the bulk (x_b) by the relation

$$\frac{x_S}{1-x_S}=\frac{x_b}{1-x_b}\exp[(\gamma_B-\gamma_A)\sigma/kT]\ , \tag{1}$$

where $\gamma_{A(B)}$ is the surface tension of pure A (B) and σ is the average area per atom.

Application of Eq. (1) to real systems does not provide good quantitative results since it does not consider any deviations from the bulk concentration for layers beneath the surface. Multilayer models which include layer compositions under the surface were formulated by Williams and Nason[4] among others[5] by employing a regular solution method. Later work by Kumar *et al.*[6,7] extended the surface segregation problem to take into account nonrandom distribution of atoms in the alloy using the quasichemical approximation (QCA).[8] Moran-Lopez and Falicov[9] have also looked at the segregation problem along with surface order and/or disorder transitions using Kikuchi's method.[10]

The primary driving force for surface segregation for an alloy A_xB_{1-x} can be thought of as the difference in energy on moving an A or B atom to the surface. As noted above, one contribution is the bond-breaking mechanism, whereby the element with the lowest surface free energy segregates to the top. Experiments[1] have shown that this is not the only mechanism that causes one species to prefer to be on the surface. In dilute alloys where there is a size difference in the component species, the dilute element in the compound will segregate to the top to alleviate the strain energy due to lattice mismatch. An example of this is seen in CuAu, where in the dilute Au alloy case, Au is predominantly on the surface.[1] Looking only at the bond-breaking mechanism, however, Cu is predicted to be the segregating element since it has the lower surface tension. It can be explained by the strain release argument since the Au atom is approximately 15% larger than the Cu atom. The connection between strain release and surface segregation was first pointed out by McLean[11] in his study on grain boundaries. Incorporation of both strain energy and bond breaking in surface segregation studies has been done by Wynblatt and Ku.[12,13]

In the theoretical and experimental work done so far, the materials studied have primarily been transition and noble metals. The effect of surface segregation on metals has important implications when dealing with catalysis, corrosion, and grain boundaries. In this paper, however,

we will look at the effects of surface segregation on semiconductor alloys. In particular, we are interested in the II-VI systems $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$. The suitability of these materials for infrared detectors is determined to a large extent by the control of the surface composition and the capability for passivation. Because of the structural weakness of $Hg_{1-x}Cd_xTe$, it has been suggested that $Hg_{1-x}Zn_xTe$ may be better suited for infrared detectors. We are interested in how the structural and electronic properties are modified by the presence of surfaces and interfaces. In regards to surface segregation, we are interested in examining the extent of Hg migration to the crystal surface.

In the first part of this paper we will discuss models that describe the thermodynamics and amount of surface segregation that occur in pseudobinary binary alloys of the form $A_xB_{1-x}C$. These models are modifications of the regular solution by Williams and Nason[4] and the QCA theory by Kumar[6,7] to include first- and second-nearest neighbors in zinc-blende structures. We note that the shortest range of interaction that can effect the atomic distribution in a bulk pseudobinary semiconductor alloy is the second-neighbor interaction (e.g., A-A, A-B, and B-B), because this is the closest distance between the alloying atoms. Thus, in the QCA, the second-neighbor model for a pseudobinary alloy is equivalent to a first-neighbor model for the usual metal binary alloys. While the first-neighbor interactions (A-C and B-C) in a pseudobinary alloy have no effect on the bulk statistics, their contributions are required for the study of surface segregation. These first-neighbor terms will become part of the "bond-breaking" energies which may become the dominant factor for surface segregation of an alloy component. Our formalism will also incorporate the release of the strain energy when an atom positions itself on the surface. Section III presents numerical results for the layer compositions of the II-VI alloys $Hg_{1-x}Zn_xTe$ and $Hg_{1-x}Cd_xTe$. We will adopt the experimental and theoretical mixing enthalpies quoted in Ref. 14 for the second-neighbor energy parameters. The energies required to move a cation from the alloy bulk and place it on the surface, calculated for these alloys in Ref. 15, will be used as the bond-breaking energies. All temperatures considered will be above the critical temperature for phase separation. In addition, we will only consider substitutional segregation. In Sec. IV we will discuss the results and give ideas for further work on this topic.

II. THEORY

Consider a zinc-blende alloy of the type $A_xB_{1-x}C$, where the A and B atoms occupy one fcc lattice while the C atoms occupy the other fcc lattice. In this configuration, an A (or B) atom is surrounded by four C nearest-neighbor atoms and a total of 12 second-nearest neighbor A and B atoms. Assuming the alloy is semi-infinite, divide the alloy into layers parallel to the surface, where each layer i has a fractional concentration of A atoms, x_i ($i=1$ being the surface layer). The layers considered here refer to layers which contain the alloying atoms (A and B). Note that for the zinc-blende structure, a layer of C atoms is always between adjacent alloy atom layers. Neglect any gas vapor effects by assuming a vacuum interface with the surface.

To determine the composition distribution at thermal equilibrium the total free energy of the system F is minimized with respect to the layer concentrations $\{x_i\}$ with the constraint that the total number of A atoms is a constant:

$$\sum_i x_i = \text{const} . \tag{2}$$

For layers deep inside the alloy, any effects of the surface are negligible and the concentration is that of the bulk, x_b. Through the use of Lagrange multipliers, the conditions stated above can be written as

$$\frac{\partial F}{\partial x_i} = \frac{\partial F}{\partial x_b} \tag{3}$$

for all x_i. This is equivalent to saying that the chemical potentials of each layer, defined as $\mu_i = \partial F/\partial x_i$, must be equal to each other:

$$\mu(x_i) = \mu(x_b) . \tag{4}$$

We will write the total free energy per number of atoms in a plane for the layers system as a sum of layer contributions:

$$F = \sum_i U_i - T \sum_i S_i , \tag{5}$$

where U_i is the configurational enthalpy and S_i is the configurational entropy of the ith layer. As will be seen later, U_i and S_i actually depend on contributions from neighboring layers.

In the regular or QCA treatments, the interactions in the enthalpy term arise from pair interactions only, the interactions in our case being A-C and B-C for the nearest neighbor and A-A, A-B, and B-B for the second nearest neighbor. These bond enthalpies are written as ϵ^s_{ij} and ϵ_{ij} for the i-j pair interactions at the surface and below the surface, respectively. As we shall see later, the distinction between surface and nonsurface bond enthalpies is done in order to consider the effects of strain release.

In a given layer, an A or B atom is surrounded by

$$\begin{aligned} Z^{(1)} &= Z_L^{(1)} + Z_B^{(1)} + Z_T^{(1)} , \\ Z^{(1)} &= Z_L^{(1)} + Z_B^{(1)} \ (\text{surface}) , \end{aligned} \tag{6}$$

nearest-neighbor C atoms. In (6), $Z_L^{(1)}$ is the number of nearest neighbors on the same layer and $Z_T^{(1)}$ and $Z_B^{(1)}$ are the number of nearest neighbors on the top and bottom layers, respectively. Similarly, for the second neighbor A-A, A-B, and B-B interactions, the number of second-nearest neighbors at and below the surface will be expressed as

$$\begin{aligned} Z &= Z_L + Z_B + Z_T , \\ Z &= Z_L + Z_B \ (\text{surface}) . \end{aligned} \tag{7}$$

Regular solution. In the context of the regular solution, the distribution of atoms in a given layer is assumed to be random and the entropy term per layer is simply the form for random distribution of atoms on a given layer

$$S_i = -k\,[x_i \ln x_i + (1-x_i)\ln(1-x_i)]\,. \tag{8}$$

Tne enthalpy for the first layer is given by

$$U_1 = Z_L^{(1)}[x_1\varepsilon^s_{AC} + (1-x_1)\varepsilon^s_{BC}] + \frac{Z_B^{(1)}}{2}[x_1\varepsilon_{AC} + (1-x_1)\varepsilon_{BC}] + \frac{Z_L}{2}[x_1^2\varepsilon^s_{AA} + 2x_1(1-x_1)\varepsilon^s_{AB} + (1-x_1)^2\varepsilon^s_{BB}]\,,$$

$$+\frac{Z_B}{2}[x_1x_2\varepsilon_{AA} + x_1(1-x_2)\varepsilon_{AB} + x_2(1-x_1)\varepsilon_{AB} + (1-x_1)(1-x_2)\varepsilon_{BB}]\,, \tag{9}$$

while for all other layers the enthalpy is

$$U_i = \left[Z_L^{(1)} + \frac{Z_B^{(1)} + Z_T^{(1)}}{2}\right][x_i\varepsilon_{AC} + (1-x_i)\varepsilon_{BC}]$$

$$+\sum_{j=i-1}^{j=i+1}\frac{Z_{ij}}{2}[x_ix_j\varepsilon_{AA} + x_j(1-x_i)\varepsilon_{AB} + x_i(1-x_j)\varepsilon_{AB} + (1-x_j)(1-x_j)\varepsilon_{BB}]\,, \tag{10}$$

where

$$Z_{ij} = \begin{cases} Z_T, & j = i-1 \\ Z_L, & j = i \\ Z_B, & j = i+1 \end{cases} \tag{11}$$

and the $\frac{1}{2}$ terms are there to prevent double counting. For those layers in the bulk, the free energy, F_b, is obtained from Eqs. (10) and (11) by setting $x_{i+1} = x_i = x_{i-1} = x_b$,

$$F_b = \left[Z_L^{(1)} + \frac{Z_B^{(1)} + Z_T^{(1)}}{2}\right][x_b\varepsilon_{AC} + (1-x_b)\varepsilon_{BC}]$$

$$+\frac{Z}{2}[x_b^2\varepsilon_{AA} + 2x_b(1-x_b)\varepsilon_{AB} + (1-x_b)^2\varepsilon_{BB}]$$

$$+kT[x_b\ln x_b + (1-x_b)\ln(1-x_b)]\,. \tag{12}$$

Now that we have the total free energy of our semi-infinite alloy we can use our condition for thermal equilibrium [Eq. (3)] to relate the layer concentrations to the bulk layers as

$$\frac{x_i}{1-x_i} = \frac{x_b}{1-x_b}e^{\Delta F_i^{seg}/kT}\,, \tag{13}$$

where the enthalpy of segregation ΔF_i^{seg} is given by

$$\Delta F_1^{seg} = \Delta E_A - \Delta E_B - Z_L\varepsilon^s(1-2x_1) + 2\varepsilon\left[\frac{Z+Z_L}{4} + \left[\frac{Z_T+Z_B}{2}\right]x_2 - Zx_b\right]\,, \tag{14a}$$

$$\Delta F_i^{seg} = 2\varepsilon\left[Z_Lx_i\left[\frac{Z_T+Z_B}{2}\right](x_{i+1}+x_{i-1}) - Zx_b\right] \quad (i\neq 1)\,. \tag{14b}$$

In Eq. (14), the interaction parameter is defined as

$$\varepsilon = \varepsilon_{AB} - \tfrac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB})\,, \tag{15}$$

with ε^s defined similarly. The first term in Eq. (14a) is given by

$$\Delta E_A = \left[Z_L^{(1)} + \frac{Z_T^{(1)}}{2}\right]\varepsilon_{AC} - Z_L^{(1)}\varepsilon^s_{AC} - \frac{Z_L}{2}\varepsilon^s_{AA} + \frac{\varepsilon_{AA}}{2}\left[Z_L + \frac{Z_T+Z_B}{2}\right] \tag{16}$$

with a similar term for ΔE_B. Note that the first-neighbor interactions only enter the surface term, ΔF_1^{seg}. Also note that the terms $\{x_i\}$ before Eq. (13) are an arbitrary set of layer concentrations while in Eqs. (13) and (14) and hereafter $\{x_i\}$ are to be defined as the equilibrium layer concentrations.

In the ideal solution case (i.e., $\varepsilon = 0$) only the surface layer is different from the bulk and the degree of surface segregation is determined entirely by the term $\Delta E_A - \Delta E_B$. Comparing Eqs. (1), (13), and (14a) we see that $\Delta E_A - \Delta E_B$ is related to the surface tension.

QCA solution. For a nonzero interaction parameter [Eq. (15)], *A-A* and *B-B* bonds are favored when ε is positive while *A-B* bonds are favored when ε is negative. This contradicts the entropy term in the regular solution [Eq. (8)], which was formulated with the assumption of a random distribution of the atoms in the alloy. The QCA formalism takes into account this nonrandomness via short-range disorder parameters. In the standard QCA for bulk alloys, only one short-range disorder parameter is required. As pointed out by Kumar,[6] however, when one is considering different layer concentrations the distinction must be made between intralayer and interlayer short-range disorder.

For the QCA, the segregation equation becomes

$$\frac{x_i}{1-x_i} = \frac{x_b}{1-x_b}e^{\Delta F_1^{seg}/kT}e^{-\Delta S_i^{seg}/k}\,, \tag{17}$$

where ΔS_i^{seg} is the entropy of segregation. Our QCA result for the enthalpy of segregation is basically the same as that for Kumar[7] except that only the lateral interactions on the surface contain the ε^s term

$$\Delta F_1^{seg} = \Delta E_A - \Delta E_B - Z_L\varepsilon^s(1-2x_1)\alpha_1 + 2\varepsilon\left[\frac{Z\alpha_b}{2} + \left[\frac{Z_T+Z_B}{4}\right][1-2(1-x_2)\beta_1] - Zx_b\alpha_b\right]\,, \tag{18a}$$

$$\Delta F_i^{\rm seg} = 2\varepsilon \left\{ \frac{Z\alpha_b - Z_L\alpha_i - (Z_T + Z_B)\beta_i}{2} + \left| \frac{Z_T + Z_B}{2} \right| (x_{i-1}\beta_i + x_{i-1}\beta_{i-1}) + Z_L x_i\alpha_i - Zx_b\alpha_b \right| , \quad (18b)$$

where α_i and β_i are intralayer and interlayer short-range disorder parameters. These short-range disorder parameters give a measure of the degree of randomness in the alloy. As $\alpha_i \to 1$ and $\beta_i \to 1$, the alloy becomes a random alloy and the enthalpy of segregation term reverts back to Eq. (14). These parameters can be obtained from the mass-action laws,[6,7]

$$\frac{\varepsilon}{kT} + \frac{1}{2}\ln\left| \frac{p^{ii}_{AB}p^{ii}_{BA}}{p^{ii}_{AA}p^{ii}_{BB}} \right| = 0 , \quad (19a)$$

$$\frac{\varepsilon}{kT} + \frac{1}{2}\ln\left| \frac{p^{i,i-1}_{AB}p^{i,i-1}_{BA}}{p^{i,i-1}_{AA}p^{i,i-1}_{BB}} \right| = 0 , \quad (19b)$$

where p^{ij}_{XY} is the probability that an X atoms is on the ith layer and a Y atom is in an adjacent position on the jth layer. For the surface layer ($i=1$) ε is replaced by ε^s in (19). The probabilities are summarized below:

$$\begin{aligned}
p^{ii}_{AA} &= x_i[1-(1-x_i)\alpha_i] , \\
p^{i,i-1}_{AA} &= x_i[1-(1-x_{i-1})\beta_i] , \\
p^{ii}_{BB} &= (1-x_i)(1-x_i\alpha_i) , \\
p^{i,i-1}_{BB} &= (1-x_i)(1-x_{i-1}\beta_i) , \\
p^{ii}_{AB} &= x_i(1-x_i)\alpha_i , \\
p^{i,i-1}_{AB} &= x_i(1-x_{i-1})\beta_i , \\
p^{ii}_{BA} &= x_i(1-x_i)\alpha_i , \\
p^{i,i-1}_{AB} &= x_{i-1}(1-x_i)\beta_i .
\end{aligned} \quad (20)$$

Note that additional interlayer interaction short-range parameter β'_i. Since concentration varies with layer, two interlayer short-range disorder parameters are required. However, the β'_i is related to β_i from the constraint

$$N^{i,i-1}_{AA} + N^{i,i-1}_{BA} = Z_T N^{i-1}_A , \quad (21)$$

where N^{ij}_{XY} is the number of X-Y pairs when X is on the ith layer and Y is on the jth layer and N^i_A is the number of A atoms on the ith layer. See Refs. 6 and 7 for complete details.

The entropy of segregation term is the same as that derived in Ref. 7 and is given by

$$\Delta S_i^{\rm seg}/k = Z\ln\left| \frac{x_b}{1-x_b} \right| - Z^i\ln\left| \frac{x_i}{1-x_i} \right| - \frac{Z}{2}h^b_{AB} + \frac{Z_L}{2}h^i_{AB} + \left| \frac{Z_T + Z_B}{2} \right| \left| \ln\left| \frac{p^{i,i-1}_{AA}}{p^{i,i-1}_{BA}} \right| - \ln\left| \frac{p^{i,i-1}_{BB}}{p^{i,i-1}_{BA}} \right| + \beta_i(1-x_{i+1})g^{ii}_{AB} - \beta_{i-1}x_{x-1}g^{i,i-1}_{AB} \right| , \quad (22)$$

where Z^i is the total number of second-nearest neighbors at layer i and h^i_{AB} and g^{ij}_{AB} are defined as

$$h^i_{AB} = \ln\left| \frac{p^{ii}_{AA}}{p^{ii}_{BB}} \right| + \alpha_i(1-2x_i)g^{ii}_{AB} , \quad (23)$$

$$g^{ij}_{AB} = \ln\left| \frac{p^{ij}_{AB}p^{ij}_{BA}}{p^{ij}_{AA}p^{ij}_{BB}} \right| . \quad (24)$$

As the short-range disorder parameters approach 1 and the entropy of segregation term goes to zero.

III. RESULTS

Numerical results are now presented for the (111) face that terminates with the cation surface of $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$ for temperatures are the critical temperature for phase separation. For the (111) face, the second-nearest-neighbor coordination numbers are $Z=12$, $Z_L=6$, and $Z_T=Z_B=3$. Since accurate values for the surface tension of a solid are difficult to obtain, workers in metal alloys have related this term to the heat of vaporization ($\Delta H^{\rm vap}$) via[3]

$$\gamma_i\sigma = \frac{Z_T}{Z}\Delta H^{\rm vap} . \quad (25)$$

As Wynblatt and Ku[12,13] have pointed out, however, this relation overestimates the surface energy. [For the (100) fcc face, the overestimation is a factor of 2.] The empirical relationship

$$\gamma_i\sigma = \Delta H^{\rm vap}/6 \quad (26)$$

has given better numerical results for liquid metals.[3] In our study of semiconductor alloys, however, we have used the vacancy formation energies, calculated by Berding *et al.*,[15] for $\Delta E_{A(B)}$. This vacancy formation energy is defined as the energy required to take an atom from the bulk and place it on a particular surface. These quantities have been calculated using Harrison's tight-binding theory[16] with corrections made for the long-range Coulomb energies resulting from charge redistribution and rehybridization of the dangling bonds. Unlike other authors who have used concentration-independent and/or empirical parameters, the vacancy formation energy from Ref. 15 is concentration dependent. The values for $Hg_{1-x}Zn_xTe$ are given by

$$\begin{aligned}
\Delta E_{\rm Hg} &= -(0.23x_b + 0.89)\ {\rm eV} , \\
\Delta E_{\rm Zn} &= -(0.26x_b + 0.98)\ {\rm eV} ,
\end{aligned} \quad (27)$$

and for $Hg_{1-x}Cd_xTe$

$$\Delta E_{Hg} = -(0.14x_b + 0.89)\ \text{eV}\ ,$$
$$\Delta E_{Cd} = -(0.14x_b + 1.02)\ \text{eV}\ . \tag{28}$$

The pair interaction terms for the bulk (ε) and surface (ε^s) are provided by Chen and Sher,[14] who determined these parameters from substitutional energies when an A or B atom is removed from the bulk. These substitutional energies are calculated using Harrison's bonding theory[16] along with the valence-force-field model.[17] Effects incorporated in the substitutional energies include strain and chemical terms that arise as the alloy relaxes upon removal of an atom. For our surface interaction parameter, however, the strain term in the substitutional energy is set to zero to allow for strain release. Our approach to adding strain effects differs from that of Wynblatt and Ku.[12,13] Their enthalpy of segregation was obtained by taking Eq. (14) and adding a term derived by Friedel[18] who used elastic continuum theory to calculate elastic strain energy in a dilute alloy. They did not differentiate between surface and bulk interaction parameters and considered only the top surface as being different from the bulk.

With these parameters, numerical layer compositions were calculated using Eqs. (13) and (14) for the regular solution and Eqs. (17) and (18) for the QCA solution. The assumption was made that a certain layer N is the bulk layer and those layers above the Nth layer ($i < N$) will have different layer compositions than the bulk. A total of N nonlinear coupled equations were then solved for the layer compositions. In order to make sure enough layers were taken, the calculation was repeated for $N = N + 1$ until convergence close to the bulk value was reached. For most systems and temperatures presented in this report, four layers was enough. For most of our

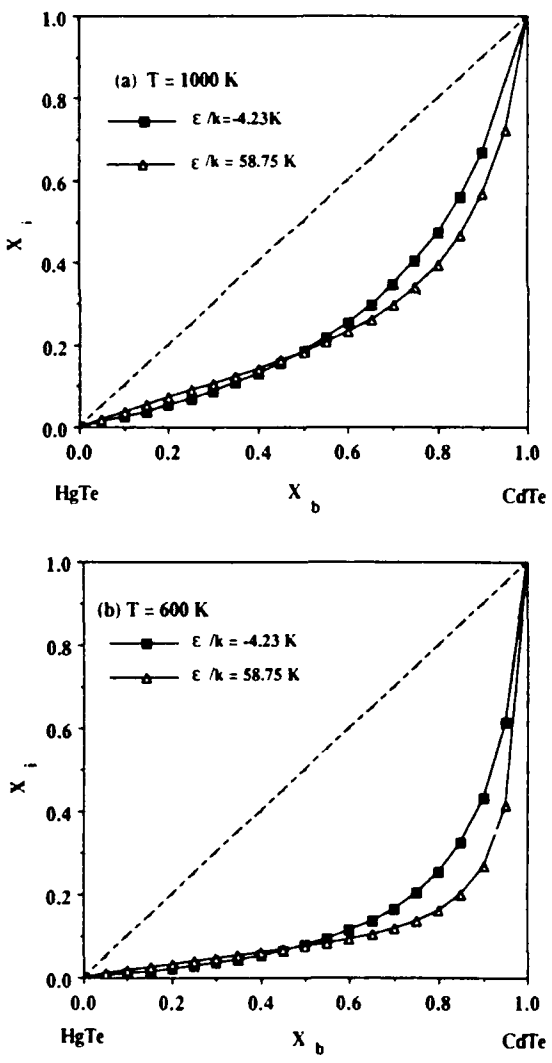


FIG. 1. Fractional surface concentration (x_i) of CdTe as a function of the fractional bulk CdTe concentration (x_b) at the (111) face of $Hg_{1-x}Cd_xTe$. Results are from the quasichemical approximation using an experimental (■) and theoretical value (△) for the bulk interaction parameter. (a) $T = 1000$ K. (b) $T = 600$ K.

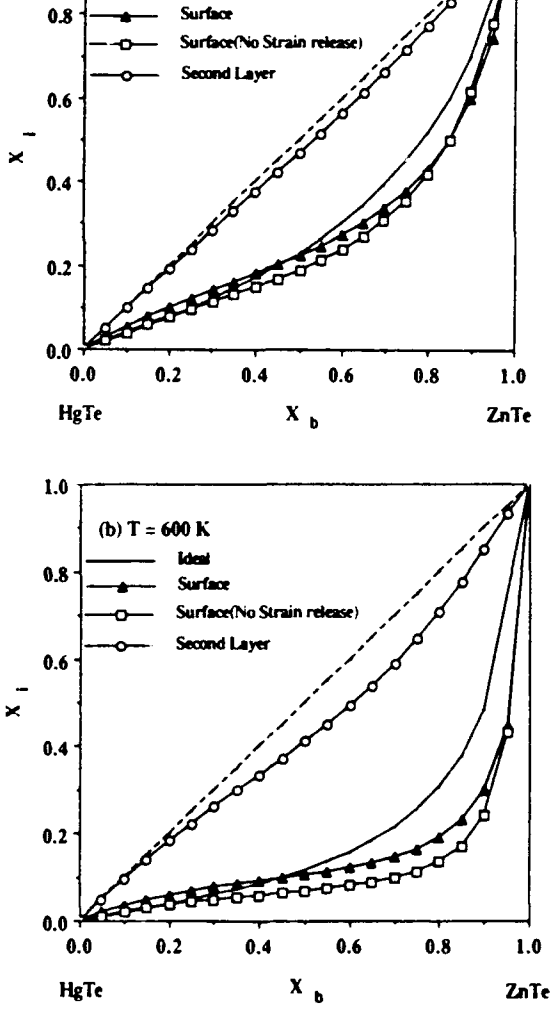


FIG. 2. Fractional surface concentration (x_i) of ZnTe as a function of the fractional bulk ZnTe concentration (x_b) at the (111) face of $Hg_{1-x}Zn_xTe$ at temperatures of (a) $T = 1000$ K and (b) $T = 600$ K. All nonideal curves were calculated in the quasichemical approximation with the theoretical interaction parameter from Chen and Sher ($\varepsilon/k = 62.94$ K).

cases we have looked at, the regular and QCA formalisms gave similar results, so only the QCA are presented. This is borne out when one calculates the short range disorder parameters, which were always close to one for all interaction parameters and temperatures considered.

Figures 1(a) and 1(b) show the results for $Hg_{1-x}Cd_xTe$ at temperatures of 1000 and 600 K. For the bulk interaction parameter, the theoretical value from Chen and Sher[14] is $\varepsilon/k = -2.92$ K, while an experimental value[16] of $\varepsilon/k = 58.75$ K is also shown for comparison. For both cases the surface interaction parameter is $\varepsilon^S/k = -4.23$ K, calculated the same way as the theoretical bulk value but with the strain term set equal to zero. Since the theoretical parameter is small, both regular and QCA results are very close to the ideal case. The layers concentrations beneath the surface are very close to the bulk so are not shown. As expected, all figures show that for all alloy concentrations the component with the weaker bonds (Hg) has segregated to the surface. Also note that as the temperature increases the amount of segregation decreases because the entropy term ($-T\Delta S$) is now dominating to create more disorder (i.e., less segregation).

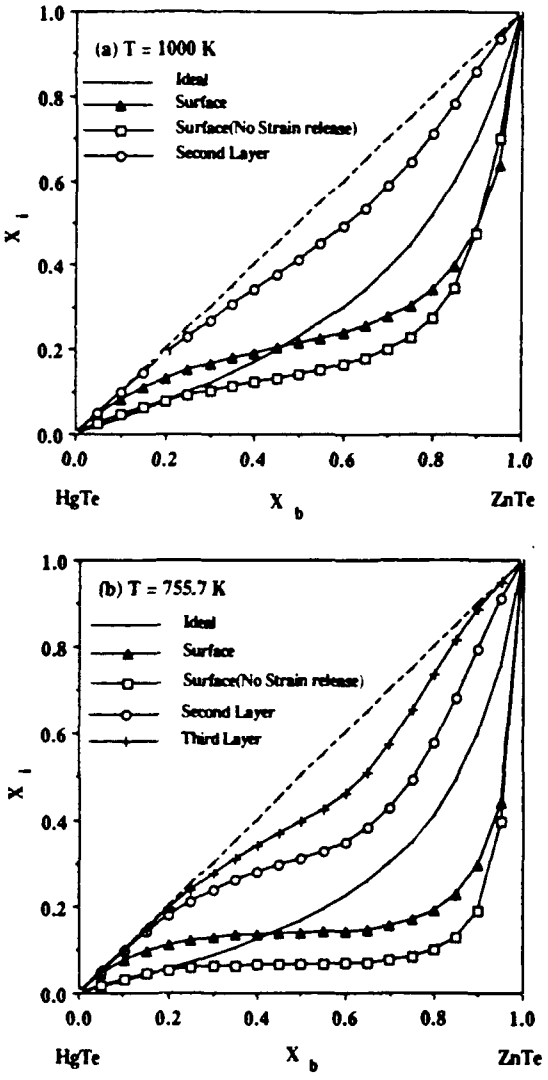


FIG. 3. Fractional surface concentration (x_s) of ZnTe as a function of the fractional bulk ZnTe concentration (x_b) at the (111) face of $Hg_{1-x}Zn_xTe$ at temperatures of (a) $T = 1000$ K and (b) $T = 755.7$ K. All nonideal curves were calculated in the quasichemical approximation with the experimental interaction parameter ($\varepsilon/k = 125.96$ K).

Presented in this paper are results for $Hg_{1-x}Zn_xTe$ using both an experimental interaction parameter, $\varepsilon/k = 125.96$ K, and a theoretical value $\varepsilon/k = 62.94$ K. In all cases the surface interaction parameter is set at $\varepsilon^S/k = 6.04$ K. Figures 2(a) and 2(b) show the segregation curves using the theoretical interaction parameter at temperatures of 1000 and 600 K. As with the $Hg_{1-x}Cd_xTe$ case, Hg segregates to the surface layers with the amount of segregation increasing with decreasing temperature. In addition, the layers below the surface begin to show appreciable deviations from bulk concentrations as the temperature decreases. In order to show the effect of strain release in our calculations, each figure also contains calculated results where the surface interaction parameter is set equal to the bulk interaction parameter. Note that in each case, the amount of segregation due to strain release is reduced in the HgTe-rich side and to a smaller degree the amount of segregation is enhanced in the ZnTe-rich side. This confirms the statements of Wynblatt and Wu[12,13] in regards to segregation in dilute compounds.

Figures 3(a) and 3(b) show the results using the experimental interaction parameter at $T = 1000$ and 755.7 K. The last temperature is the critical temperature for miscibility calculated in the regular solution theory ($T_c = Z\varepsilon/2k$). As expected, the trends are the same as with the previous case. More segregation is observed because of the increased value of ε/kT.

IV. DISCUSSION

Models for surface segregation within the framework of the regular and quasichemical approximations have been applied to the II-VI compounds $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$. As mentioned in the introduction, it has been suggested that $Hg_{1-x}Zn_xTe$ might be a better candidate for an infrared material than $Hg_{1-x}Cd_xTe$. In regards to surface segregation, our results show this to be the case. Let us compare the amount of Hg enrichment at the surface around the bulk concentration $x_b = 0.1-0.3$, the typical region for infrared applications. Referring back to Figs. 1–3, for any given temperature, more Hg segregation is seen in $Hg_{1-x}Cd_xTe$ than in $Hg_{1-x}Zn_xTe$, with the effects of strain release providing even less Hg segregation in the Hg-rich side of $Hg_{1-x}Zn_xTe$. Figure 4 shows the layer concentrations at the growth temperature of 973 K for $Hg_{0.80}Cd_{0.20}Te$ and $Hg_{0.85}Zn_{0.15}Te$. At these concentrations, both have approximately the same band gap. The experimental interaction parameters were used in the calculations. The reduction in Cd at the surface layer for $Hg_{0.80}Cd_{0.20}Te$ is 65% while the reduction in Zn at the surface of $Hg_{0.85}Zn_{0.15}Te$ is only 28%.

In general, let us reiterate the basic conclusions that have come about from our studies.

(i) The amount of surface segregation on the top layer

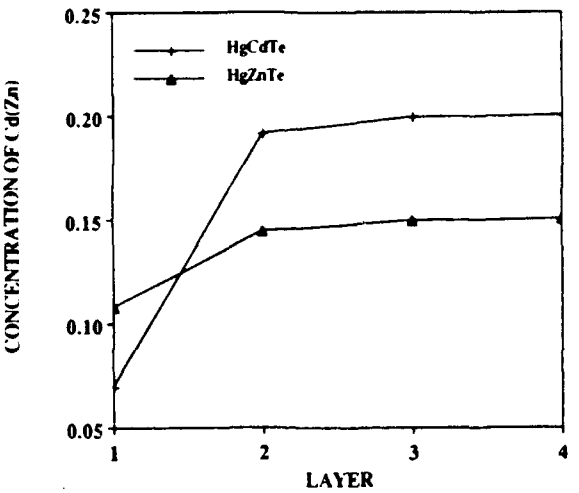


FIG. 4. Layer concentrations at the growth temperature of 973 K for $Hg_{0.80}Cd_{0.20}Te$ and $Hg_{0.85}Zn_{0.15}Te$. Experimental interaction parameters were used.

depends primarily on the difference in the vacancy formation energies, $\Delta E_A - \Delta E_B$, which is the difference in the energies required to move the constituent atoms from the bulk to the surface in the alloy. As expected, the weaker-bonding elements segregate to the top layers.

(ii) The strength of the second-neighbor term ε/kT determines how far from the ideal case the curves will be and how many layers below the surface some appreciable amount of segregation will occur.

(iii) The effect of strain release on segregation is determined in this formalism by the difference in our surface and bulk interaction parameters. As with dilute alloy cases, if A is the atom that goes to the surface most easily in the absence of strain, then strain release in $A_xB_{1-x}C$ pseudobinary alloys will reduce segregation at the BC-rich side and enhance it somewhat at the AC-rich side.

The models presented here have considered only pair interactions. Cluster models, where the alloy is divided up into microclusters of arbitrary size and all interactions within the cluster considered, are better suited for zinc-blende structures and provide a measure of short-range correlation in the alloy.[19] The application of a cluster model in surface segregation studies should provide more exact results, especially at temperatures below the critical temperature for order and/or disorder. However, to progress beyond predictions of trends more accurate estimates of the segregation interaction parameters are needed.

While our results for II-VI alloys have agreed qualitatively with metallic systems in regards to their surface behavior, we have not found any specific surface segregation experiments or semiconductor alloys to compare with. We call for experimentalists to look into this phenomenon in this important class of alloys.

ACKNOWLEDGMENTS

This work was supported by the U.S. Air Force Office of Scientific Research (AFOSR) under Contract No. F49620-85-K-0023 and by NASA under Contract Nos. NAG-1-708, NAS1-18232, and NAS1-18226.

[1]T. M. Buck in *Chemistry and Physics of Solid Surfaces IV*, edited by R. Vanselow and R. Howe (Springer-Verlag, Berlin, 1982), and references therein.

[2]J. W. Gibbs, *The Scientific Papers of J. Willard Gibbs* (Dover, New York, 1961), Vol. 1, p. 219.

[3]S. H. Overbury, P. A. Betrand, and G. A. Somorjai, Chem. Rev. **75**, 547 (1975).

[4]F. L. Williams and D. Nason, Surf. Sci. **45**, 377 (1974).

[5]D. Kumar, A. Mookerjee, and V. Kumar, J. Phys. F **5**, 725 (1976).

[6]V. Kumar, D. Kumar, and S. K. Joshi, Phys. Rev. B **19**, 1954 (1979).

[7]V. Kumar, Phys. Rev. B **23**, 3756 (1981).

[8]E. A. Guggenheim, *Mixtures* (Clarendon, Oxford, 1952).

[9]J. L. Moran-Lopez and L. M. Falicov, Phys. Rev. B **18**, 2542 (1978); **18**, 2549 (1978).

[10]R. Kikuchi, Phys. Rev. **81**, 988 (1951).

[11]D. McLean, *Grain Boundaries in Metals* (Clarendon, Oxford, 1957).

[12]P. Wynblatt and R. C. Ku, Surf. Sci. **65**, 511 (1977).

[13]P. Wynblatt and R. C. Ku, in *Interfacial Segregation*, edited by W. C. Johnson and J. M. Blakely (American Society for Metals, Metal Park, Ohio, 1977).

[14]A.-B. Chen and A. Sher, Phys. Rev. B **32**, 3695 (1985).

[15]M. A. Berding, A. Sher, and A.-B. Chen, J. Vac. Sci. Technol. A **5**, 3009 (1987); and unpublished. In the latter paper deviations of ΔE from linearity are discussed. Since the deviations are small, we have used the linear approximation here.

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

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

[18]J. Friedel, Adv. Phys. **3**, 446 (1954).

[19]A. Sher, M. van Schilfgaarde, A.-B. Chen, and W. Chen, Phys. Rev. B **34**, 4279 (1987).

ON PASSIVATION OF $Hg_{1-x}Cd_xTe$ AND $Hg_{1-x}Zn_xTe$

A. Sher, and M.A. Berding
SRI International, Menlo Park, CA 94025

R.S. Patrick, and A.-B. Chen, Physics Department
Auburn University, Auburn AL 36849

ABSTRACT

We argue that passivation of HgCdTe and HgZnTe is related to tendencies of the alloy constituents (and impurities) to segregate to interfaces. In the worst case, the surface is so Hg-rich that it is a semimetal. We present a model that offers insight into the mechanisms driving segregation for these alloys to vacuum, CdTe, ZnTe, ZnS, and various oxide interfaces, The crystal-orientation-dependence of the effect is also discussed. We conclude that segregation is minimized by using CdTe or ZnTe as the passivant.

I INTRODUCTION

This paper presents a model of the behavior of various passivants used on $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$ alloys. The model is capable of explaining many of the phenomena that impact surface leakage, which surfaces are passivated most easily, and how various passivants modify active device elements. The model treats segregation of the alloy constituents at interfaces and is extendable to impurities. In all alloys, the equilibrium concentration is never constant from the bulk to an interface (particularly a free

* This paper was presented as an added paper at the 1989 meeting.

surface); one species always segregates preferentially.[1] The interface concentration generally returns to its bulk value in a few atomic spacings, depending on the material in the passivating coating and the effective temperature at which the equilibrium is established. Segregation of this sort is well-documented in metal alloys,[2,3] but has received little attention in semiconductors. Concepts developed in the metals literature have been incorporated into our formalism. At this point ,many quantitative details of the model are still incomplete; however, we have identified the major underlying physical phenomena and trends among a broad range of effects can be deduced from the model's logic structure.

II DISCUSSION

The vacuum surface of an $A_{1-x}B_xC$ alloy will be discussed first. For definiteness, consider a (111) A surface. Under ideal circumstances, this surface is cation-terminated, and each cation makes three bonds to the underlying anion layer and having one dangling hybrid pointing to the vacuum. In our calculation, the solid is thought of as a collection of anion and cation layer pairs.[1] Interactions among the atoms in each layer and between the layers are summed to get the enthalpy of the solid. The entropy is also evaluated using a new quasichemical combinatorial method.[4] The free energy is minimized to find the composition x_i, i = 1,2,3,... of each layer. Details of this method can be found in prior publications.[1,4]

---

[1]Patrick, R.S., A.-B. Chen, A. Sher, and M.A. Berding, 1989: "Surface Segregation in Pseudobinary Alloys," *Phys. Rev. B*, Vol. 39, No. 9, pp. 5980-5986. Reference to much of the relevant literature of this subject can be found in this paper.

[2] Williams, F.L., and D. Nason, 1974: *Surf. Sci.*, Vol. 45, p. 377.

[3] Kumar, V., D. Kumar, and S.K. Joshi, 1979: *Phys. Rev. B*, Vol. 19, p. 725.

[4] Sher,. A., M. van Schilfgaarde, A.-B. Chen, and W. Chen, 1987: "Quasichemical Approximations in Binary Alloys," *Phys. Rev. B*, Vol. 36, No. 8, pp. 4279-4295 (September).

Two major effects drive segregation. The first, referred to as the *chemical driving energy*, is related to the fact that the solid can minimize its enthalpy by having the constituent on the surface that makes the weaker bond. However, having all of one class of atoms on the surface is unfavorable to the entropy. Hence, the surface concentration is a compromise. The second driving mechanism is *strain release*. An impurity that has a bond length mismatched with the major constituent has less strain energy on the surface than in the bulk. Hence, the low concentration constituent of a bond-length-mismatched alloy tends to go to the surface, e.g., in $Hg_{1-x}Zn_xTe$ for $x > 0.15$, the Zn is driven to the surface. Near $x = 0.5$, neither species is preferentially driven to the surface by strain release, and in the high-concentration region the opposite species goes to the surface. If the low concentration species also makes the weak bond, then the two mechanisms reinforce one another and enhance segregation; in the opposite case, they compete and segregation is reduced. In $Hg_{1-x}Zn_xTe$ for $x < 0.25$, the mechanisms compete so as to reduce segregation. Strain release plays little role in $Hg_{1-x}Cd_xTe$, because the components are nearly bond-length-matched.

In both $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$, the equilibrium surfaces are Hg-rich relative to the bulk. The layer concentration returns almost to that of the bulk by the second layer. However, in both cases, for x in the range used for long-wave detectors, the surfaces are semimetals that look n-type. Thus, any passivation procedure must begin by carefully removing any equilibrium surface layers. Figures 1 and 2 show the equilibrium vacuum surface concentration versus the bulk concentration for a sample equilibrated at 973 K. If a lower temperature had been used, the segregation would have been more severe. One problem with HgCdTe may be that diffusion persists to low temperature; hence, the effective equilibration temperature may be low.

The (111) A surface has one dangling cation bond. The (110) surface also has one dangling cation bond, but the bond density is higher than on the (111) A surface. Hence, the (110) surface segregates slightly more than (111) A. However, there are two and three dangling bonds per cation, respectively, on the (100) and (111) B surfaces. Hence, those regions of the surface that are cation-terminated should be progressively more difficult to passivate, at least in $Hg_{1-x}Cd_xTe$, where the chemical term dominates. However, particularly for the (111) B surface, a fair portion of the surface will be Te-terminated. The underlying cations will then make four bonds to Te, just as the bulk atoms do. Thus,

passivation should be easy in these regions. Because the Te-terminated patches are easy to passivate, and because the cation-terminated patches are difficult to passivate, this is another source of variability for imaging arrays.

The way these effects are modified by CdTe and ZnTe capping layers is also shown in Figs. 1 and 2. For $Hg_{1-x}Cd_xTe$, the segregation in Fig. 1 is almost completely removed by both passivants. The reason is that the cations at the interface between the HgCdTe and the capping layer now make four Te bonds, just as they do in the bulk. As a consequence, the only energy difference between a Hg or Cd in the bulk and at the interface stem from small metallization-induced local bond and small back-bond changes. The effect of these passivants on $Hg_{1-x}Z_xTe$ seen in Fig. 2 is less complete because of the strain release contribution. Here, the ZnTe leaves the interface slightly Zn-rich, while CdTe leaves the interface slightly Hg-rich relative to the bulk. However, both passivant materials still are reasonably effective.

The results in Figs. 1 and 2 are idealized: They are calculated as if the CdTe and ZnTe passivation layers stay fixed at the equilibration temperature, 973 K in these figures. This is unrealistic, the CdTe will certainly interdiffuse with, for example, the HgCdTe to leave a graded concentration region at the interface that is Cd-rich. This grading would have a range determined by the time diffusion is permitted to occur and could be many layers thick. On the other hand, the ZnTe will not interdiffuse into the HgCdTe, if after it is deposited, all processing remains below the critical spinodal transition temperature. Thus, ZnTe may prove to be a more temperature-stable passivant. Only a few atomic layers of ZnTe are needed to serve as the passivant; accordingly, strain-induced misfit dislocations can be avoided. If a thicker insulator, or a protective layer is needed (rather than just a layer to cut down surface leakage), CdTe or some insulator can be put on top of the ZnTe to finish off the passivation.

The effect of ZnS, Photox, SiO_2, or anotic oxides can also be deduced from this model. Take ZnS as an example. Now on the (111) A surface, the extra cation bond is made to a sulfur. The CdS and HgS bonds are much stronger than CdTe or HgTe, respectively, and the CdS bond is much stronger than HgS. Thus, the CdS bond dominates the behavior, and Cd preferentially segregates to the interface. While this is preferable to having a Hg-rich interface, because it opens the bandgap rather than narrowing it, other troublesome side effects may be caused. We have not yet discussed dopant

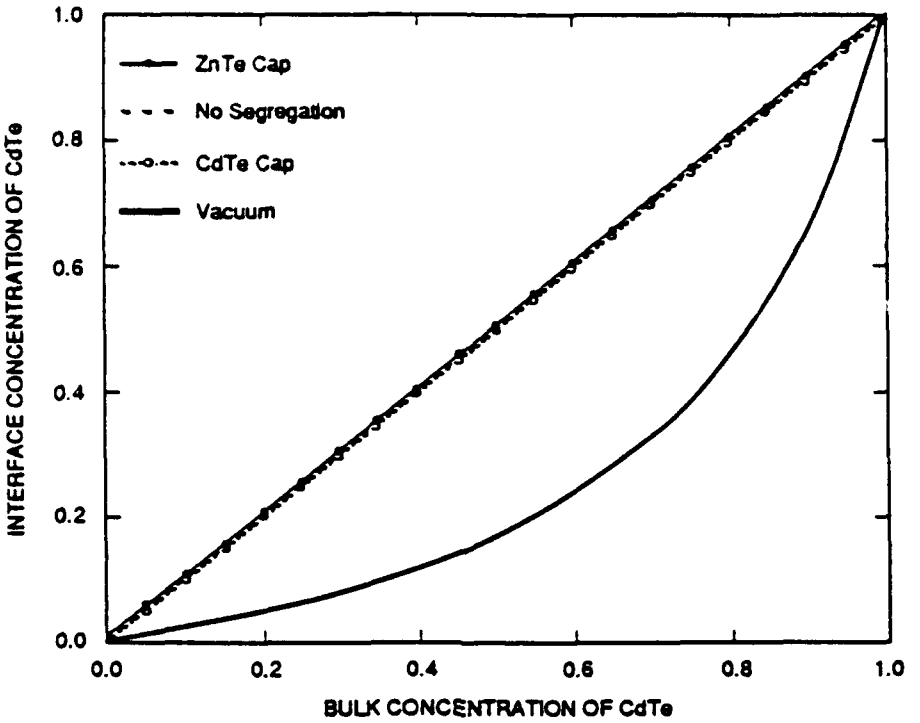


FIGURE 1 INTERFACE CONCENTRATION OF Cd IN $Hg_{1-x}Cd_xTe$ FOR VARIOUS PASSIVATING CAPPING MATERIALS

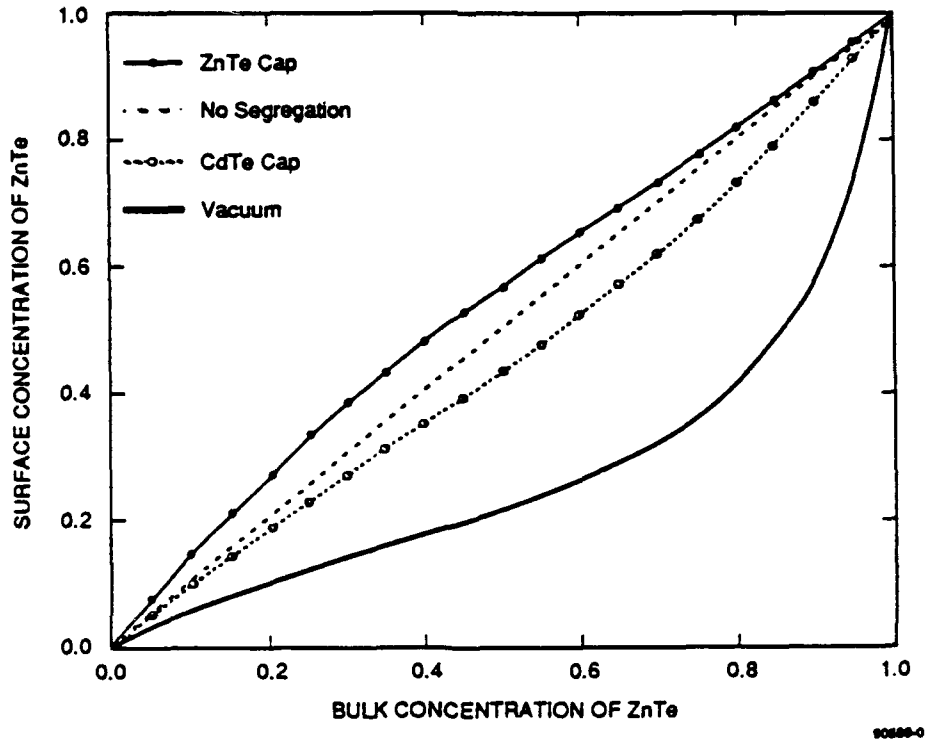


FIGURE 2 INTERFACE CONCENTRATION OF ZnTe IN $Hg_{1-x}Zn_xTe$ FOR VARIOUS PASSIVATING CAPPING MATERIALS

(nonisoelectronic impurity) segregation. If the dopants preferentially segregate to the interface, that too can cause high surface leakage. Most impurities also make stronger bonds to S than to Te, and so will tend to segregate to the interface. No quantitative analysis of these effects have been done yet, but the direction of trends is clear.

Oxides will behave like the sulfides, but the likelihood of epitaxial interfaces over substantial regions is smaller. Hence at a (111) A interface, for example, where the bond density of the semiconductor and capping layer are grossly different, one expects, a distribution of cation-oxygen and dangling bonds. Thus, segregation will be driven by the statistical average of these effects over a given domain. If different domains have different oxide crystal orientations, stoichiometries, or (in the native oxide case) different compositions (e.g., tellurium oxide, cadmium oxide), or mercury oxide, then there will be segregation patterns over the $Hg_{1-x}Cd_xTe$ surface with different cd and dopant concentrations. As long as the interface is always Cd-rich, this may cause minimal problems, but it is bound to affect the uniformity of array performance.

III SUMMARY

The model presented here is incomplete, but broad guidelines for $Hg_{1-x}Cd_xTe$ passivation can be stated:

- The top few layers of material must be removed before deposition of a passivation layer.
- Te-bearing compounds are the best materials choice to minimize interface concentration profiles.
- Epitaxial layers are best, but if polycrystalline materials are used, they should at least be properly oriented polycrystals.
- ZnTe may be more stable against interface thermal degradation than CdTe.
- The easiest orientations to passivate are the (111) A and (110), and progressively harder ones are cation patches on the (100) and (111) B surfaces and on these latter surfaces careful treatments to avoid cation patches may be especially helpful.
- Nothing has been said about radiation damage hardness yet, but the model obviously offers a logic structure in terms of which to understand the question.

ACKNOWLEDGMENTS

This work was supported in part by NASA Contract NAS1-18226 (Task 2) and Office of Naval Research Contract N00014-88-C-0096.

Semicond Sci Technol 5 (1990) S100–S102 Printed in the UK

Band structures of HgCdTe and HgZnTe alloys and superlattices

A-B Chen†, Y-M Lai-Hsu†, S Krishnamurthy‡, M A Berding‡ and A Sher‡

†Physics Department, Auburn University, Auburn, AL 36849, USA
‡SRI International, Menlo Park, CA 94025, USA

Received in accepted form 18 September 1989

Abstract. The band structures of HgTe, CdTe, ZnTe, the alloys $Hg_{1-x}Cd_xTe$ (HCT) and $Hg_{1-x}Zn_xTe$ (HZT), and several small-gap superlattices (SL) are calculated using a tight-binding model. Our calculations show a nearly linear dependence of the energy gap on the concentration in HCT, but a strong non-linear variation in HZT. The electron mass as a function of the band gap is found to be the same in HCT and HZT in the small-gap region. Our calculated SL bands, in fair agreement with the most recent experiments and theories, support the assumption of a large valence band offset (350 meV) between HgTe and CdTe.

1. Introduction

Because of structural weakness in $Hg_{1-x}Cd_xTe$ (HCT), the $Hg_{1-x}Zn_xTe$ alloys (HZT) and related superlattices (SL) may become competitive infrared materials. In this paper we examine the important parts of the band structures of HCT, HZT and several SL systems. Our study is based on a second-neighbour tight-binding (SNTB) model, which has been implemented for all the systems concerned, including the pure compounds, the alloys and the SL. The spin-orbit interactions and the strain effects in the SL are also included. The alloy disorder is treated within the molecular coherent potential approximation (MCPA) [1, 2]. The SL band structures are calculated using a difference equation approach [3].

2. Pure compounds HgTe, CdTe and ZnTe

The SNTB model uses four orthonormal local s- and p-orbitals per atom. The first- and second-neighbour interactions are assumed to be of the two-centre type. The s-term values ε_s are taken such that the $-\varepsilon_s$ are the experimental first ionisation energies of the atoms. The p-term values ε_x and other interaction parameters are adjusted to fit the important band quantities of the pure compounds. Figures 1(*a*)-1(*c*) display the calculated bands for the three pure compounds. In comparing the present model with the most popular TB band model for the HgTe CdTe SL by Schulman and Chang (SC) [4], we found that the bands around the gaps from both models are rather similar. The largest discrepancy between the two models is in the valence band width. Our model puts the X_3, level at the experimental values [5] of -5 and -5.5 eV respectively for CdTe and HgTe, while the CS model sets it about 2 eV deeper. Since this discrepancy happens at an energy far from the fundamental gap, it should not strongly affect the band structures of the SL near the gap.

3. HCT and HZT alloys

The band parameters of the pure compounds are used in the alloy calculation without any further adjustment. The dominant disorder in HCT and HZT comes from the large difference in the s-term values between the cations. However, the fluctuations in the off-diagonal matrix elements will also affect the detailed band structures. This alloy disorder is treated within the MCPA as described in [1, 2].

The major results for the alloys are presented in figure 2, where the band gap is plotted as a function of the alloy concentration, and in figure 3, where the conduction band mass is plotted as a function of the band gap. For HCT the band gap E_g deviates only slightly from the straight-line average of the pure compound values, E_g'. However, HZT shows a significant bowing below the average value. Figure 3 indicates that the m_c^*-value for a device operating at the same wavelength in the narrow-gap region is nearly the same whether HCT or HZT is used. Finally we note that the present calculation agrees with [2] in the MCPA self-energy, indicating that the disorder effect on the lifetime is important only for states well

0268-1242/90/03S100 + 03 $03.50

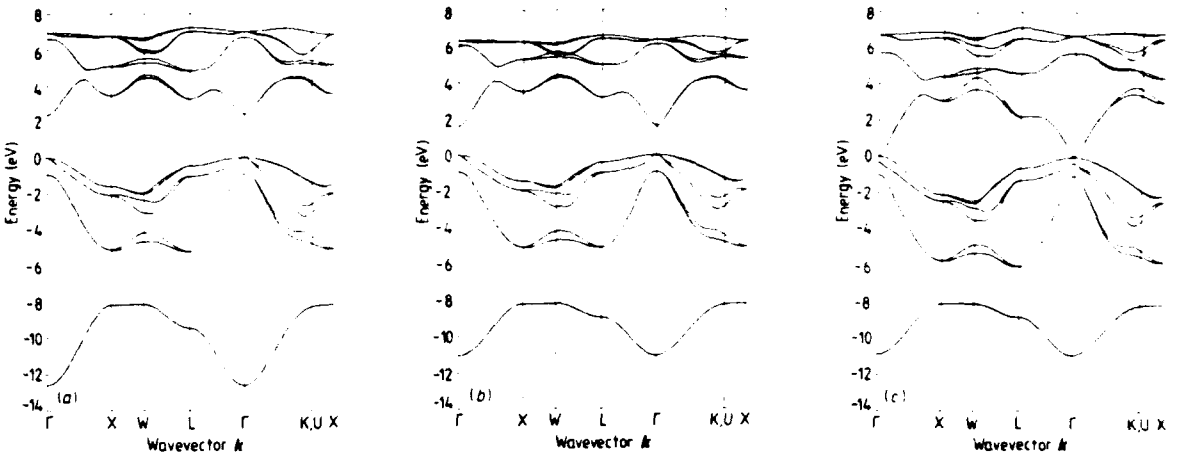


Figure 1. *The calculated band structures of (a) ZnTe, (b) CdTe and (c) HgTe.*

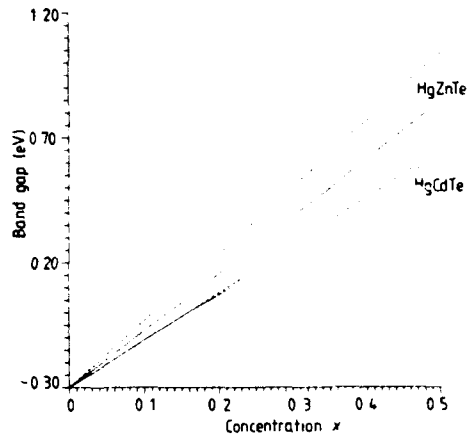


Figure 2. The band gaps of $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$ as a function of the alloy concentration x.

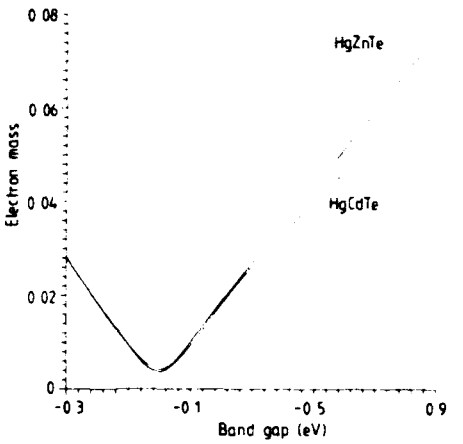


Figure 3. The electron masses of $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$ as a function of the band gap

above and below the band edges in both HCT and HZT, and is not a limiting factor on the mobilities of electrons and holes in the low-field transport in HCT and HZT.

4. Superlattices

To use the TB model to treat the SL, we further allow a rigid shift of all the term values in a slab with respect to the other to produce the desired valence band offset ΔE_v. Table 1 compares our calculated band gaps at $k = 0$ for the (001) SL $HgTe/Cd_{0.85}Hg_{0.15}Te$ with the experimental and theoretical values based on the SC model quoted in [6] for several combinations of slab thickness represented by the numbers of double layers of atoms (n, m) in the slabs. First we see a small but consistent discrepancy between the two theoretical models. The SC model yields a gap about 10–30 meV smaller than ours. The experimental data are in better agreement with the calculations with a valence band offset $\Delta E_v = 350$ meV between HgTe and CdTe than with $\Delta E_v = 40$ meV. However, the agreements are not really satisfactory. For example, it is apparent from the theory that the band shrinks as the ratio n,m increases, whereas the experimental data show an increase in the gap going from the (16, 16) to (17, 15) samples. For those samples in which there is a clear gap, our calculations with $\Delta E_v = 350$ meV are seen to agree with experiments slightly better than the SC model. The comparisons for the first three systems with (n,m) equal to (24, 9), (25, 10) and (24, 15) are less transparent, because the SL band structure has a nearly zero or negative gap. While we only list the gap at $k = 0$, the actual gaps may occur at other k points. The complexity of the bands may be illustrated by comparing our bands for the sample with $(n, m) = (26, 9)$ shown in figure 4 with those in figure 9 of [6] for the same system using the same $\Delta E_v = 350$ meV.

In figure 4 the bands are plotted along k_z (perpendicular to the slab) and k_x (parallel to the slabs). Along k_z, all the bands which are derived from the heavy-hole bands of the slabs are practically horizontal lines owing to the large ΔE_v used. The only two bands with appreciable dispersion are those derived from the conduction band and the light-hole band. The one derived from the

Table 1. Comparison between the calculated band gaps (in meV) in the present work with the experimental and theoretical values of [6] for the (001)$HgTe/Hg_{0.15}Cd_{0.85}Te$ superlattices

No of layers (HgTe/HgCdTe)	Theory				Experiment [6]	
	ΔE_v = 40 meV		ΔE_v = 350 meV		Transport	Optical
	Present	[6]	Present	[6]		
(24,9)	33	5	−1	−30	< 5	−16
(24,10)	37	12	2	16 − 26	< 8	−10
(24,15)	65	51	16	1	< 5	−3
(19,12)	89	68	43	25	35	41
(18,14)	107	89	56	42	54	53
(16,16)	135	119	81	70	56	
(17,15)	121	104	68	55	80	64
(18,18)	117	104	61	50	81	

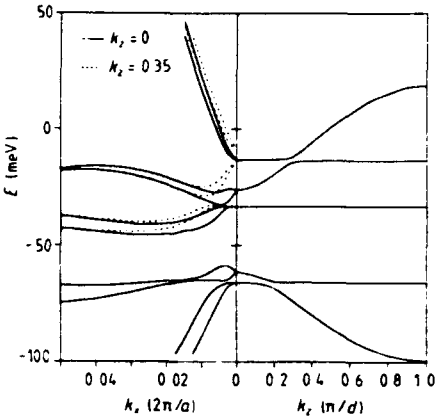


Figure 4. The calculated band structure of (001) $HgTe/Hg_{0.15}Cd_{0.85}Te$ superlattice with the numbers of layers (n, m) = (26, 9) and with a valence band offset ΔE_v = 350 meV between HgTe and CdTe.

conduction band crosses a heavy-hole band at $k_z = k_d = 0.35\pi/d$ in our model and at $k_d = 0.42\pi/d$ in [9], where d is the SL period. This is therefore a zero-gap system. The light-hole mass for the band along k_x for $k_z = k_d$, as shown by the broken curve, is extremely small, $m^* < 0.001$; this was also found in the Shubnikov-de Haas experiment on the hole orbit of this system by Seiler *et al* [7], who also found a sudden increase in the mass going from the low to the high magnetic field. They indicated that the rapidly changing mass as a function of k_x can explain the experimental result. The largest difference between the two calculations is the location of the band originating from the slab conduction band at $k = 0$. Our calculation gives an inverted gap of −12 meV while the CS model gave a value of about −40 meV. We also note that in the side bands (along k_x) with $k_z = 0$, the top valence band goes down and then up as k_x increases. This is due mainly to the inclusion of the strain in the HgTe slabs. This intrinsic complexity in the SL band structures and other extrinsic complications such as interfacial diffusion and non-uniformity of the Hg concentration are among the difficulties in pinning down the band structures in the small-gap SL.

5. Conclusions

The band structures for the pure compounds HgTe, CdTe and ZnTe, their alloys and the SL made from these systems have been studied using an SNTB model. The results for the important parts of the band structures for the pure compounds and alloys are in substantial agreement with our previous results [2] based on a more elaborate scheme. The most important result for the alloys is that HZT is as good an infrared material as HCT in terms of the electronic structures in the small-gap region. Our results for the SL, in fair agreement with the most recent theoretical and experimental studies, support the assumption of a large valence band offset of 350 meV between HgTe and CdTe. However, the SL band structure is less understood than that in the alloys. More detailed comparisons between theories and experiments are needed to improve our understanding of the band structures of small-gap superlattices.

Acknowledgments

This work was supported in part by the US NASA, AFOSR and ONR.

References

[1] Hass K C, Lampert R J and Ehrenreich H 1984 *Phys. Rev. Lett.* **52** 77
[2] Berding M A, Krishnamurthy S, Sher A and Chen A-B 1987 *J. Vac. Sci. Technol.* A **5** 3014
[3] Chen A-B, Lai-Hsu Y-M and Chen W 1989 *Phys. Rev.* B **39** 923
[4] Schulman J N and Chang Y-C 1986 *Phys. Rev.* B **33** 2594
[5] Spicer W E, Silberman J A, Morgen J, Lindau A, Wilson J A, Chen A-B and Sher A 1982 *Phys. Rev. Lett.* **13** 948
[6] Hoffman G A, Meyer J R, Bartoli F J, Han J W, Cook J W Jr, Schetzina J F and Schulman J N 1989 *Phys. Rev.* B **39** 5208
[7] Seiler D G, Ward G B, Justice R J, Koestner R J, Goodwin M W, Kinch M A and Meyer J R 1989 *J. Appl. Phys.*

Full-potential Korringa-Kohn-Rostoker band theory applied to the Mathieu potential

Chin-Yu Yeh and A.-B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849

D. M. Nicholson and W. H. Butler
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830
(Received 21 May 1990; revised manuscript received 8 August 1990)

The band theory of Korringa, Kohn, and Rostoker (KKR) based on the Green-function method is extended to space-filling potentials. A numerical test using the Mathieu potential shows good convergence for the bands up to 1.5 Ry with $l \leq 4$ included in the angular-momentum expansion for the wave functions. Our results strongly support the applicability of the full-potential KKR to bulk electronic-structure problems.

I. INTRODUCTION

The Korringa, Kohn, and Rostoker (KKR) band theory[1,2] is an elegant theory for the one-electron energy bands in a closed-packed crystal for which the muffin-tin (MT) construction for the potential is a reasonable approximation. To expand the scope of application, considerable effort has been expended to extend the KKR theory to full crystal potentials.[3-11] One concern about such extension is related to the so-called near-field corrections (NFC) (Refs. 3–5) arising from the expansion of the KKR Green function beyond the muffin-tin region. Although there are proofs[2,9,11] showing that NFC do not exist, questions have been raised about the applicability of the theory.[12] Since space-filling potentials are nonspherical and the Wigner-Seitz cell boundary is not smooth, we are further concerned about the speed of convergence in terms of angular-momentum (l) expansions. In this paper the integral equation approach of Kohn and Rostoker[2] (KR) is used to derive the full-potential KKR (FP-KKR) equation explicitly. One advantage of our derivation is that all the quantities involved are functions of $\mathbf{r}$ within a unit cell. Thus we can avoid the uncertainty in extending the wave function beyond the unit cell encountered in some other derivations.[6,11] We have also tested the convergence by comparing the numerical results with the exact solution for the Mathieu potential[13-15] in the simple-cubic crystal. Excellent results for the band structure in the energy range of interest are obtained with a maximum value of $l=4$ included in this expansion.

The fact that the Mathieu potential is exactly soluble gives it an advantage for testing purposes over working with realistic potentials.[4,9,17] Our test complements the empty-lattice potential[18-20] to provide a stringent test for the FP-KKR theory. The strong angular-momentum dependence in the Mathieu potential gives a good representation of the anisotropy that is present in the open structures pertaining to many semiconductors and insulators. The restriction of the KKR to closed-packed metals imposed by the muffin-tin approximation is lifted by the full-potential method discussed here. The results obtained here should encourage the application of this theory to real crystals.

II. THE FULL-POTENTIAL KKR EQUATION

In this section we want to show that the Kohn-Rostoker integral equation can be simply extended to obtain the full potential KKR theory. The Schrödinger equation in the band calculation

$$[-\nabla^2+V(\mathbf{r})]\psi_{\mathbf{k}}(\mathbf{r})=E\psi_{\mathbf{k}}(\mathbf{r}) \tag{1}$$

for a full crystal potential $V(\mathbf{r})$ is equivalent to solving the following integral equation:[2]

$$\psi_{\mathbf{k}}(E,\mathbf{r})=\int_{\tau} G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')V(\mathbf{r}')\psi_{\mathbf{k}}(E,\mathbf{r}')d\mathbf{r}' , \tag{2}$$

where the integration is over the Wigner-Seitz cell of volume τ, and $\mathbf{k}$ is a crystal wave vector. $G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')$ in Eq. (3) is the KKR free-electron Green's function[2]

$$G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')=-\frac{1}{\tau}\sum_{n}\frac{\exp[i(\mathbf{K}_n+\mathbf{k})\cdot(\mathbf{r}-\mathbf{r}')]}{(\mathbf{K}_n+\mathbf{k})^2-E} , \tag{3}$$

where $\mathbf{K}_n$ are the reciprocal-lattice vectors. Alternatively $G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')$ can be expressed as[2]

$$G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')=-\frac{1}{4\pi}\sum_{s}\frac{\exp(i\kappa|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|)}{|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|}\exp(i\mathbf{k}\cdot\mathbf{R}_s) , \tag{4}$$

where $\kappa=\sqrt{E}$ for $E>0$ and $\kappa=i\sqrt{E}$ for $E<0$, and $\mathbf{R}_s$ are the lattice translation vectors. To derive the FP-KKR equation, we first observe that Eq. (3) can be cast into a surface integral,

$$\int_{S_\tau}[G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')\nabla'\psi_{\mathbf{k}}(E,\mathbf{r}') - \psi_{\mathbf{k}}(E,\mathbf{r}')\nabla'G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')]\cdot d\mathbf{S}'=0 , \tag{5}$$

where S_τ is the surface of the Wigner-Seitz cell.

Since the $\mathbf{r}'$ in the surface integral exceeds the muffin-tin radius r_m, we need to consider the expansion of

the Green function beyond the original range of Kohn and Rostoker. Several authors have already considered this point. For simplicity, we shall only consider the case with one atom per unit cell. Since this expansion is a central point of controversy, we rederive the results explicitly in the Appendix for the range of $\mathbf{r}$ and $\mathbf{r}'$ needed here. We show that the expansion

$$G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')=\sum_{L}\left\{\sum_{L'}[i^{l-l'}B_{L,L'}(\mathbf{k},E)J_L(\kappa\mathbf{r})J_{L'}(\kappa\mathbf{r}')+\kappa\delta_{LL'}J_L(\kappa\mathbf{r})N_L(\kappa\mathbf{r}')]\right\} \tag{6}$$

is valid as long as both $\mathbf{r}$ and $\mathbf{r}'$ are inside τ and satisfy the following condition:

$$r<r'<|\mathbf{R}_s| \quad \text{for all } \mathbf{R}_s\neq 0 . \tag{7}$$

In Eq. (6) the notations $J_L(\kappa\mathbf{r})=j_l(\kappa r)Y_L(\hat{\mathbf{r}})$ and $N_L(\kappa\mathbf{r})=n_l(\kappa r)Y_L(\hat{\mathbf{r}})$ are used, where j_l and n_l are, respectively, the spherical Bessel and Neumann functions, Y_L is a real spherical harmonics, and L represents the double indices (l,m). $B_{L,L'}(\mathbf{k},E)$ is the usual KKR structure constant.[2,13] We note that for any $|\mathbf{r}|$ smaller than r_m, the condition in Eq. (7) is satisfied for all r' contributing to the surface integration in Eq. (5). The condition $|\mathbf{r}'|<\mathbf{R}_s$ in Eq. (7) holds for most lattices; exceptions are those, for example, with long narrow cells. For such cases, this condition can be satisfied by breaking the unit cell into smaller cells including so-called "empty cells" which do not contain an atomic nucleus.

The wave function inside the cell τ can be expanded in a basis set $\{\Phi_L(E,\mathbf{r})\}$ as

$$\psi_{\mathbf{k}}(E,\mathbf{r})=\sum_{L}a_L(\mathbf{k},E)\Phi_L(E,\mathbf{r}) . \tag{8}$$

The basis function $\Phi_L(E,\mathbf{r})$ is a regular solution to the Schrödinger equation inside τ,

$$[-\nabla^2+V(\mathbf{r})]\Phi_L(E,\mathbf{r})=E\Phi_L(E,\mathbf{r}) , \tag{9}$$

and behaves like $J_L(\kappa\mathbf{r})$ at the origin $r=r_0\to 0$, which is typically the location of the atomic nucleus. This basis set can be calculated using the following integral equation:[4]

$$\Phi_L(E,\mathbf{r})=J_L(\kappa\mathbf{r})+\sum_{L'}\int_{r_0}^{r}g_L(E;\mathbf{r},\mathbf{r}')V(\mathbf{r}')\Phi_{L'}(E,\mathbf{r}')d^3r' , \tag{10}$$

where $g_L(E;\mathbf{r},\mathbf{r}')$ is a free-particle Green's function and is defined as

$$g_L(E;\mathbf{r},\mathbf{r}')=\kappa[J_L(\kappa\mathbf{r})N_L(\kappa\mathbf{r}')-N_L(\kappa\mathbf{r})J_L(\kappa\mathbf{r}')] . \tag{11}$$

We note that the basis function $\Phi_L(E,\mathbf{r})$ is coupled to other angular-momentum channels for $r>r_0$, because the crystal potential $V(\mathbf{r})$ is not spherical.

The expansions of G in Eq. (6) and of ψ in Eq. (8) can be substituted in Eq. (5) to obtain

$$\sum_{L}J_L(\kappa\mathbf{r})\sum_{L'}\left[\left\{\sum_{L''}i^{l-l''}B_{LL''}(\mathbf{k},E)S_{L''L'}(E)\right\}+\kappa C_{LL'}(E)\right]a_{L'}(\mathbf{k},E)=0, \quad r<r_m , \tag{12}$$

where

$$S_{L'L}(E)=\kappa\int_{S_-}[J_{L'}(\kappa\mathbf{r}'),\Phi_L(E,\mathbf{r}')]\cdot d\mathbf{S}' , \tag{13}$$

and

$$C_{LL'}(E)=\kappa\int_{S_-}[N_L(\kappa\mathbf{r}'),\Phi_{L'}(E,\mathbf{r}')]\cdot d\mathbf{S}' . \tag{14}$$

In the above, the notation $[F_1,F_2]\equiv F_1\nabla'F_2-F_2\nabla'F_1$ has been used. The surface integrals in Eqs. (13) and (14) are over the boundaries of τ as indicated by S_-. Since $J_L(\kappa\mathbf{r})$ in Eq. (12) are linearly independent functions, the following set of homogeneous equations holds:

$$\sum_{L'}\left[\left\{\sum_{L''}i^{l-l''}B_{LL''}(\mathbf{k},E)S_{L''L'}(E)\right\}+\kappa C_{LL'}(E)\right]a_{L'}(\mathbf{k},E)=0 . \tag{15}$$

This is the FP-KKR equation that we are after.

We note that our derivation is similar to Nesbet's derivation.[11] We hope, however, that the above explicit derivation may be more accessible to some readers. It is also useful for establishing the notation necessary for the description of the application of FP-KKR theory to the solution of the Mathieu potential which constitutes the main result of this paper.

III. CALCULATION OF S AND C MATRICES

The surface integrals for the S and C matrices in Eqs. (13) and (14) can be very time consuming. It is desirable to seek simplification of these calculations in a real application. One plausible approximation which is consistent with the KKR spirit is to expand every quantity involved in angular-momentum components. Equations (13) and (14) are equivalent to the volume integrations

$$S_{L'L}(E)=\kappa\int_{\tau}J_{L'}(\kappa\mathbf{r})V(\mathbf{r})\Phi_L(E,\mathbf{r})d\mathbf{r} \tag{16}$$

and

$$C_{L'L}(E)=-\delta_{L,L'}+\kappa\int N_{L'}(\kappa r)V(\mathbf{r})\Phi_L(E,\mathbf{r})d\tau\ . \tag{17}$$

One can free the limits of these integrals by replacing the crystal potential V by the truncated potential V^T,

$$V^T(\mathbf{r})=V(\mathbf{r})\sigma(\mathbf{r})\ , \tag{18}$$

where $\sigma(\mathbf{r})$ is a step function and is defined as

$$\sigma(\mathbf{r})=\begin{cases}1, & \text{for } \mathbf{r} \text{ within } \tau\\ 0, & \text{otherwise}\ .\end{cases} \tag{19}$$

The angular-momentum expansion for the basis function is assumed to be

$$\Phi_L(E,\mathbf{r})=\sum_{L'}\phi_{L'L}(E,r)Y_{L'}(\mathbf{r})\ , \tag{20}$$

and the truncated potential V^T is expanded as

$$V^T(\mathbf{r})=\sum_L V_L^T(r)Y_L(\mathbf{r})\ . \tag{21}$$

The integrations in Eqs. (16) and (17) can be reduced, respectively, to the simple radial integrations

$$S_{L'L}(E)=\kappa\sum_{L''}\int_{r_0}^{r_s} j_{L'}(\kappa r)V_{L'L''}(r)\phi_{L''L}(E,r)r^2dr \tag{22}$$

and

$$C_{L'L}(E)=-\delta_{L,L'}+\kappa\sum_{L''}\int_{r_0}^{r_s} n_{L'}(\kappa r)V_{L'L''}(r)\times\phi_{L''L}(E,r)r^2dr\ , \tag{23}$$

where r_s is the radius of the circumscribing sphere of the Wigner-Seitz cell. The $V_{LL'}(r)$ is given by

$$V_{L'L''}(r)=\int Y_{L'}(\mathbf{r})V^T(\mathbf{r})Y_{L''}(\mathbf{r})d\Omega=\sum_L C_{L'L''}^L V_L^T(r)\ , \tag{24}$$

where

$$C_{L'L''}^L\equiv\int Y_{L'}(\mathbf{r})Y_L(\mathbf{r})Y_{L''}(\mathbf{r})d\Omega \tag{25}$$

is a Gaunt coefficient.

Note that in the above the basis function $\Phi_L(E,\mathbf{r})$ is assumed to be calculated from Eq. (10), where $V(\mathbf{r})$ is the full crystal potential. This is the same procedure used by Brown and Ciftan (BC).[6] The original Williams-Morgan[3] (WM) approach, however, used the truncated potential V^T for the calculation of the basis function in Eq. (10). If expansions of the potential and Φ_L in Eq. (10) include all the angular-momentum components, both approaches probably will give the same results for the band structure, provided both converge.[11,21] In practice, the expansion is limited to a certain l_{max}; therefore, these two approaches yield different results.

In the actual calculation of the basis functions using either V or V^T, we first write the potential as the sum of V_0 and ΔV, where V_0 is the spherical part of the potential and ΔV is the rest. We then solve for the radial wave function f_l corresponding to V_0. Similarly, the basis is written as $\Phi_L=F_L+\Delta\Phi_L$, where $F_L=f_l Y_L$, and $\Delta\Phi_L$ is solved from the integral equation

$$\Delta\Phi_L(E;\mathbf{r})=\sum_{L'}\int_{r_0}^{r} g_{L'}(E;\mathbf{r},\mathbf{r}')\Delta V(\mathbf{r}')F_L(E;\mathbf{r}')d\mathbf{r}' + \sum_{L'}\int_{r_0}^{r} g_{L'}(E;\mathbf{r},\mathbf{r}')V(\mathbf{r}')\Delta\Phi_L(E;\mathbf{r}')d\mathbf{r}'\ , \tag{26}$$

iteratively using angular-momentum expansions for all quantities involved.

IV. APPLICATION TO MATHIEU POTENTIAL

To test the accuracy of the FP-KKR equation and the convergence in angular expansion described above, we applied the theory to the Mathieu potential[13-16] of the form

$$V(\mathbf{r})=-U_0\left[\cos\frac{2\pi x}{a}+\cos\frac{2\pi y}{a}+\cos\frac{2\pi z}{a}\right]\ , \tag{27}$$

where we took the lattice constant a to be 2π times the Bohr radius and the potential parameter U_0 to be 0.5 Ry. Because the potential is separable, the eigenvalue problem reduces to three one-dimensional problems. The band structures and corresponding wave functions can be computed to the precision of the computer and can be regarded as "exact" in the numerical comparison.

The Mathieu potential is poorly represented by the MT approximation, because the simple cubic structure is rather open and the potential has a large variation in the

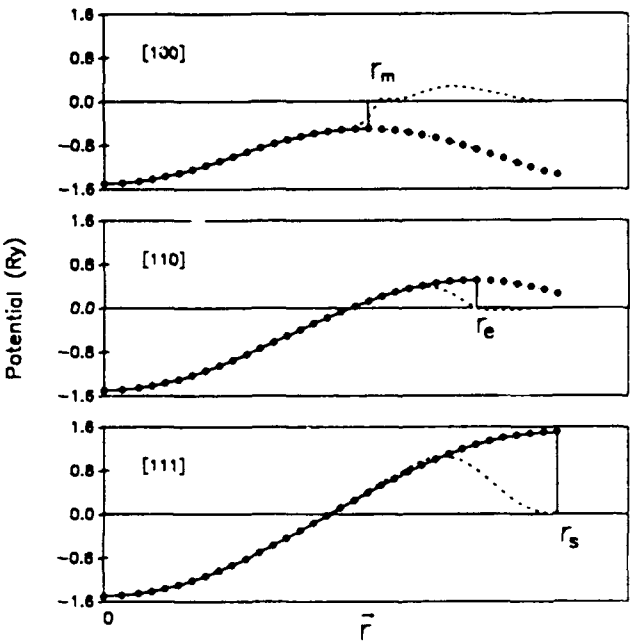


FIG. 1. Angular-momentum expansion of the Mathieu potential along [100], [110], and [111]. The solid circles represent the continuous crystal potential $V(\mathbf{r})$, and the solid lines are the truncated potential $V^T(\mathbf{r})$. The dotted and the dashed lines are the sums of the angular-momentum components up to $l=8$ for $V(\mathbf{r})$ and $V^T(\mathbf{r})$, respectively. r_m, r_e, and r_s are the distances between the origin and the face, edge, and corner of the cube, respectively. Notice that the dotted line and the solid circles are not distinguishable in the figure.

interstitial region. For example, with U_0 set to be 0.5 Ry, the MT constant potential is $V_c = U_0[9/\pi(6-\pi)] \simeq 0.501\,116$ Ry, while the actual value of the potential varies from -0.5 Ry at $(\frac{1}{2},0,0)\alpha$ to 1.5 Ry at $(\frac{1}{2},\frac{1}{2},\frac{1}{2})\alpha$.

When the full potential in Eq. (27) is expanded in cubic harmonics $V(\mathbf{r}) = \sum_L V_L(r) K_L(\mathbf{r})$, $V_L(r)$ is proportional to $-U_0 j_l(2\pi r/a)$, and the series converges very fast. With an $l_{max} = 8$, one can achieve a converged $V(\mathbf{r})$, as shown in Fig. 1. However, in the expansion for the truncated potential, $V^T(\mathbf{r}) = V(\mathbf{r})\sigma(\mathbf{r}) = \sum_L V_L^T(r) K_L(\mathbf{r})$, the components

$$V_L^T(r) = \int K_L(\mathbf{r}) V(\mathbf{r}) \sigma(\mathbf{r}) d\mathbf{r} \tag{28}$$

have to be carried out numerically with great care. Because of sharp edges and corners in $V^T(\mathbf{r})$, the angular-momentum expansion is only slowly converging. This is evident in Fig. 1, which shows sizable errors made in all three directions [100], [110], and [111] in the expansion of V^T up to $l_{max} = 8$.

We have carried out the FP-KKR calculation using the wave-function expansion in Eqs. (10) and (24) up to $l_{max} = 4$. The basis sets are calculated using both the BC and WM approaches with the potentials expanded up to $l_{max} = 8$. Results from the MT-KKR approximation are also obtained for comparison.

In Fig. 2(a), the solid lines represent the "exact" band structures for the Mathieu potential. The dots are the MT-KKR results. Despite the crude approximation in the MT potential, the lowest band is still reasonable. The MT approximation becomes worse at the higher energies,

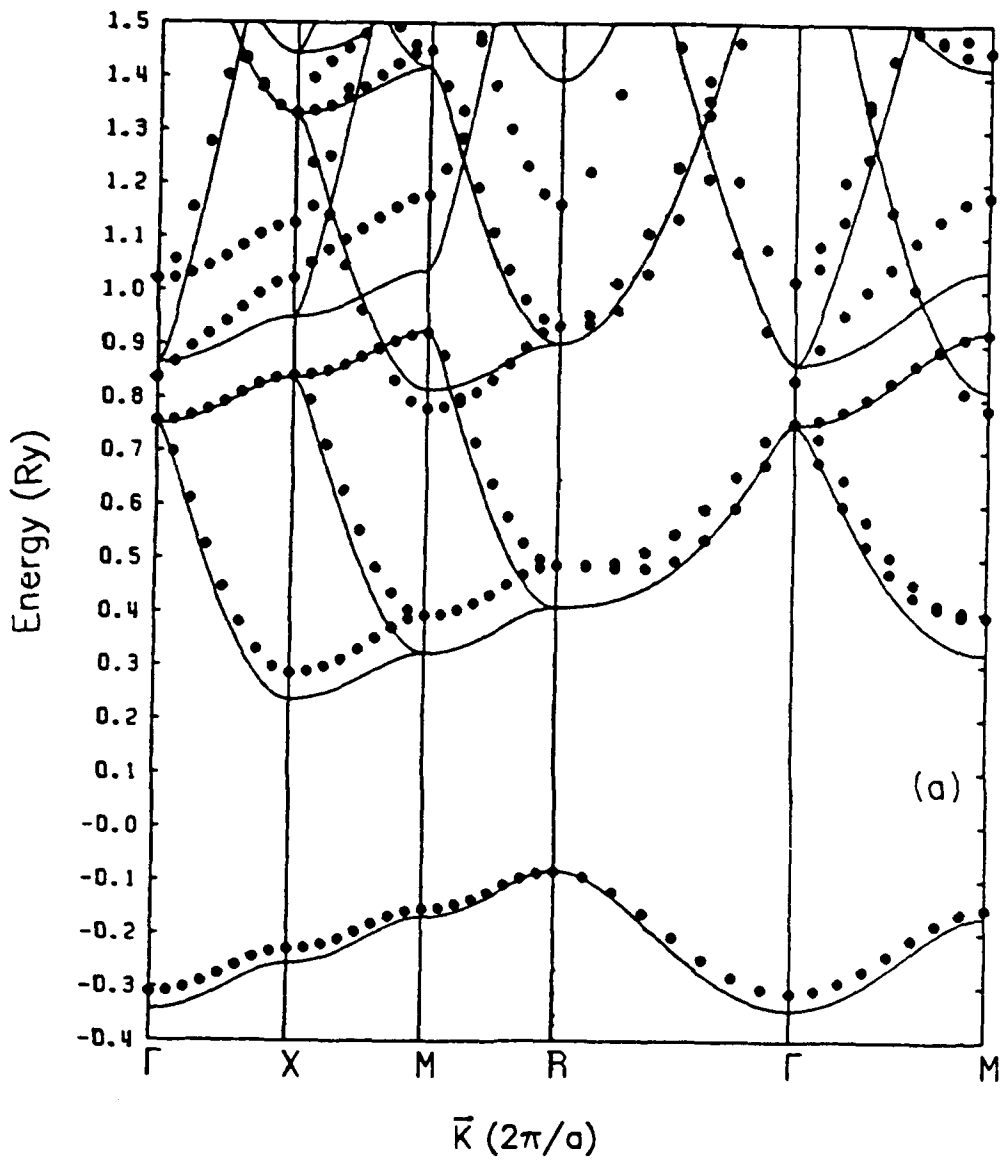


FIG. 2. Comparison of (a) the muffin-tin KKR and (b) the FP-KKR band structures (the dots) with the exact results (the solid lines) for the Mathieu potential. The symmetry points Γ, X, M, and R correspond to the wave vector at (0,0,0), $(\frac{1}{2},0,0)$, $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, and $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, respectively, in units of $2\pi/a$.

as indicated by the large energy deviations and splittings of the levels. For example, the "exact" bands from Γ to X from Γ to M around 0.9 Ry are degenerate due to separability of the Mathieu potential, while the MT approximation lifts this "accidental" degeneracy.

The full-potential KKR results are compared with the "exact" band structure in Fig. 2(b). The dots are now the FP-KKR results and are calculated based on the BC approach. The agreement is excellent and rather uniform up to 1.3 Ry. The calculation even preserves the accidental degeneracy at Γ at energy 0.88 Ry. The lowest band has a detectable deviation of 0.016 Ry at R, but has very small root-mean-square (rms) deviation. The deviations at R and some other energy states are probably due to the truncation in the angular-momentum expansions. The FP-KKR bands based on the WM approach are not noticeably different from those based on the BC approach plotted in Fig. 2(b). However, there are slight differences between the results of the two approaches. For reference, we list the deviations of both the BC and WM approaches and the "exact" energies in Table I. While the WM approach gives a larger deviation in the lowest-energy band around R, the overall rms deviations of these two approaches are similarly small. These results imply some freedom in the choice of basis functions. Provided that reasonable approximations are made in the representation of the cell potential V^T and in the calculation of $S_{LL'}$ and $C_{LL'}$ from Eqs. (21) and (22), it appears that the FP-KKR equation will give reasonable bands independent of the exact algorithm for obtaining the Φ_L, e.g., from $V(\mathbf{r})$, $V^T(\mathbf{r})$, or other smooth potentials augmented to V^T.

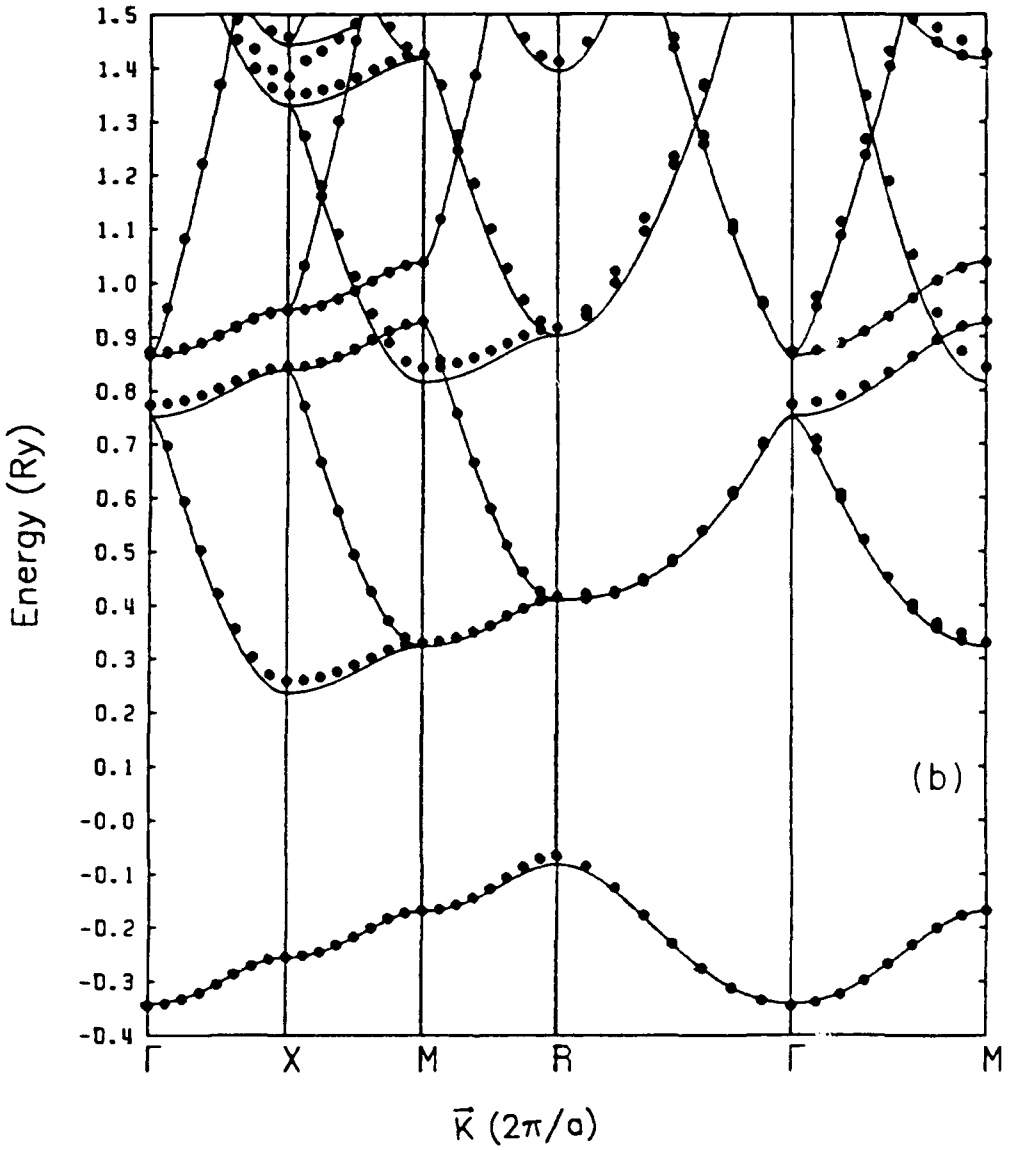


FIG. 2. (*Continued*).

TABLE I. Deviations of the FP-KKR band energies ΔE from the exact values E_{exact} for the Mathieu potential at several symmetry points. The subscripts WM and BC stand, respectively, for the Williams-Morgan and Brown-Ciftan approaches described in the text. All energies are in Ry.

Symmetry states	ΔE_{WM}	ΔE_{BC}	E_{exact}
Γ_1	−0.0053	−0.0034	−0.3414
Γ_{15}	0.0243	0.0223	0.7517
Γ_1	−0.0015	0.0024	0.8653
Γ_{12}	0.0034	0.0069	0.8653
R_1	0.0274	0.0160	−0.0827
R_{15}	0.0089	0.0064	0.4097
R_{25}	0.0156	0.0145	0.9020
R_2	0.0184	0.0171	1.3943
X_1	0.0024	0.0005	−0.2551
X_4	0.0214	0.0216	0.2372
X_5	0.0060	0.0056	0.8379
X_1	−0.0010	−0.0021	0.9515
X_2	−0.0047	−0.0036	0.9515
M_1	0.0016	0.0002	−0.1689
M_5	0.0072	0.0067	0.3234
M_3	0.0270	0.0259	0.8158
M_4	0.0033	0.0037	0.9241
M_1	0.0058	−0.0005	1.0377

V. SUMMARY

The main purpose of this paper is to test the accuracy of the FP-KKR theory in band-structure calculations. To help eliminate doubts about this theory, we have derived the FP-KKR equation explicitly from the Kohn-Rostoker integral equation.[2] This FP-KKR theory still preserves the clear separation between the structural and potential information possessed in the MT-KKR equation. The potential information is contained in the S and C matrices, which can be easily calculated if the basis functions and potentials are expressed in angular-momentum expansions. Such expansions are desirable in a realistic calculation. The whole procedure has been tested against the exactly soluble Mathieu potential in the simple-cubic structure. Because of the openness of the structure and the high anisotropy of the potential, this potential provides a challenging model to test against any band-structure theory. Our results show that with wave functions expanded up to $l_{max}=4$ and the potential up to $l_{max}=8$, the FP-KKR theory as described above gives excellent results for the bands in the energy range needed for solid-state applications. With this method, one should be able to deal with solids having open structures, such as semiconductors, for which MT-KKR is not suitable.

ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with Dr. B. Segall, Dr. J. S. Faulkner, and Dr. A Gonis. The research at Auburn University was supported in part by Office of Naval Research (ONR) Contract No. N00014-88-C0096 and Air Force Office of Scientific Research (AFOSR) Contract No. F49620-88-K-0009. The calculation was done by using the Alabama Supercomputer Network. The work at Oak Ridge National Laboratory was sponsored by the U.S. Department of Energy Division of Materials Sciences, Office of Basic Energy Sciences through Contract No. DE-AC05-8400R-21400 with Martin Marietta Energy Systems, Inc.

APPENDIX: GREEN-FUNCTION EXPANSION

Here we want to show that Eq. (6) is valid when Eq. (7) is satisfied. Following Kohn and Rostoker,[2] we separate $G_{\mathbf{k}}$ of Eq. (5) into two parts,

$$G_{\mathbf{k}}(E;\mathbf{r},\mathbf{r}')=g_0(\mathbf{k},E;\mathbf{r},\mathbf{r}')+g_1(\mathbf{k},E;\mathbf{r},\mathbf{r}') , \tag{A1}$$

where g_0 is the singular part,

$$g_0(\mathbf{k},E;\mathbf{r},\mathbf{r}')=-\frac{1}{4\pi}\frac{\exp(i\kappa|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} , \tag{A2}$$

and

$$g_1(\mathbf{k},E;\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\sum_{s\neq 0}\frac{\exp(i\kappa|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|)}{|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|}\exp(i\mathbf{k}\cdot\mathbf{R}_s) . \tag{A3}$$

For $r<r'<R_s$ and for $\mathbf{r}$ and $\mathbf{r}'$ inside τ, the first part has the expansion $g_0=-i\kappa\sum_L J_L(\kappa r)H_L^{\dagger}(\kappa r')$, where $H_L^{\dagger}(\kappa r)=J_L(\kappa r)+iN_L(\kappa r)$. Under the same condition for $\mathbf{r}$ and $\mathbf{r}'$, $|\mathbf{r}|<|\mathbf{r}'+\mathbf{R}_s|$ also holds for a Wigner-Seitz cell, so that the following expansion is valid:[6,8]

$$\frac{\exp(i\kappa|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|)}{|\mathbf{r}-\mathbf{r}'-\mathbf{R}_s|} = -i\kappa\sum_L J_L(\kappa r)H_L^{\dagger}(\kappa(\mathbf{r}'+\mathbf{R}_s)) = -i\kappa\sum_L\left[\sum_{L'}J_L(\kappa r)\mathcal{H}_{LL'}^{\dagger}(\kappa\mathbf{R}_s)J_{L'}(\kappa r')\right] , \tag{A4}$$

where

$$\mathcal{H}_{LL'}^{\dagger}(\kappa\mathbf{R}_s)=4\pi\sum_{L_2}i^{(l'+l_2-l)}C_{L,L'}^{L_2}H_{L_2}^{\dagger}(\kappa\mathbf{R}_s) , \tag{A5}$$

where $C_{L,L'}^{L_2}$ is given in Eq. (25). Therefore the Green function has the expansion in Eq. (6) with the structure constant given by

$$B_{LL'}(\mathbf{k},E)=-i\kappa\left[\delta_{LL'}+i^{(l'-l)}\sum_{s\neq 0}\mathcal{H}_{LL'}^{\dagger}(\kappa\mathbf{R}_s)\exp(i\mathbf{k}\cdot\mathbf{R}_s)\right] . \tag{A6}$$

[1]J. Korringa, Physica **13**, 392 (1947).
[2]W. Kohn and N. Rostoker, Phys. Rev. **94**, 1111 (1954).
[3]A. R. Willaims and J. van W. Morgan, J. Phys. C **7**, 37 (1974).
[4]J. S. Faulkner, Phys. Rev. B **19**, 6186 (1979).
[5]J. S. Faulkner and T. P. Beaulac, Phys. Rev. B **26**, 1597 (1982).
[6]R. G. Brown and M. Ciftan, Phys. Rev. B **27**, 4564 (1983).
[7]A. Gonis, Phys. Rev. B **33**, 5914 (1986).
[8]R. Zeller, J. Phys. C **20**, 2347 (1987).
[9]A. Gonis, X.-G. Zhang, and D. M. Nicholson, Phys. Rev. B **38**, 3564 (1988); **40**, 947 (1989).
[10]X.-G. Zhang and A. Gonis, Phys. Rev. B **39**, 10373 (1989).
[11]R. K. Nesbet, Phys. Rev. B **41**, 4948 (1990).
[12]E. Badralexe and A. J. Freeman, Phys. Rev. B **36**, 1378 (1988).
[13]F. S. Ham and B. Segall, Phys. Rev. **124**, 1786 (1961).
[14]B. Segall, J. Phys. Chem. Solids **8**, 371 (1959).
[15]P. M. Morse, Phys. Rev. **35**, 1310 (1930); *Tables Relating to Mathieu Functions* (Colmbia University, New York, 1951).
[16]J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1965), Vol. 2, pp. 158-166.
[17]D. M. Nicholson and J. S. Faulkner, Phys. Rev. B **39**, 8187 (1989).
[18]R. Zeller, Phys. Rev. B **38**, 5993 (1988).
[19]J. S. Faulkner, Phys. Rev. B **32**, 1339 (1985).
[20]R. G. Brown and M. Ciftan, Phys. Rev. B **39**, 3543 (1989).
[21]Reference 11 showed that the two procedures led to the same basis functions Φ_L.

Semicond. Sci. Technol. 5 (1990) S86–S89. Printed in the UK

Vacancies and surface segregation in HgCdTe and HgZnTe

M A Berding†, A Sher†, A-B Chen‡ and R Patrick‡

† SRI International, Menlo Park, California, USA
‡ Auburn University, Auburn, Alabama, USA

Received in accepted form 18 September 1989

Abstract. Vacancies are known to play an important role in the Hg-based narrow-gap alloys HgCdTe and HgZnTe. In this paper we summarise our recent calculations of the vacancy formation energies in HgCdTe and HgZnTe. We find that the vacancy formation energy in these alloys varies non-linearly with alloy concentration, resulting in higher vacancy densities than those predicted on the basis of a linear variation of the vacancy energies. Surface segregation in these alloys is driven by bond strength and bond length differences. We review our recent calculations which show that the chemical terms dominate in HgCdTe and result in Hg-rich surfaces, while in HgZnTe the strain terms contribute as well and result in a less Hg-rich surface, suggesting that HgZnTe surfaces may be more amenable to surface processes such as passivation.

1. Introduction

The non-idealities of semiconductors always control device properties; this is true to an exceptional degree in the Hg-based narrow-gap alloys (HgCdTe and HgZnTe). In this paper we review our recent work on two of these important non-idealities, namely vacancies and surface segregation. *Vacancies* in HgCdTe are responsible for the intrinsic doping in the material and their abundance has been attributed to the weak HgTe bond. Sher *et al* (1985) have suggested that HgZnTe may be a better candidate than HgCdTe for infrared device applications on the basis of calculations which showed the HgTe bond to be stronger in HgZnTe than in HgCdTe, implying a lower tendency for Hg vacancy formation. Below we review our calculations (Berding *et al* 1987, Berding, Chen and Sher, unpublished) of the vacancy formation energies in these two alloys. The surfaces of these alloys play an important role in the various device-processing steps. *Surface segregation* in metal alloys is a well known phenomenon. Measurements (Buck 1982) have shown that the composition of the surface may differ from that of the bulk, and such segregation is also expected to occur in the semiconductor alloys. The same forces that drive vacancy formation in these alloys also drive surface segregation, namely the relative strengths of the HgTe, CdTe and ZnTe bonds in the various alloy (and surface) environments. In this paper we review our recent work in this area (Patrick *et al* 1989).

2. Vacancies

The calculation of the properties of vacancies in semiconductors is a difficult problem to which substantial effort has been directed with various degrees of success. Most of these calculations have focused on the localised electronic levels. In HgCdTe and HgZnTe we are interested in the relative ease with which the vacancies form in the materials, and how the vacancy concentration varies across the composition range. In particular, we want to know whether, for an alloy concentration corresponding to a given band gap, vacancies will be more of a problem in HgZnTe or in HgCdTe. To answer this question, we have developed a tight-binding model of the vacancies to calculate the vacancy formation energies E_v in the semiconductor compounds and alloys. This model is summarised below.

For the compound semiconductors a Green function method can be used to treat the vacancy as a point defect in the otherwise perfect lattice. In the alloys the lattice is not perfect and is in general disordered owing to the presence of the two atom species located on the cation sublattice. Because of this disorder, an embedded cluster model of the vacancy is preferable for the alloys, so that the near-alloy environment of the vacancy can be modelled explicitly while the far environment can be treated approximately (e.g. in the virtual crystal approximation or VCA).

We have calculated E_v for ZnTe, CdTe and HgTe

0268-1242/90/030S86 + 04 $03.50

using a tight-binding model based on Harrison's (1981) parameters with the repulsive interaction energy selected to give correct agreement with the experimental cohesive energies in the compounds. The final state for the removed atom in free space has been used as a reference, although various other final states can also be calculated. Atom clusters with up to second-, third- and fourth-neighbour shells of atoms were used, and each was coupled to an extended bulk using perturbation theory at the cluster boundary. Clusters are centred at the vacancy formation site, and E_v is calculated from a difference in total energies of the cluster before and after vacancy formation.

The Green function method was also used to calculate total energies and E_v for the compounds, using the same Hamiltonian. For the cation vacancy the cluster calculations converge very quickly to the Green function (effectively an infinite-cluster result). For the anion vacancy the convergence is not as fast, an effect attributed to the more extended nature of the defect states produced by the anion vacancy. The relative magnitude of the anion or cation vacancy formation energies for most compounds is not changed from the cluster to the Green function result, so comparisons among the various compounds should not change. Thus we can use the clusters to calculate E_v in the alloys, with a fixed correction across the composition range to account for the cluster truncation.

The vacancy formation energies in the alloys are calculated in a similar manner, with differences of cluster total energies with and without a vacancy used to calculate E_v. The alloy environment for the first- and second-nearest neighbours is treated explicitly, with vacancy formation energies being calculated for the various particular arrangements of cations. For example, for the Te vacancy in HgCdTe, E_v for five particular near-neighbour environments of the Te are calculated: Hg_0Cd_4, Hg_1Cd_3, Hg_2Cd_2, Hg_3Cd_1 and Hg_4Cd_0. For the cation vacancy, E_v for the various cation arrangements in the second-nearest-neighbour shell must be considered. For both cation and anion vacancies the alloy environment in the third shell and beyond is modelled in the VCA.

Results for E_v against the first- or second-nearest-neighbour composition are shown in figure 1. The effects of the alloy environment are most dramatic for the anion vacancy, as expected, because the cation substitution occurs in the first-neighbour shell about the vacancy. A minimum occurs in E_v for Te in HgCdTe and HgZnTe and arises from the occupation of the defect states, which in this one-electron picture lie at higher energies in ZnTe and CdTe than in HgTe. Thus in clusters containing at least one Hg in the first-neighbour shell the defect electrons will occupy the lower-lying Hg-like defect state.

The effect of the alloy environment is less for the cation than for the anion vacancy. For all cations, E_v is found to decrease as one goes to Hg-rich clusters in HgCdTe and HgZnTe. Earlier calculations (Sher *et al* 1985) based on bond strength modifications in the alloys predicted that Hg vacancies will occur more readily in HgCdTe than in HgTe, in contrast to the present predictions. The current results show that the simple bulk bond strength arguments are insufficient to predict trends in E_v, because they do not include the effects of back-bond strength modifications due to the vacancy. Thus, although the HgTe bond may be weakened in the alloy, when an Hg vacancy is formed, adjacent bonds strengthen, modifying E_v in the alloy.

The vacancy concentrations in the random alloys have been calculated from the above results and are shown in figure 2. Shown for comparison are the vacancy concentrations we get if we assume a linear variation of E_v with near-neighbour environment, i.e. a straight-line variation between the end points in figure 1. For any given alloy concentration x, an appropriate probability-weighted average over the clusters is taken to calculate the results in figure 2. From figure 2 one sees that the effects of the non-linearity of E_v on vacancy concentrations are quite dramatic, even for the cation vacancy. In particular, our results predict higher concentrations of vacancies in the alloys than one would expect on the basis of simple arguments (i.e. a linear interpolation of E_v from the compounds). We emphasise here the overall behaviour observed in figures 1 and 2, because the magnitude of E_v and the vacancy concentrations will depend on the final states available to the atoms forming the vacancies.

Before concluding, we note that vacancy formation energies for final states of the cations on the (111)A surface of the compound can be estimated if one approximates the (111)A localised states before and after an A atom is added to it, by similar defect states at the anion and cation vacancy sites. A difference of the vacancy formation energies for the free atom final state and the (111)A surface final state yields an estimate of the sublimation energy of an A atom from the (111)A surface. These numbers are related to the energies necessary to calculate the surface segregation in HgCdTe and HgZnTe.

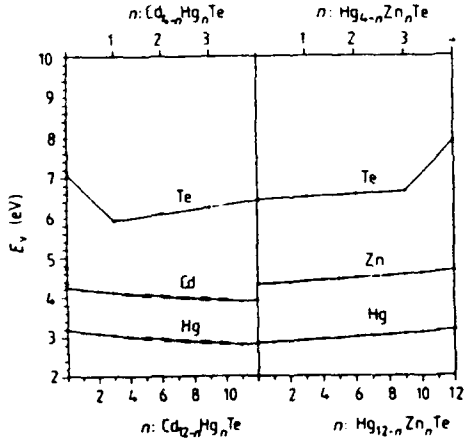


Figure 1. Extraction energies in the Te common-anion alloys as a function of near-neighbour configuration for bulk (VCA) concentration $x = 0.5$. The full curves correspond to the average extraction energies and the broken curves to the RMS deviations.

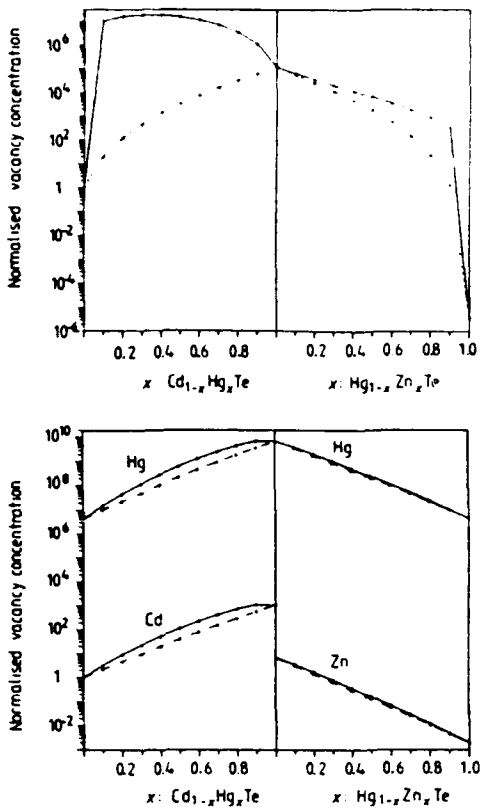


Figure 2. Relative vacancy concentrations in HgCdTe and HgZnTe. The full curves correspond to results using E_v in figure 1; the broken curves are based on a linear interpolation from E_v for the compounds.

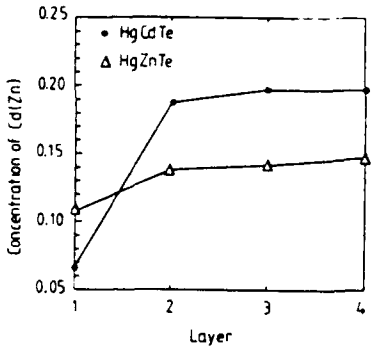


Figure 3. Layer concentrations at the growth temperature of 973 K for $Hg_{0.80}Cd_{0.20}Te$ and $Hg_{0.85}Zn_{0.15}Te$.

3. Surface segregation

Two major contributions drive surface segregation in the semiconductor alloys. Chemical energies arise from differences in bond-breaking energies of the constituent compounds. This can drive surface segregation, because fewer bonds are made by atoms at the surface than in the bulk, and one expects the less-well-bound species (e.g. Hg) to segregate to the surface. In lattice-mismatched alloys, strain energies can also drive surface segregation, because the removal of a mismatched atom to the surface results in a lowering of the strain energy in the bulk. The strain energy release would be expected to be the largest for the dilute species segregating to the surface. In lattice-matched HgCdTe, only the chemical energies contribute significantly to the segregation, while in HgZnTe both strain and chemical energies contribute.

The surface concentration profile is calculated (Patrick *et al* 1989) by equating the chemical potential (defined as the first derivative of the free energy F with respect to concentration) in each layer, because at equilibrium the chemical potential in each layer must be equal. The layers are coupled to one another through the configuration entropy contribution to F. Calculations were done within the quasi-chemical approximation (Kumar *et al* 1979, Kumar 1981) in which pair interactions only are included, and the regular solution model (Williams and Nason 1974) in which randomness is assumed within each layer.

Surface segregation profiles were calculated for the (111)A surface of HgCdTe and HgZnTe. Results corresponding to the equilibrium for the ideal (111)A surfaces at 973 K are shown in figure 3. Results shown are for bulk concentrations of $Hg_{0.85}Zn_{0.15}Te$ and $Hg_{0.80}Cd_{0.20}Te$ corresponding to the same band gap. A larger surface segregation is found in HgCdTe than in HgZnTe, although for both materials at this high temperature the concentration decays to the bulk value within four atom layers (a layer consisting of an anion and cation pair of atom planes). For both alloys the surfaces are found to be Hg-rich, as expected on simple chemical bond strength arguments. In HgZnTe the strain contribution tends to drive Zn, the minority species, to the surface. Thus in HgCdTe the chemical energies alone drive Hg to the surfaces, while in HgZnTe the chemical and strain energies oppose one another, lessening the magnitude of the Hg surface concentration enhancement.

Calculations are in progress for the surface segregation on the (110) cleavage plane and the (100) surface, which is important in epitaxial growth. On the basis of bond strength arguments alone, one would expect the surface segregation to be larger on the (100) growth surface than on the (111)A, because of the fewer bonds made by the surface atoms (the cations on the (111)A surface make three bonds, while on the (100) only two surface bonds are made).

4. Conclusions

The above results demonstrate some of the complexities encountered when dealing with the semiconductor alloys. The vacancy formation energies are found to exhibit a non-linear variation with alloy concentration, resulting in larger vacancy concentrations in the alloys than otherwise expected. Similar effects in other II-VI and III-V alloys have been predicted (Berding, Chen and Sher, unpublished). Surface segregation has been found

to a greater extent for the (111)A surface in HgCdTe than in HgZnTe, indicating that HgZnTe may pose fewer problems in the surface-related properties and processing steps. Results from calculations currently in progress on the (110) and (100) surfaces, and including a vapour phase, will more fully answer the question of surface segregation on technologically important surfaces.

Acknowledgments

This work has been supported by the National Aeronautics and Space Administration under Contract NAS1-18226, by the Office of Naval Research under Contract N00014-85-K-0448 and by the Air Force Office of Scientific Research under Contract F49620-88-K-0009.

References

Berding M A, Chen A-B and Sher A 1990 *Phys. Rev.* B submitted for publication.
Berding M A, Krishnamurthy S, Sher A and Chen A-B 1987 *J. Vac. Sci. Technol.* A **5** 3014
Buck T M 1982 *Chemistry and Physics of Solid Surfaces IV* ed. R Vanselow and R Howe (Berlin: Springer)
Harrison W 1981 *Phys. Rev.* B **24** 5835
Kumar V 1981 *Phys. Rev.* B **23** 3756
Kumar V, Kumar D and Joshi S K 1979 *Phys. Rev.* B **19** 1954
Patrick R, Chen A-B, Sher A and Berding M A 1989 *Phys. Rev.* B **39** 5980
Sher A, Chen A-B, Spicer W E and Shih C K 1985 *J. Vac. Sci. Technol.* A **3** 105
Williams F L and Nason D 1974 *Surf. Sci.* **45** 377

Vacancy formation and extraction energies in semiconductor compounds and alloys

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

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

(Received 7 May 1990; accepted for publication 2 August 1990)

Extraction energies for diamond and zinc-blende semiconductor compounds and pseudobinary alloys are calculated using a tight-binding cluster method, where the final state of the removed atom is in a free-atom state. The extraction energies provide a convenient reference from which other final states of the removed atoms can be calculated. In the elemental and compound semiconductors, the convergence of the cluster calculation was verified using a Green's function calculation with the same Hamiltonian. For the elemental semiconductors, vacancy (or Schottky defect) formation energies, in which the final state of the removed atom is on the surface, have been calculated. For pseudobinary alloys of the form $A_{1-x}B_xC$, we find extraction energies to be very sensitive to the local environment, exhibiting a nonlinear variation between the A- and B-rich local environments; the nonlinearity is especially pronounced for the removal of a C atom. Nonlinearities are found to arise primarily from the occupation of localized vacancy states. The impact that these alloy variations will have on measurable properties are discussed.

I. INTRODUCTION

Knowing the ease with which vacancies form in semiconductor compounds and alloys is essential to understanding many properties of these materials and the way they respond to device processing. For example, the deep levels often associated with vacancies can be detrimental to device performance. Unfortunately, limited experimental values exist for the vacancy formation energies in semiconductors, and little effort has been directed toward understanding the variations of the vacancy formation energies in the alloys. Interpretation of such experiments is complicated because, in addition to the various types of vacancies possible (anion, cation, and, in alloys, particular environments for the anion or cation), the presence of other native defects, such as interstitials, antisites, and impurities, also contribute to the measured quantities such as carrier concentrations and diffusion. Thus vacancies are rarely truly isolated defects in crystals and are influenced by the presence of other defects and the compound stoichiometry. In addition, diffusion measurements which have been used to deduce the vacancy formation energy are complicated by contributions from the migration energy, as well as by interstitial diffusion and possible diffusion paths along extended defects, such as dislocations. Because of this, the magnitude of vacancy formation energies in many semiconductor compounds is not well established experimentally.

Many efforts have been directed at understanding the important electronic properties of vacancies and deep states in semiconductors.[1-4] Most of these calculations have focused on the localized states produced by the vacancy using cluster and slab methods.[5-11] Fewer models have been used to calculate formation in semiconductors.[11-16]

In this paper we present a model for the calculation of the extraction and vacancy formation energies in diamond-cubic and zinc-blende semiconductors. We use a cluster method based on Harrison's tight-binding theory.[17] Harrison's theory has been found to give semiquantitative agreement with experiment, yielding proper trends in structure-related properties of pure semiconductor compounds and their alloys.The elimination of surfacelike states associated with the cluster boundary have been addressed by including only complete bonds in the cluster, thereby allowing atoms with one or two hybrid orbitals missing at the cluster edge, and then coupling the cluster states to bond orbitals outside of the cluster by using second-order perturbation theory. We have included in this study (1) a verification of the convergence with cluster size using a Green's function method; (2) an application to both the anions and cations in the compound semiconductors; (3) a systematic comparison among the various group-IV, -III-V, and -II-VI semiconductors; and (4) an extension to the pseudobinary alloys to study the explicit effect of alloy composition, by directly considering the various possible clusters of atoms about the vacancy site. Such a detailed examination of the dependence of the extraction energy on the alloy environment cannot be easily obtained in a Green's function method using an effective-medium theory. Because of the difficulty associated with correctly predicting the energy positions of the deep levels in the gap, no attempt is made to do so here. Additionally, only the zero-temperature, ground-state energies have been calculated. Preliminary results of this work have been published previously.[18-21]

The remainder of this paper is organized as follows. In Sec. II we briefly discuss the method used to calculate the extraction and vacancy formation energies in the diamond and zinc-blende semiconductors. The extension of this method to the alloys is given in Sec. III. Results for both compounds and alloys are presented in Sec. IV, and concluding remarks are presented in Sec. V.

 0021-8979/90/225064-13$03.00

II. EXTRACTION AND VACANCY FORMATION ENERGIES IN SEMICONDUCTORS

The method we use is a tight-binding model, based on Harrison's universal parameters[22] and term values taken from Chen and Sher.[23] We begin by calculating the extraction energy, defined as the energy necessary to remove an atom from a bulk lattice site to a free-atom state, leaving behind an isolated bulk vacancy. The vacancy is described by using a cluster of atoms surrounding the vacancy site. A sp^3 hybrid basis is used, and atoms at the edge of the cluster are truncated so that no dangling bonds are included. The bonds at the cluster edge are coupled to the infinite crystal using second-order perturbation theory. Cation- or anion-centered clusters are used for the calculation of their respective extraction energies. Total electronic energies of a cluster, with and without the center atom removed, are calculated from a sum of one-electron energies, obtained from the cluster diagonalization. The extraction energy is then calculated from

$$E_x = (E_f - E_i), \tag{1}$$

where E_f is the total energy of the cluster with the center atom removed plus the removed atom in a free-atom state, and E_i is the total energy of the cluster before the removal of the central atom. The cluster total energies include a contribution from the hybrid overlap interaction in each bond: this overlap energy V_{ij}, is chosen for each material so that energy per bond for the initial cluster agrees with experiment. Convergence of the extraction energies with cluster size was verified using a Green's function calculation,[24] in which the vacancy is simulated by setting the term values at the vacancy site to infinity.

Several additional contributions to E_x were also estimated. First, the charge redistribution that occurs when an atom is removed from the cluster produces a shift in both the Madelung energy ΔK and the on-site Coulomb energy ΔU. The calculation of these energies is discussed in Appendix A. Second, the atoms about the vacancy site may move from their ideal lattice positions as a result of the formation of a vacancy, relaxing into the minimum energy configuration. The calculation of the energy resulting from radial relaxation about the vacancy site is given in Appendix B. Finally, the second-neighbor interaction of the dangling hybrids at the vacancy site and the related Jahn–Teller distortion can lower the vacancy formation energy. An estimate of this correction is given in Appendix C.

Vacancy formation or Schottky defect formation energies are calculated from, in the elemental semiconductors,

$$E_v = E_x - E_{\rm coh}, \tag{2}$$

or in compound semiconductors as

$$E_v = \tfrac{1}{2}(E^a_x + E^c_x) - E_{\rm coh}, \tag{3}$$

where $E_{\rm coh}$ is the bulk cohesive energy per atom. Equation (2) corresponds to the removal of an atom from a bulk site to a surface, leaving a bulk vacancy behind, and thereby increasing the number of sites of the solid by one while leaving the nature of the surface effectively unchanged. For the compound semiconductors, Eq. (3) corresponds to the creation of an unbound cation-anion vacancy pair and the addition of a unit cell to the crystal, here normalized per atom. We cannot consider the surface to be unchanged after the removal of only an anion or cation from the bulk to the surface, because in doing so the surface stoichiometry, by necessity, will be changed. Another way to visualize this is to realize that the cohesive energy appearing in Eq. (2) can only be defined per cation-anion pair in a compound, and the cohesive energy per anion or per cation alone cannot be generally defined. Thus, in compound semiconductors, additional information, such as stoichiometry and other dominant defects, must be supplied to calculate the vacancy concentration. Hence, while energies $E^{a(c)}_v = E^{a(c)}_x - E_{\rm coh}$ can be algebraically defined, they are not generally physically meaningful in the sense that they correspond to a specific event, nor will they alone in statistical equations predict measurable quantities. For similar reasons, the Schottky defect formation energies in the alloys are not rigorously defined.

III. EXTRACTION ENERGIES IN ALLOYS

Extraction energies in the pseudobinary semiconductor alloys are calculated in a manner similar to the elemental and compound semiconductors. Beyond the cation and anion, there are two distinct classes of vacancies which we must consider: the removal of an atom of the common species and the removal of an atom of the substituted species. Below, we briefly discuss the calculations for these two classes.

We begin by discussing the extraction energy for an atom of the common species, i.e., a C atom in the pseudobinary alloy $A_{1-x}B_xC$, where C can be either a cation or anion. Consider the four first-neighbor positions occupied by $(4-n)$ A atoms and n B atoms, where $n = 0$–4. Because of the tetrahedral symmetry, there is only one unique arrangement for the $(4-n)$ A and n B atoms. Because the number of possible arrangements of A and B atoms in the third-shell neighbor sites and beyond is large, and because we expect the primary effects of alloying on the extraction energy to have a short range, we use a virtual crystal average (VCA) for the medium beyond the second shell.

For an alloy of a given concentration x, the average bond lengths of the AC and BC bonds in the alloys have been shown to be well represented by[25]

$$d_{AC} = d^0_{AC} + x(d^0_{BC} - d^0_{AC})/4, \tag{4a}$$

and

$$d_{BC} = d^0_{BC} - (1-x)(d^0_{BC} - d^0_{AC})/4, \tag{4b}$$

where d^0_{AC} and d^0_{BC} are the bond lengths of the constituent compounds. This is also in excellent agreement with extended x-ray absorption fine-structure (EXAFS) experiment. Because in a random alloy the deviation of the bond lengths in various classes of clusters about this value is small,[26] we have used bond lengths calculated from Eq. (4) for all classes of clusters (i.e., values of n). Because the breathing-mode lattice relaxation about the vacancy site was found to be small for elemental and compound semiconductors (see Sec. IV), lattice relaxation about the vacancy in the alloys has been ignored.

For a given bulk concentration x, the extraction energy

of a C atom from a $A_{4-n}B_n$ cluster, $E_x^C(x,n)$, is calculated by taking the difference between the final- and initial-state cluster energies with the same configuration of A and B atoms. In a random alloy $A_{1-x}B_xC$, the average extraction energy of a C atom is given by

$$E_x^C(x) = \sum_{n=0}^{4} P_n E_x^C(x,n), \tag{5}$$

where P_n is the random probability of a cluster with n B atoms, given by

$$P_n = p_n x^n (1-x)^{4-n}, \tag{6}$$

and where p_n is the binomial coefficient

$$p_n = \binom{4}{n}. \tag{7}$$

The above expression for $E_x^C(x)$ can be readily extended to include nonrandom cluster distributions by replacing P_n with the appropriate nonrandom probabilities. Note that it is the particular values of $E_x(x,n)$ which enter into the calculation of many physical properties, not the average calculated in Eq. (5), as shall be discussed below.

We now consider a B vacancy in the alloy $A_{1-x}B_xC$. To include specific alloy arrangements, one must consider the G_n unique ways of arranging $(12-n)$ A atoms and n B atoms in the second-nearest-neighbor positions, where $n = 0$–12. Because of couplings to other second-neighbor sites through the third-neighbor atoms, all second-neighbor sites are not equivalent, and in general G_n is large. Thus we approximate by assuming the cluster is a Bethe lattice, thereby decreasing the connectivity of the lattice and subsequently the number of unique arrangements to G'_n. Because the number of possible arrangements of A and B atoms in the fourth-neighbor sites and beyond is large, and because the detailed arrangement of the atoms in these positions should not significantly affect the extraction energy, a VCA average beyond the third shell is used.

The extraction energy, $E_v^n(x,n,g_n)$, for each arrangement g_n of n B atoms and $(12-n)$ A atoms in the second-neighbor positions is computed by taking the difference between the initial- and final-state cluster energies. The average extraction energy for a given concentration x is expressed by

$$E_v^B(x) = \sum_{n=0}^{12} P_n \sum_{g_n=1}^{G'_n} P_{g_n} E_v^n(x,n,g_n) + E_{\text{Bethe}}(x), \tag{8}$$

where P_n is the random probability of a cluster with n A atoms, given by

$$P_n = p_n x^n (1-x)^{12-n}, \tag{9}$$

and p_n is the binomial coefficient

$$p_n = \binom{12}{n}. \tag{10}$$

P_{g_n} is the probability of the g_n^{th} particular arrangement of $(12-n)$ A atoms and n B atoms in the second-neighbor sites of the cluster. Because it is well known that the properties of solids are poorly represented by approximations such as the Bethe tree, we do not focus on the quantitative values in presenting the results for a vacancy for the alloys, but rather on the overall trends observed. A correction $E_{\text{Bethe}}(x)$ is added to the values calculated using the Bethe approximation to produce agreement with the non-Bethe results for the compounds. The correction in the alloy is calculated by first computing it for the pure compounds, AC and BC, e.g., $E_{v,\text{Bethe}}^{AC} = E_v^{AC}(\text{zinc blende}) - E_v^{AC}(\text{Bethe lattice})$, using the same cluster size, then interpolating to calculate the correction in the alloy, via $E_{\text{Bethe}}(x) = (1-x)E_{\text{Bethe}}^{AC} + (x)E_{\text{Bethe}}^{BC}$. This correction is typically <0.3 eV. As above, these expressions can be readily extended to include nonrandom cluster distributions.

IV. RESULTS

A. Extraction energies

Extraction energies E for anions and cations in the group-IV, -III-V, and -II-VI semiconductors were calculated for cluster sizes from 13 to 59 atoms and compared with the results of a Green's function calculation using the same matrix elements. Convergence to the Green's function values for the cation vacancies was better than for the anion vacancies. This was found to be due to the differences in the nature of the states associated with the dangling hybrids at the vacancy site. The localized states at the cation vacancy site arise primarily from the anion dangling hybrids, while the states at the anion vacancy site arise primarily from the cation dangling hybrids. The anion hybrid-derived states are more valence band like and thus more localized; in contrast the cation hybrid-derived states are more conduction band like and more delocalized; thus finite cluster effects are more important. Calculated extraction energies using the Green's function are given in Table I. Also shown for comparison are

TABLE I. Extraction and vacancy formation energies for zinc-blende semiconductors. The vacancy formation energy for the compounds are normalized per atom. Also shown for comparison are the cohesive energy per atom in the ideal solids.

Compound	$E_{v,\text{sub}}$ – eV	E^{v} – eV	E^{v} – eV	E_c – eV
C	7.36	19.2	19.2	11.9
Si	4.64	10.3	10.3	5.7
Ge	3.88	8.1	8.1	4.2
Sn	3.12	6.3	6.3	3.2
AlP	4.26	10.6	11.1	6.6
GaP	3.56	8.3	8.8	5.0
InP	3.48	8.2	8.8	5.1
AlAs	3.78	8.8	9.6	5.4
GaAs	3.26	6.7	7.8	4.0
InAs	3.10	6.8	7.5	4.0
AlSb	2.40	4.7	6.1	3.0
GaSb	2.96	5.9	6.8	3.4
InSb	2.80	6.0	6.3	3.3
ZnS	3.18	6.9	10.0	5.3
CdS	2.84	6.1	8.8	4.6
HgS	2.04	4.5	7.8	4.1
ZnSe	2.58	5.7	8.7	4.6
CdSe	2.42	5.0	7.8	4.0
HgSe	1.70	3.4	6.8	3.4
ZnTe	2.40	5.1	8.1	4.2
CdTe	2.20	4.7	7.4	3.8
HgTe	1.64	3.1	6.4	3.1

the cohesive energies of the ideal solids. We see that estimated errors attributable to the cluster approximation are $\lesssim 0.1$ eV for the cation vacancies and $\lesssim 0.5$ eV for the anion vacancies. Similar errors are expected in the alloy cluster calculations.

Although the major contribution to E_v is from breaking four bonds at the vacancy site, there are additional terms that contribute to E_v and that differ for the cation and anion vacancies. These differences can be attributed to several sources. First, when an anion is removed from the solid, the electrons on the removed anion are depromoted from the sp^3 hybrids, while the electrons on the four cations adjacent to vacancy remain in excited hybrid states; when a cation is removed, it is the cation electrons that are depromoted from the hybrids, while the anion orbitals remain hybridized. Because the bond-breaking term $2E_{coh}$ is referenced to the free cation and anion, the energy of the electrons which remain promoted to sp^3 hybrids must be included explicitly. This contribution to E_v is larger for the cation vacancy because both the number of electrons and the promotion energy per electron is greater for the four anions at the vacancy site. A second difference in E_v arises from metallization-induced[17] shifts in the dangling hybrid energies. This shift is upward for the anion hybrids and downward for the cation hybrids, yielding net positive (negative) contribution to E_v for the cation (anion) vacancy. Finally, a third difference in E_v arises from back-bond energy shifts which result from the cation (anion) dangling hybrids at the anion (cation) vacancy site coupling to adjacent unbroken bonds and antibonds in the lattice. Resulting shifts from this source are large and negative for the cation vacancy because of the larger magnitude of the coupling V_1, and the deeper energy and resulting stronger coupling of the anion dangling hybrid state, with respect to the cation states at the anion vacancy site.

The contribution from relaxation and Coulomb energies to both anion and cation extraction energies are comparable and small. The nearest-neighbor atoms are found to relax away from the vacancy site when a cation is removed, while they relax toward the vacancy site when an anion vacancy is created. For both cation and anion vacancies, calculated bond-length shifts are small, with the first-neighbor bond lengths shifting by less than 2%. The breathing-mode relaxation at a vacancy site has often been assumed to be similar to the relaxation occurring at a (111) surface[27,28] where the atoms are found to move toward the bulk, shortening the bond lengths of nearest-neighbor bonds by 8%–15%. In contrast to some previous work, we find that the bond-length and energy shifts in the relaxed cluster are small compared to the accuracy of the model; inclusion of relaxation in the calculation of E_f and E_v is not important. The Jahn–Teller energy E_{JT} resulting from the direct interaction of the dangling hybrids has been estimated (see Appendix C), and energies of the order of -0.6 eV are obtained.

B. Schottky defects

From the extraction energies and Eqs. (2) and (3) above, Schottky defect formation energies were calculated and are given in Table I. As discussed above, in the compounds the Schottky defect formation energy E_f is physically meaningful only for a cation-anion (bound or unbound) vacancy pair, since the creation of a new unit cell by bringing atoms to the surface requires creation of two lattice sites. A comparison of E_f with the cohesive energy is shown in Fig. 1. We find, as expected, that the E_f increases with increasing cohesive energy, and that the group-IV, -III-V, and -II-VI compounds fall into groups that each vary nearly linearly with the cohesive energy, but that have different slopes and initial values. Also, for a given cohesive energy, E_f is larger for the more ionic materials. Compensation effects, which are expected to be largest in the wide-gap ionic materials, are not included in this calculation. Observed trends may be modified by the inclusion of this effect. Also shown in this figure is the quantity $E_x^{a(c)} - E_{coh}$ for the compound semiconductors, to illustrate the variation of extraction energy with cohesive energy. The difference in this energy between the anion and cation is larger for the more ionic, group-II-VI compound, as expected. We also find $E_x^c < E_x^a$, consistent with the observations of measured extraction energies in CdS and CdTe.[29,30]

The vacancy concentration can be found by minimizing the change in the Gibb's free energy ΔG between the pure bulk crystal and the crystal containing n vacancies. For noninteracting vacancies, the Gibbs free energy is given by

$$\Delta G = nE_f - TS^{\text{config}}(n) + F(n,T), \tag{11}$$

where T is the temperature and $S^{\text{config}}(n)$ is the configurational entropy of the n vacancies. $F(n,T)$ accounts for free-energy changes from the modification of the phonon fre-

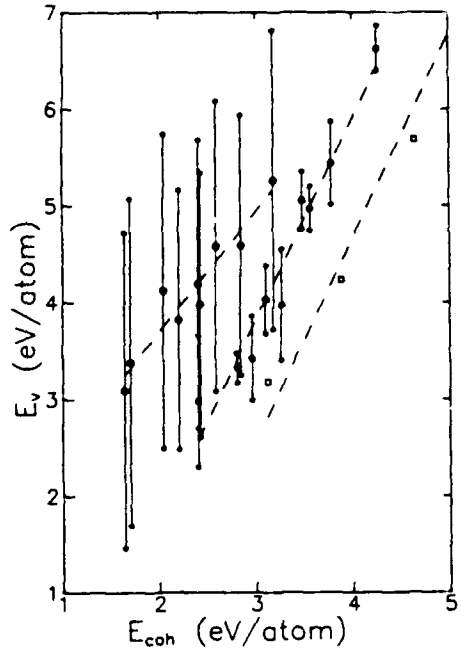


FIG. 1. Comparison of the Schottky defect formation energy (normalized per atom) and the cohesive energy. Squares are the group-IV elements, solid circles are group-III-V compounds, and open circles are group-II-VI compounds. Dashed lines represent the least-squares fit through each series of compounds, where, for the group-IV elements, carbon (not shown) has been included in the fit. Upper (lower) $\times$'s represent the quantity $E_x^{a(c)} - E_{coh}$ for a given compound.

quencies because of the presence of n vacancies, as well as additional free-energy changes, such as entropy associated with the Jahn–Teller distortion at the vacancy sites. The case most often treated in textbooks[31,32] is where the number of surface sites is ignored in the counting of the entropy. The number of vacancies n is given by[31] as

$$n = \eta N e^{-E_f/kT}, \tag{12}$$

where

$$\eta = \exp\left(\frac{-1}{kT}\frac{\partial F(n,T)}{\partial n}\right), \tag{13}$$

and n is the number of sites for the vacancy. If the change in the surface is included explicitly in the calculation of the Schottky defect formation energy, one must include the complete information about the surface. Simply including the entropy of the vacancy atoms removed to the surface in the Bragg–Williams approximation, assuming a perfectly flat surface when there are zero vacancies (see, for example, Landsberg and Canagaratna),[33] is not sufficient. A full account of the nature of the surface, to include the atom/surface-vacancy distributions,[34] the appropriate entropy, multiple surface layers, and even surface reconstruction, is necessary so that the true surface-bulk equilibrium can be described. Because the problem can be decoupled into two equilibria, i.e., the surface with the ideal bulk and the vacancies with the ideal bulk, the equilibrium vacancy concentration can be expressed by Eq. (12) in most cases.

Because of the difficulty of measuring the vacancy formation energy, only a limited number of values exists in the literature. Even for silicon, the most studied of the semiconductors, the value for the vacancy formation enthalpy is not firmly established, because of the complication of diffusion experiments by the possibility of multiple simultaneous diffusion processes such as vacancies and interstitials. High-temperature diffusion in Si is found to follow a typical Arrhenius behavior, with measured activation energy of 4–6 eV.[35] The activation energy in diffusion E_a, is given by the sum of the vacancy formation energy E_f and the migration energy E_m. Migration energies measured by Watkins[36] are of the order of 0.25 eV. Combining our vacancy formation energy with the measured migration energies, we find $E_a \cong 6$ eV, consistent with the measured activation energies. Our results are also in excellent agreement with local density, pseudopotential, Green's function calculations of Car *et al.*[14] Note that both the present results and those of Car *et al.* are approximately twice the value of E_f from earlier estimates.[12,37]

The extraction energies are more difficult to measure in the semiconductor compounds because of the presence of two classes of defects on the cation and anion sublattice. In the annealing experiments typically done, assumptions about which native defects are dominant must be inferred from the experiments, and additionally, much of the analysis must employ values of the donor or acceptor levels of the vacancies, which are in themselves not conclusively identified in most compounds. The extraction energies in a number of II-VI compounds have been deduced from annealing experiments, characterized by Hall measurements. Ray and Kröger[38] have measured the Zn extraction energy for ZnSe as $E_x^{Zn} = 4.6$ eV, which compares to our calculated value of $E_x = 5.7$ eV. In CdS, Kumar and Kröger[30] have measured $E_x^{Cd} = 4.0$ eV and $E_x^{S} = 7.3$ eV compared to our calculated values of $E_x^{Cd} = 6.1$ eV and $E_x^{S} = 8.8$ eV. In CdTe, Chern, Vydyanath, and Kröger[29] have measured $E_x^{Cd} = 4.7$ eV and $E_x^{Te} = 6.0$ eV compared to our calculated values of $E_x^{Cd} = 4.7$ eV and $E_x^{Te} = 7.4$ eV. Finally, in $Hg_{1-x}Cd_xTe$ ($x = 0.2$), Vydyanath[39] measures $E_x^{Hg} = 2.2$ eV, while we find $E_x^{Hg} = 1.6$ eV. Note that in general our calculated values are larger than the measured values, although for the cation vacancy in CdTe our agreement with the measured values is quite good. We do find that $E_x^{anion} > E_x^{cation}$, consistent with experiment. Because of possible contribution from compensation effects, for which one must know about the localized levels in the gap, care must be exercised when comparing our results directly with the measured Schottky defect formation energies, in order to ensure that the experimental situation approximates the initial and final states treated in our calculations.

Baraff and Schlüter[15] have calculated the total energies for the cation and anion vacancies in GaAs using a pseudopotential Green's function calculation within the local density approximation. We compare our results with their value for the Schottky defect energy, $E_f^{GaAs} = 3.35$ eV/atom, where both vacancies are neutral.[40] The corresponding values from the present work is $E_f^{GaAs} = 4.0$ eV. Jansen and Sankey[16] have also calculated the vacancy formation energies for several compounds using an *ab initio* pseudo-atomic-orbital method. Our calculated energies are in good agreement with their values. They calculate formation energies of 4.8, 4.25, 4.3, and 3.7 eV in GaP, GaAs, ZnSe, and ZnTe, respectively. These compare to our values of 5.0, 4.0, 4.6, and 4.2 eV for the same compounds. As in the comparison with experiment, we find that the present tight-binding model in general yields energy values which are higher than the *ab initio* values.

C. II-VI pseudobinary alloys

While the cluster approximation is expected to introduce some error in the calculated extraction energies, as noted above, the results presented here emphasize the relative energies of different local clusters in the alloys, and hence the cluster calculation should not change our major conclusions. Corrections from Jahn–Teller distortion and Coulomb energies have not been included in this alloy study.

1. Common anion alloys

We begin by examining in detail the extraction energies in one II-VI common anion alloy. The Te common anion alloy system has been chosen as an example because of the technological importance of these alloys. HgCdTe and HgZnTe are both narrow-gap alloys in the Hg-rich concentrations with infrared device applications, and CdZnTe is often employed as a substrate material for HgCdTe and HgZnTe. Additionally, the alloy variation of E_x in this system is representative of the behavior found in other alloy systems.

Cation and anion extraction energies in the Te common

anion alloys $Zn_{0.5}Cd_{0.5}Te$, $Cd_{0.5}Hg_{0.5}Te$, and $Hg_{0.5}Zn_{0.5}Te$ are shown in Fig. 2 as a function of the near-neighbor composition. These and the following alloy results are for clusters with 53 atoms. For the Te anion vacancy, the five near-neighbor arrangements are individually unique and correspond to the values shown in Fig. 2. The values for the Zn, Cd, and Hg cation extraction energies in Fig. 2 correspond to an average over the unique arrangements of atoms in the second-neighbor positions, as shown explicitly in Fig. 3. The random-probability-weighted average extraction energy as a function of composition x is shown as the solid line in Fig. 4, with the rms deviations shown as dashed lines.

First, we discuss the anion extraction energies in the CdHgTe alloy, shown in the center panel in Fig. 2. In CdHgTe, the dependence of the Te extraction energy on n, the number of Hg atoms in the first-neighbor sites, is easily understood within the context of Harrison's bond-orbital model.[17] Although results are interpreted using bond-orbital model concepts, the results presented are those from a full cluster calculation. There are two primary contributions to the extraction energy which dominate the overall dependence of E_x on n. The first contribution to E_x accounts for a gradual increase in E_x in going from the Cd_4 Hg_1 to the Hg_4 cluster and is due primarily to the differences in the bonding energy e_h (Ref. 17) of the four bonds which are broken upon vacancy formation. Since the bonding energy level of the HgTe bond is lower than for the CdTe bond, i.e., $e_h^{HgTe} < e_h^{CdTe}$, because of the deeper s and p states on the Hg compared to Cd, this contribution to the E_x will result in an increase with an increase in the number of Hg cations adjacent to the vacancy site. The breaking of the four bonds also results in regaining the repulsive bond energy V_{ij}, which is smaller for the slightly shorter bond-length compound CdTe. This contribution to E_x will produce a decrease in E_x with the increasing number of HgTe bonds broken, but the magnitude is small because of the nearly equal lengths of the CdTe and HgTe bonds. The net result of both these bond-breaking contributions to E_x is to produce an increase in E_x with an increasing number of Hg atoms adjacent to the vacancy, as seen in Fig. 2. The second contribution to E_x that dominates the n dependence is related to the energy of

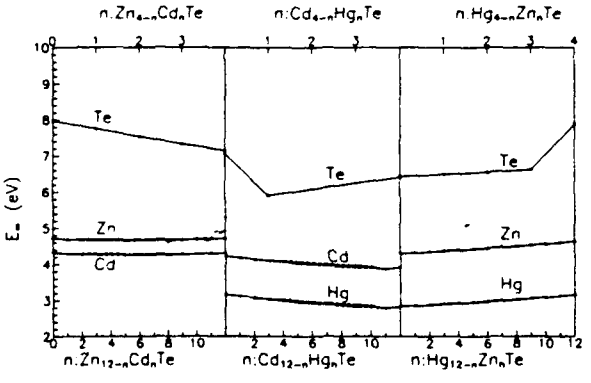


FIG. 2. Extraction energies in the Te common anion alloys as a function of near-neighbor configuration for bulk concentration $x = 0.5$. In Figs. 2 and 3–14, the solid line corresponds to the average extraction energies and the dashed lines to the rms deviations.

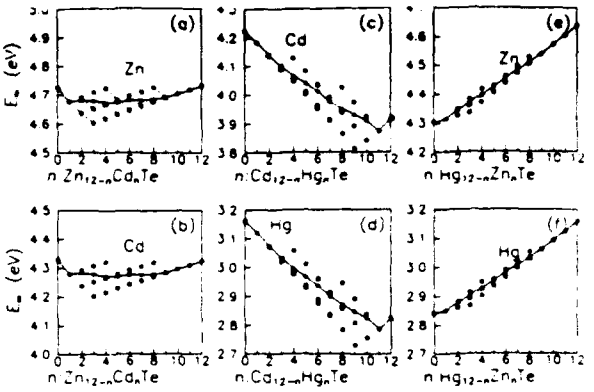


FIG. 3. Cation extraction energy in Te common anion alloys plotted for the unique configurations of second-nearest-neighbor cations, for bulk concentration $x = 0.5$. The open circles correspond to the extraction energies of unique configurations; the filled circles correspond to the probability-weighted average.

the electrons in the dangling hybrid levels. When a Te atom is removed from CdHgTe, the six outer-shell anion electrons are removed with the neutral Te atom, to leave two electrons in the four cation dangling hybrids. The energy levels that these electrons occupy depend on the cation species which are adjacent to the vacancy site. The atomic sp^3 hybrid energy of Hg is lower than that of Cd, $\epsilon_h^{Hg} < \epsilon_h^{Cd}$ (ϵ is used throughout to refer to the free atomic values, and e is used to refer to the calculated cluster values). As expected, a similar relationship is found between the calculated dangling hybrid energies at the vacancy site, where we have $e_h^{Hg} < e_h^{Cd}$. Thus, if at least one Hg atom is adjacent to the vacancy site, the two electrons will occupy the lower-energy Hg dangling hybrid level, and their contribution to the extraction energy will be approximately constant. For the Cd_4 cluster, the two electrons must occupy the higher-energy Cd dangling hybrid orbitals; this accounts for the relatively large jump in the extraction energy in going from the Cd_4 to Cd_3Hg_1 cluster in CdHgTe, as seen in Fig. 2.

A similar interpretation follows for the anion extraction energy in HgZnTe, shown in the third panel of Fig. 2. For the HgZnTe system, we have $e_h^{Zn} > e_h^{Hg}$, thus the two electrons which occupy the dangling hybrids will occupy the lower-

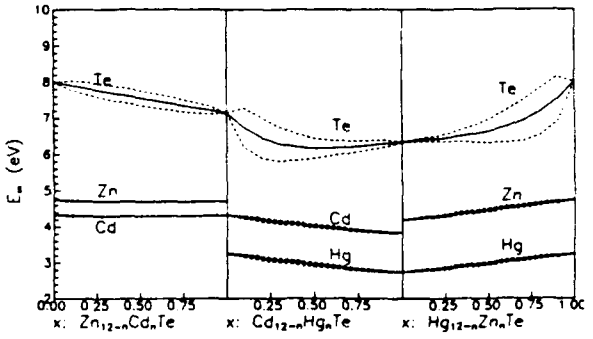


FIG. 4. Extraction energies in the Te common anion alloys as a function of composition x.

lying Hg levels. For the Hg_4, Hg_3Zn_1, Hg_2Zn_2, and Hg_1Zn_3 clusters, this contribution to E_x is nearly constant and it abruptly increases for the Zn_4 cluster, as is observed in Fig. 2. The other primary contributions to E_x come from the bonding energy and repulsive energy of the bond breaking. These two contributions produce opposite slopes of comparable magnitude between the Hg_4 and Zn_4 clusters, resulting in $E_x(x,n)$ nearly independent of n. Finally, for the Te vacancy in ZnCdTe, shown in the first panel of Fig. 2, the dependence of the contribution to E_x from the electrons in the dangling hybrids on the number of Zn and Cd atoms adjacent to the vacancy site is not as important as in CdHgTe and HgZnTe, because the difference in the dangling hybrid energies e_h^{Zn} and e_h^{Cd}, is small. For the bonding energy contribution to E_x, we find $e_h^{ZnTe} < e_h^{CdTe}$ and $V_0^{ZnTe} > V_0^{CdTe}$, resulting in a net decrease in E_x with an increasing number of CdTe bonds.

As noted above, the cation extraction energies in Fig. 2 represent an average over the unique configuration of cations in the second-neighbor shell. E_x for the unique second-neighbor cation configurations are shown in Fig. 3. The regular structure observed in Fig. 3 is in part due to the use of a Bethe lattice and the additional degeneracies it superposes on the lattice. The variations in E_x arises from shifts in the energy of the four broken bonds and the vacancy localized states through interaction with the second-neighbor alloy bonds. For a true zinc-blende lattice, we expect a smaller dispersion about the average value shown in Fig. 3.

The overall dependence of the anion extraction energy on x, shown in Fig. 4, follows the same general trends as in Fig. 2. In CdHgTe and HgZnTe, the downward bowing of $E_x(n)$ results in a downward bowing in $E_x(x)$. Also, the comparatively large variation of the Te extraction energy relative to changes in the near-neighbor environment results in a large rms deviation, particularly for the Cd-rich compositions of CdHgTe and Zn-rich compositions of HgZnTe. For the cation vacancy, the variation of the extraction energy with cluster composition is of smaller magnitude than for the anion vacancy, thus resulting in a relatively small rms deviation, although the downward bowing of the cation vacancy plot as a function of n results in a small downward bowing in the x-dependent function.

For the anion extraction energy in all of the common anion II-VI alloys, the break from linearity of the extraction energy versus cluster composition n always occurs at the $n = 1$ or 3 cluster, as was observed for the Te system in Fig. 2. This is because the two electrons which occupy the dangling hybrids will always occupy the lowest available hybrid level. For an alloy $A_{4-n}B_nC$, where $e_h^A > e_h^B$, these two electrons will occupy a B dangling hybrid, except in the $n = 0$ cluster where no B atoms are present, resulting in an abrupt increase in E_x. Because this break from linearity is mostly related to the cation hybrid energy only, one notes that if the break from linearity occurs, for example, at $n = 1$ for $A_{4-n}B_nC$, it will also occur at $n = 1$ for all $A_{4-n}B_nD$ alloys, as is observed by comparing Figs. 2, 5, and 6. This break in linearity of the curves can be predicted, since $\epsilon_h^{Hg} < \epsilon_h^{Cd}$, $\epsilon_h^{Hg} < \epsilon_h^{Zn}$, and $\epsilon_h^{Zn} \cong \epsilon_h^{Cd}$. Additionally, the sign of the slopes can be predicted within the simple bond-orbital model by comparing the major alloy-dependent contribution, $E_{bi} = 8\epsilon_h + 4V_0$, for the different systems. The bond-orbital model predicts $E_{bi}^{ZnX} > E_{bi}^{HgX} > E_{bi}^{CdX}$ for X = S, Se, and Te. This is in agreement with the sign of the slopes calculated within the cluster calculation and shown in Figs. 2, 5, and 6. Although the bond-orbital model is sufficient to predict the general trends observed, the magnitudes of the slopes must be determined from the cluster calculation, which includes, for example, the back-branch metallization effects not calculated in the simple bond-orbital analysis above.

As was found for the Te common anion alloy, the average cation extraction energies in other II-VI alloys vary nearly linearly with the cation concentration in the second-nearest-neighbor shell. The cation extraction energies in the II-VI alloys vary, depending on the particular arrangement of atoms in the second-shell sites, as was shown for Te in Fig. 3. Although these values for each alloy system are not shown, a measure of the variation is shown in the rms deviations in the plots of $E_x(n)$.

2. Common cation alloys

The structure in the cation extraction energy versus n curves for the II-VI common cation alloys, Figs. 7–9, can be interpreted in a manner similar to the anion extraction in the II-VI common anion alloys. From the bond-orbital model, we find $E_{bi}^{XS} > E_{bi}^{XSe} > E_{bi}^{XTe}$ for X = Zn, Cd, and Hg. Thus we expect this contribution to E_x will increase linearly for increasing number of S atoms in the $XS_{4-n}Se_n$ and $XTe_{4-n}S_n$ clusters, and will increase linearly with increasing Se atoms in $XSe_{4-n}Te_n$ clusters. The second contribution that dominates the n-dependent behavior of E_x is that of the energy of the six electrons in the dangling hybrids. From the atomic term values, we find $\epsilon_h^{Te} > \epsilon_h^{Se} > \epsilon_h^{S}$; thus we would predict that the electrons will occupy, in order of preference (lowest energy first) S, Se, then Te hybrids, as is observed from the results of the cluster calculations. For example, E_x for the cluster $XS_{4-n}Se_n$ can be expressed as

$$E_x = C_0 + C_1 n + \alpha e_h^S + \beta e_h^{Se}, \tag{14}$$

where $C_0 + C_1 n$ includes the composition independent and E_{bi} (nearly linear) terms, and $\alpha = 6$, $\beta = 0$, for $n = 0,1$; $\alpha = 4$, $\beta = 2$, for $n = 2$; $\alpha = 2$, $\beta = 4$, for $n = 3$; and $\alpha = 0$,

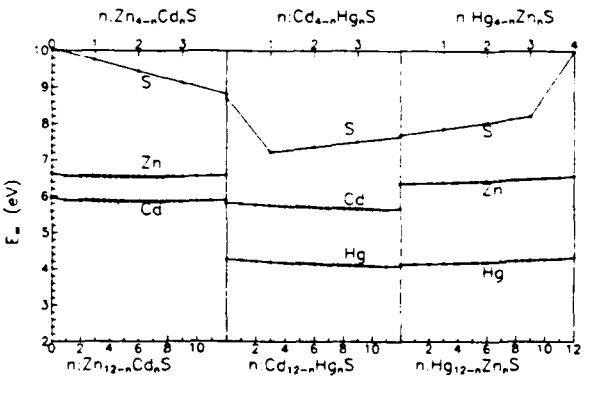


FIG. 5. Extraction energies in the S common anion alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

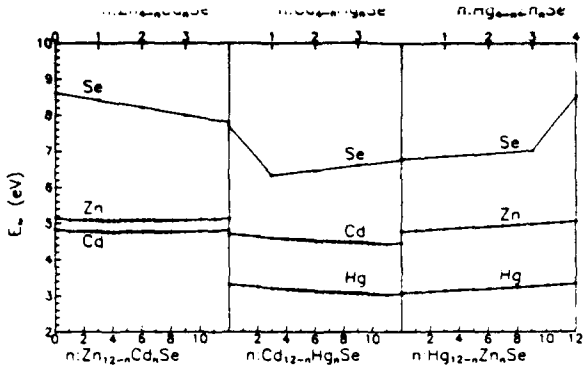


FIG. 6. Extraction energies in the Se common anion alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

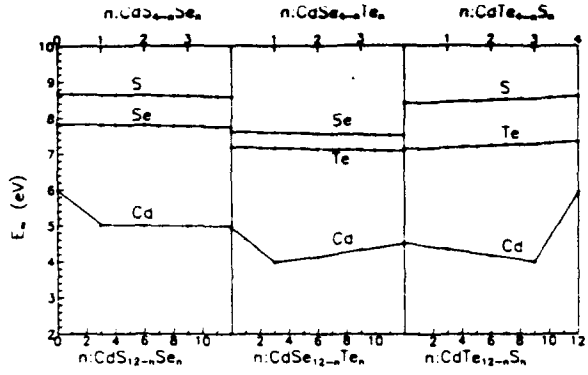


FIG. 8. Extraction energies in the Cd common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

$\beta = 6$, for $n = 4$. Note that between $n = 0$ and 1, the slope of E_x versus n is determined by the variation of E_{br} with n. For $n = 1$–4, the slope is a sum of the E_{br} contribution plus a contribution from the electrons in the dangling hybrids. Consider the Zn vacancy in $ZnS_{4-n}Se_n$. The $C_0 + C_1 n$ contribution to E_x will decrease with increasing n because $E_{br}^{ZnS} > E_{br}^{ZnSe}$. This slope can be seen between $n = 0$ and 1 in Fig. 7. The electrons in the dangling hybrids will produce an increasing contribution to E_x with increasing n because $\epsilon_h^S < \epsilon_h^{Se}$. The sum of the $C_0 + C_1 n_0$ term (positive slope) plus the energy from the electrons in the dangling hybrids (negative slope between $n = 1$ and 4) yields a net negative slope for $n = 1$–4. A similar analysis follows for the cation vacancy in other II-VI common cation alloys. Once again, the general shape and sign of the slopes of the E_x curves can thus be predicted, based on the bond-orbital analysis, although the more detailed cluster calculation is needed for a quantitative description.

The average anion vacancy in these alloys varies nearly linearly with n as expected, because the alloy dependence enters in the second-neighbor sites. A measure of the extraction energies sensitivity to the detailed arrangement in the second-neighbor sites can be gathered from the rms deviations indicated in the dashed lines in Figs. 7–9.

D. III-V pseudobinary alloys

1. Common anion alloys

For anion extraction energies in the III-V common anion alloys shown in Figs. 10–12, the analysis is similar to that in the II-VI alloys. As above, we use atomic term values and the bond-orbital model to interpret our results. We first note that $\epsilon_h^{Ga} < \epsilon_h^{Al} < \epsilon_h^{In}$ and $E_{br}^{InX} < E_{br}^{AlX} < E_{br}^{GaX}$ for X = P, As, and Sb. The analysis is as follows, using the P vacancy in $Ga_{4-n}In_nP$ (center panel of Fig. 10) as an example. Because $E_{br}^{InP} < E_{br}^{GaP}$, this energy produces a linear contribution to E_x, with negative slope, as shown. The three electrons at the vacancy site will preferentially occupy the lower-energy Ga hybrid states. Thus E_x can be written as

$$E_x = C_0 + C_1 n + \alpha e_h^{Ga} + \beta e_h^{In}, \tag{15}$$

where $\alpha = 3$, $\beta = 0$, for $n = 0$-2; $\alpha = 2$, $\beta = 1$, for $n = 3$; and $\alpha = 0$, $\beta = 3$, for $n = 4$. Thus, for $n = 0$–2, the slope of E_x vs n will be negative, dominated by the back-branch energy dependence on n. For $n = 3$ and 4, the three electrons occupying the dangling hybrid levels contribute an energy to E_x which increases with increasing n.

A similar analysis follows for the other III-V common anion alloys. With one exception, the behavior of E_x vs n follows the simple bond-orbital analysis presented above.

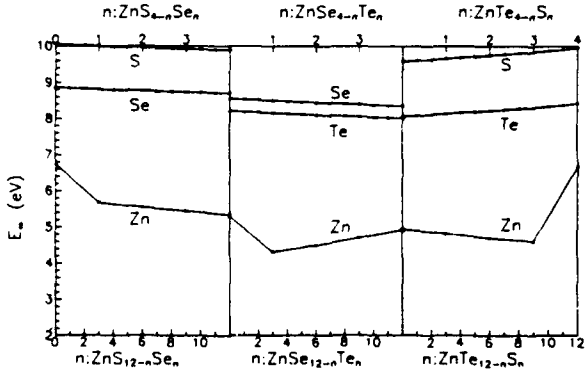


FIG. 7. Extraction energies in the Zn common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

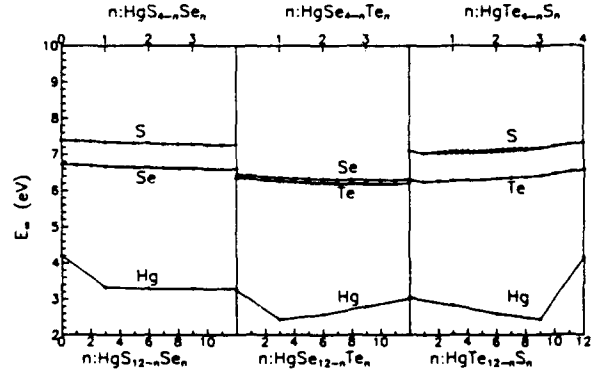


FIG. 9. Extraction energies in the Hg common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

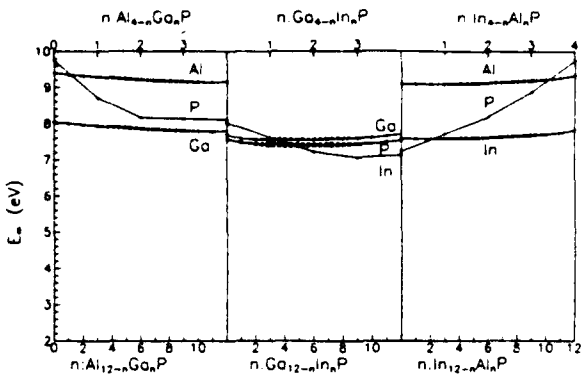


FIG. 10. Extraction energies in the P common anion alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

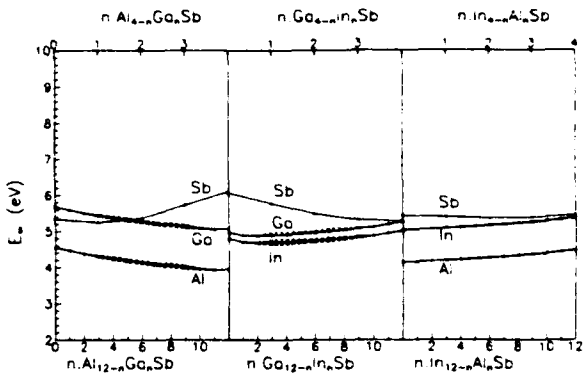


FIG. 12. Extraction energies in the Sb common anion alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

The exception is the $In_{4-n}Al_nX$ alloys, for X = P, As, and Sb. Here the atomic term values suggest that the Al hybrid levels are of lower energy than the In hybrid levels, $\epsilon_h^{Al} < \epsilon_h^{In}$, while the cluster calculations find the opposite, i.e., $e_h^{Al} > e_h^{In}$. The reasons for this reversal are related to the coupling of the dangling hybrids at the vacancy site to bonds in the rest of the cluster, which has been included in the cluster diagonalization; this can result in a significant energy shift of the dangling hybrid levels. We have found that for most systems, this shift does not alter the relative order of the hybrid energy levels.

An analysis of the cation extraction energies follows that presented for the II-VI Te alloy above. As was found for the II-VI common anion alloys, the cation extraction energies are not as sensitive to the environment as are the anion extraction energies, because the effects of alloying come in the second-neighbor sites. The rms deviations indicated as dashed lines in Figs. 10–12 provide an indication of this sensitivity.

2. Common cation alloys

For cation extraction energy in the III-V common cation alloys shown in Figs. 13–15, the analysis is similar to that in the II-VI alloys. Using the atomic term values and bond-orbital model to interpret the results, we first note that $\epsilon_h^P < \epsilon_h^{As} < \epsilon_h^{Sb}$ and $E_{br}^{XSb} < E_{br}^{XAs} < E_{br}^{XP}$ for X = Al, Ga, and In. The analysis is as follows, using the Ga vacancy in $GaAs_{4-n}Sb_n$, as shown in the center panel of Fig. 14. Because $E_{br}^{GaSb} < E_{br}^{GaAs}$, this energy produces a linear contribution to E_v with a negative slope versus n. The five electrons at the vacancy site will preferentially occupy the lower-energy As hybrid states. Thus E_v can be written as

$$E_v = C_0 + C_1 n + \alpha \epsilon_h^{As} + \beta \epsilon_h^{Sb}, \tag{16}$$

where $\alpha = 5$, $\beta = 0$, for $n = 0,1$; $\alpha = 4$, $\beta = 1$, for $n = 2$; $\alpha = 2$, $\beta = 3$, for $n = 3$; and $\alpha = 0$, $\beta = 5$, for $n = 4$. Thus, for $n = 0$–1, the slope of E_v vs n will be negative, dominated by the dependence of E_{br} on n. For $n = 2$–4, E_v will have a decreasing (with n) contribution from $C_1 n$ and an increasing contribution from the electrons in the dangling hybrids. The increase in energy occurs because the five electrons at the vacancy site occupy the higher-energy Sb hybrid levels. Thus the slope will be nearly linear between $n = 2$ and 4, because of the linear increase from the electrons in the dangling hybrid levels. Again, a similar analysis follows for the other III-V common cation alloys. All of the III-V common cation alloys follow the qualitative behavior predicted by the atomic term values ϵ_h and the bond-orbital-model results. And, as expected, the average anion extraction energies are found to vary nearly linearly with n.

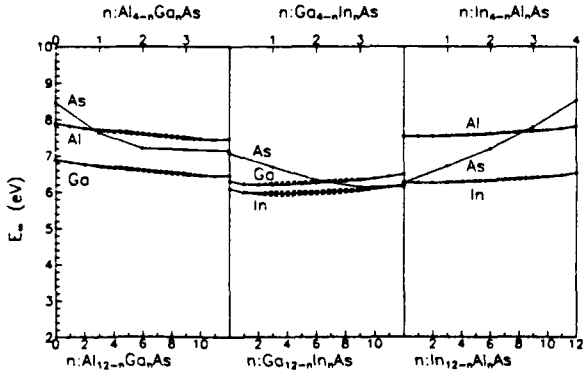


FIG. 11. Extraction energies in the As common anion alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

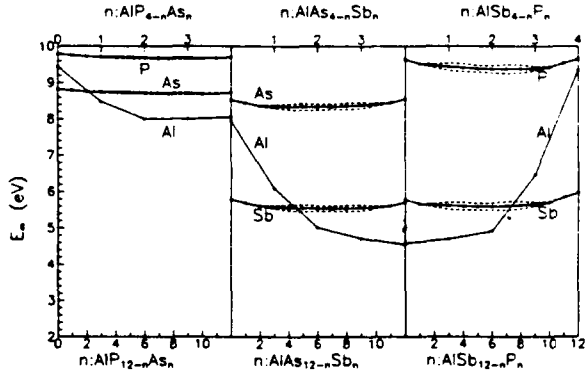


FIG. 13. Extraction energies in the Al common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

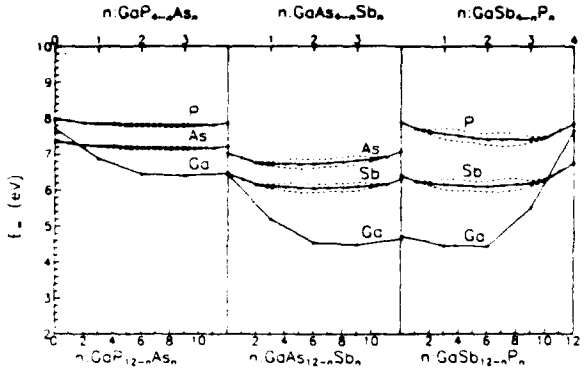


FIG. 14. Extraction energies in the Ga common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

E. Consequences of alloy variation of E_v

The predicted nonlinear variation of common-species extraction energies in alloys will have an important impact on the calculation of vacancy densities in these materials. In particular, the presence of a minimum in E_v implies larger vacancy concentrations than a simple linear variation would predict. To demonstrate this, we choose for an example the Te vacancy in the Te common anion alloys. As discussed in Sec. III B above, the definition of the Schottky defect formation energy of the individual constituents in the compounds and alloys is not rigorously defined, and as a consequence, the calculation of the vacancy concentration in these materials requires additional information about the system, e.g., stoichiometry and external phases. To demonstrate the effect that the nonlinear variation in E_v can have on properties such as the vacancy concentration, we calculate a somewhat fictitious quantity $\tilde{n}$, as defined below. This quantity has no direct significance, but does demonstrate the impact that a nonlinear variation in E_v can have on measurable properties. We begin by defining $n_i^{\mathrm{Te}}(x,n)$, which is related to the Te vacancy density for a given class of cluster n, in the alloy of composition x, as follows:

$$n_i^{\mathrm{Te}}(x,n) = \eta(x,n) N_0 P_n \exp[-E_v(x,n)/kT], \quad (17)$$

where k is Boltzmann's constant, T is the temperature in Kelvin, N_0 is the anion site density, P_n is the random probability factor, given by Eq. (6), and $\eta(x,n)$ is as defined in Eq. (13). Note that we have used the calculated zero-temperature value of E_v; more properly, the finite-temperature value, including the effect of the Fermi level, should be used here. Summing over all classes of clusters, we find

$$n_i^{\mathrm{Te}}(x) = \sum_n n_i^{\mathrm{Te}}(x,n), \quad (18)$$

and normalizing to n_v^{Te} for pure CdTe, we find

$$\tilde{n}_i^{\mathrm{Te}}(x) = \frac{n_i^{\mathrm{Te}}(x)}{n_i^{\mathrm{Te}}(\mathrm{CdTe})} = \left(\frac{a_{\mathrm{CdTe}}}{a_v}\right)^3 \exp\{[E_v^{\mathrm{Te}}(\mathrm{CdTe}) - E_v^{\mathrm{Te}}(x)]/kT\}, \quad (19)$$

where we have assumed all η's are equal. a_{CdTe} is the lattice constant of CdTe and a_v is the volume-averaged lattice constant for the alloy of composition x. Values of $\tilde{n}_i^{\mathrm{Te}}(x)$ are shown in Fig. 16 for $T = 800$ K. Shown for comparison are the calculated values based on extraction energies and assuming a simple linear interpolation between the $n = 0$ and 4 clusters of Fig. 2. The difference between the linear-interpolated values and the calculated values is quite dramatic, especially in the CdHgTe alloy where $\tilde{n}_i^{\mathrm{Te}}(x)$ differ by as much as five orders of magnitude; the difference will be smaller in calculations for which the final states on the surface are considered, and thus an energy related to E_{coh} is subtracted from E_v. Note that because of the minimum in E_v, the vacancy concentration for some composition in the alloy is found to be greater than in the constituent compounds. For the particular case of HgCdTe (center panel of Fig. 16), the difference between $\tilde{n}$ for the linear and nonlinear E_v is greatest for the Cd-rich alloys. We thus predict that even a small addition of HgTe to CdTe may result in a significant enhancement in the vacancy density in the resulting alloy, compared to that in pure CdTe. Similar results are expected for other alloys.

Next we consider the cation vacancies in the Te common anion alloys. Analogous to Eq. (18) above, we calculate

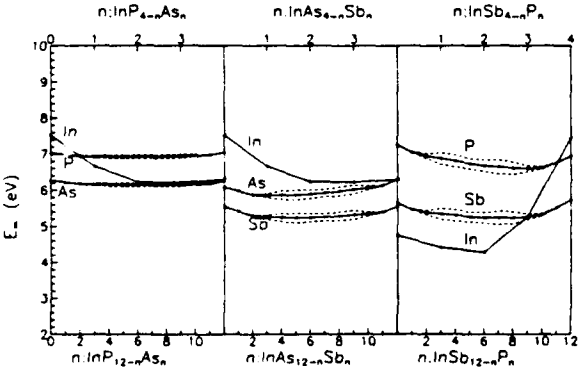


FIG. 15. Extraction energies in the In common cation alloys as a function of near-neighbor configuration for bulk concentration of $x = 0.5$.

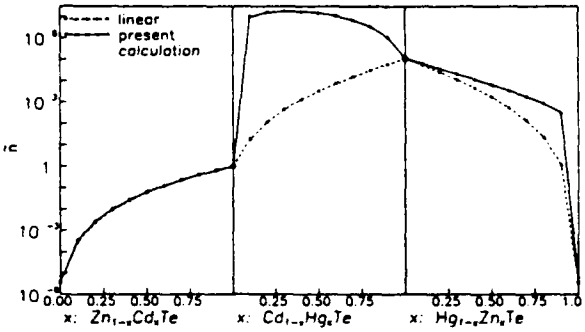


FIG. 16. Te common anion alloys using the calculated Te extraction energies from Fig. 2 (solid line) and linearly interpolated extraction energy based on the $x = 0$ and 1 values from Fig. 1 (dashed line).

$$n_i^{\text{cation}}(x,n) = N_0 \sum_{n=0}^{12} P_n \sum_{\kappa_n=1}^{C_n} \eta(x,n,g_n) P_{\kappa_n} \times \exp[-E_x(x,n,g_n)/kT] \quad (20)$$

where P_n and P_{κ_n} are as defined in Sec. III. As above, we normalize to n_i^{cation} in CdTe. Values of $\tilde{n}_i^{\text{cation}}$ are shown in Fig. 17 for $T = 800$ K. Because the nonlinearity for an atom of a substituted species was found to be less than for the common atom species, we find a less dramatic, although still substantial, difference between the linear and nonlinear results.

From the above discussion, we see that the effects of a nonlinear variation of $E_x(n)$ will be most important for the common species. One consequence of $E_x(n)$ reaching a minimum for finite alloy concentration may be a noticeable enhancement in the diffusion of the common species in the alloy, for atoms that diffuse by a simple vacancy mechanism across the complete composition range, and for materials in which the quality of the alloys is comparable to that of the constituent compounds, so that, for example, enhanced diffusion-down dislocations are not a factor. For example, Te is believed to diffuse by an interstitial mechanism[41] in $Hg_{1-x}Cd_xTe$, at least for low x; therefore, we would not expect to observe an enhanced diffusion based on a simple vacancy mechanism in this system. A second possible consequence of the nonlinear variation in E_x and subsequent effect on vacancy concentrations may be enhanced diffusion via percolation. It has been suggested[42] that percolation effects might play a role in diffusion in HgCdTe. Because vacancy densities will be greater on particular classes of sites, determined by the variations shown in Fig. 2, migration from one low-energy site to another will be energetically favored over migration to a higher-energy site, assuming migration energies do not vary by as great an amount as E_x. If there are enough low-energy sites to form percolation paths, then two parallel diffusion mechanisms may be present. This kind of multiple diffusion coefficient profile is observed for Hg in $Hg_{1-x}Cd_xTe$.[43,44] It remains to be proved that percolation, rather than diffusion-down dislocation cores or some other extrinsic mechanism, is the explanation for the Hg diffusion results. If percolation is the explanation, then based on our extraction energies, percolation should be a feature of self-diffusion in many alloys.

If deep levels exist because of the presence of vacancies, the populations of various classes of vacancies in the alloys may be measured by various techniques, such as deep-level transient spectroscopy or photoluminescence.[45,46] Results on the relative population of each class of clusters will depend on the correlation state of the alloy, its randomness, and the vacancy formation energy for the given class of cluster. A deviation from a simple random population for the various cluster types could indicate either a correlated alloy, driven by chemical or strain terms in the bulk, or a variation in the formation energy in the various classes of clusters.

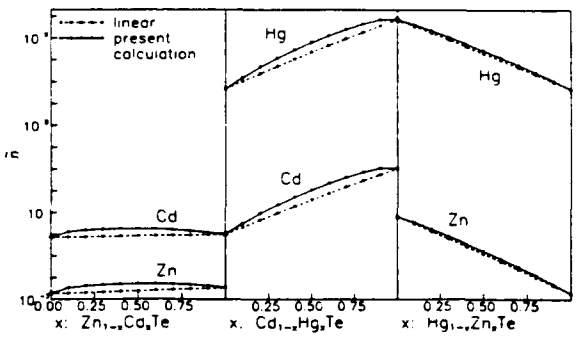


FIG. 17. Normalized cation vacancy concentrations in the Te common anion alloys using the calculated cation extraction energies from Fig. 2 (solid line) and linearly interpolated extraction energy based on the $x = 0$ and 1 values from Fig. 1 (dashed line).

V. SUMMARY AND CONCLUSION

We have found extraction energies calculated using a tight-binding cluster Hamiltonian give good agreement with the tight-binding Green's function results, in particular for the cation vacancies. Our results indicate that the extraction and Schottky defect formation energies increase with increasing cohesive energy. In a given compound, the calculated extraction energies are found to be larger for the anion than the cation, in agreement with the experimental results, with the difference between the cation and the anion being larger in the II-VI than in the III-V compounds. Comparison of our calculated values with experiments and previous theory show good agreement for Si and CdTe, although our calculated values are in general larger than experiments for those compounds for which data are available. This may be due to differences in the occupation of localized levels, which may be poorly modeled in our cluster calculation, particularly for the anion vacancy which in general creates donor levels in the gap. For the extraction energy of an atom of the substituted species in the alloys, we find a nearly linear dependence on the near-alloy environment, as expected, because the substitution is in second-neighbor sites. We have found a large variation in the extraction energy with the near-neighbor environment for the removal of an atom of the common species for which the alloy substitution is in the first-neighbor shell. In some alloys we have found the extraction energy to reach a minimum away from the compound endpoints, indicating the possibility for enhanced vacancy densities, and consequently enhanced diffusion, in the alloys.

ACKNOWLEDGMENTS

The financial support for this work through NASA Contract No. NAS1-18226, ONR Contract No. N00014-88-C0096, and AFOSR Contract No. F49620-88-K-0009 is gratefully acknowledged.

APPENDIX A: COULOMB ENERGY

In the zinc-blende semiconductor compounds, the cations and anions in the lattice carry a net charge of magnitude $|Z|$, but of opposite sign at the two lattice sites. When a neutral cation or anion is removed from the zinc-blende lattice, the net charge that previously resided at that site is redistributed in the lattice. This results in a modification to

the Coulomb energies in the crystal, both to the Madelung-enhanced K and the average local electron-electron interaction energy U. Defining Z_i as the charge initially residing on an atom in the ith shell about the vacancy site, with Z_0 the charge on the central atom, Z_1 the charge on a first-neighbor atom, etc., and δ_i as the modification to the charge on an ith-neighbor atom when the vacancy is created, the contribution to the vacancy formation energy due to the electron-electron energy ΔU and the Madelung energy ΔK are given by

$$\Delta U = \sum_i \sum_{m_i} \frac{1}{2} U_i e^2 [(Z_i + \delta_i)^2 - (Z_i)^2], \tag{A1}$$

$$\Delta K = \sum_i{}' \sum_j \sum_{m_i} \sum_{m_j} \frac{e^2}{2} \times \left(\frac{(Z_i + \delta_i)(Z_j + \delta_j)}{r_{ij}} - \frac{(Z_i)(Z_j)}{r_{ij}} \right). \tag{A2}$$

Here i and j sum over all shells of atoms, m_i and m_j sum over the atoms within the ith and jth shell, respectively, and r_{ij} is the distance between the m_ith and m_jth atoms. The prime on the sum in Eq. (A2) indicates that $i = j$ is not to be included in the sum. Values for the intra-atomic interaction parameter U_i differ at the cation and anion site and were taken from Ref. 47.

From the diagonalization of the cluster Hamiltonian containing a vacancy, we find, for most systems, that the charge redistribution is primarily to the first-neighbor shell about the vacancy. Assuming all charge shift is to the first-neighbor sites, the above expressions reduce to

$$\Delta U = -\left(\tfrac{1}{2} U_0 + \tfrac{1}{8} U_1\right) Z_0^2 e^2 \tag{A3}$$

$$\Delta K = \frac{2.17}{d} Z_0^2 e^2. \tag{A4}$$

In calculating ΔK we have included an effective reduction in the Madelung constant to account for the effects of the finite spacial extent of the atomic charge.[47]

APPENDIX B: LATTICE RELAXATION

When a vacancy is created, the atoms about the vacancy site will in general relax away from their perfect crystalline position. Here we consider only radial displacements with respect to the cluster center, Jahn–Teller distortions, which lower energies by breaking degeneracy, are not addressed. For the unrelaxed lattice we define R_n as the distance from the cluster center to the nth atom shell. The relaxation in the cluster containing a vacancy is then characterized by δ_n, defined through the relationship

$$r_n = R_n (1 - \delta_n), \tag{B1}$$

with r_n defined as the distance from the center to the nth-neighbor atom in the relaxed cluster.

The r_n are computed by minimizing E_v with respect to bond-length changes. In the energy minimization, the elastic energy of the lattice beyond the cluster is included as an elastic continuum. Covalent energy changes from both bond-length and bond-angle distortions within the cluster have been included, and rigid hybrids have been assumed. Here $V_0(d_0)$, $\partial V_0(d_0)/\partial d$, and $\partial^2 V_0(d_0)/\partial d^2$ for the initial-state cluster were fit to the experimental cohesive energy, the equilibrium bond length d_0, and the experimental bulk modulus, respectively. For small bond-length variations away from equilibrium, $V_0(d)$ can be expressed by the Taylor's series expansion

$$V_0(d) = V_0(d_0) + (d - d_0) V_0'(d_0) + \tfrac{1}{2}(d - d_0)^2 V_0''(d_0). \tag{B2}$$

This form of V_0 is used to compute the overlap energy of final-state bonds of length d. The strain energy of the elastic continuum is calculated from

$$E_{\text{out}}^{\text{elastic}} = \overline{C} \left(\tfrac{8}{3}\right)^{1/2} (d)^2 \delta_I^2, \tag{B3}$$

where I corresponds to the outermost shell of the cluster and $\overline{C}$ is an effective shear constant[23] given by

$$\overline{C} = \pi [1.6(C_{11} - C_{12}) + 4.8 C_{44}]. \tag{B4}$$

Experimental values for $C_{11} - C_{12}$ and C_{44} have been used.

APPENDIX C: DANGLING HYBRID INTERACTION

One additional correction to the vacancy formation energies should be added, because of the Jahn–Teller distortion at the vacancy site and the resulting splitting of the degeneracy and lowering of the total energy. This interaction between the dangling hybrids at the vacancy site was not included in the cluster Hamiltonian because only first-neighbor interactions have been included to this point. For large enough cluster sizes (three or more shells of neighbors about the vacancy site), the dangling hybrid states are narrowly split into a singly degenerate A_1 state and a triply degenerate T_2 state, from the coupling of the dangling hybrid levels through unbroken cluster bonds. The direct dangling hybrid interaction results in an additional splitting of the levels at the T_d symmetry vacancy site, with the A_1 level lowered by an amount 3α and the T_2 levels raised by an amount α, where α is the magnitude of the coupling constant. From self-consistent Green's function calculations, Baraff, Kane, and Schlüter[48] found a total of $A_1 - T_2$ splitting of 1.7 eV for the vacancy in silicon. Assuming half of this splitting results from the Coulomb interaction, we estimate $\alpha_{\text{Si}} = 0.21$ eV. To estimate the contribution of the splitting of the dangling hybrid energies on the vacancy formation energy in other compounds, the value of α for Si is scaled according to

$$\alpha = \alpha^{\text{Si}} \left(\frac{e_h}{e_h^{\text{Si}}}\right) \left(\frac{d^{\text{Si}}}{d}\right)^2, \tag{C1}$$

as suggested by extended Hückel theory. Here e_h is the hybrid energy for the cation or anion adjacent to the vacancy site. Although not rigorous, this simple scaling rule should yield the proper trends among the compounds.

[1] S. T. Pantelides, Rev. Mod. Phys. 50, 797 (1978).
[2] S. T. Pantelides, Ed., *Deep Centers in Semiconductors* (Gordon and Breach, New York, 1986).
[3] M. Jaros, Adv. Phys. 29, 409 (1980).
[4] M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I: Theoretical Aspects* (Springer, New York, 1981).
[5] G. D. Watkins and R. P. Messmer, in *Computational Methods for Large Molecules and Localized States in Solids*, edited by F. Herman, A. D. McLean, and R. K. Nesbet (Plenum, New York, 1973), p. 133.

[6] K. H. Johnson, Adv. Quantum Chem. **4**, 143 (1973).
[7] F. P. Larkins, J. Phys. C **4**, 3065 (1971); **4**, 3077 (1971).
[8] S. G. Louie, M. Schluter, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B **13**, 1654 (1976).
[9] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. Lett. **41**, 895 (1978).
[10] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **41**, 892 (1978).
[11] G. A. Baraff and M. Schluter, Phys. Rev. B **30**, 1853 (1984).
[12] R. A. Swalin, J. Phys. Chem. Solids **18**, 290 (1961).
[13] J. C. Phillips and J. A. Van Vechten, Phys. Rev. Lett. **30**, 220 (1975).
[14] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, Phys. Rev. Lett. **52**, 1814 (1984).
[15] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985).
[16] R. W. Jansen and O. F. Sankey, Phys. Rev. B **39**, 3192 (1989).
[17] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); Phys. Rev. B 27, 3592 (1983); Microscience (limited distribution report, SRI International, Menlo Park, CA, 1983), Vol. IV, p. 34.
[18] M. A. Berding, A. Sher, and A.-B. Chen, J. Vac. Sci. Technol. A **5**, 3009 (1987).
[19] M. A. Berding, A. Sher, A.-B. Chen, and R. S. Patrick, Semicond. Sci. Technol. **5**, 588 (1990).
[20] A. Sher, M. A. Berding, M. van Schilfgaarde, A.-B. Chen, and R. S. Patrick, J. Cryst. Growth **86**, 86 (1988).
[21] A. Sher, M. A. Berding, A.-B. Chen, and R. S. Patrick, J. Cryst. Growth **89**, 27 (1989).
[22] W. A. Harrison, Phys. Rev. B **24**, 5835 (1981).
[23] A.-B. Chen and A. Sher, Phys. Rev. B **31**, 6490 (1985).
[24] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1956).
[25] K. Shih, W. E. Spicer, W. A. Harrison, and A. Sher, Phys. Rev. B **31**, 1139 (1985).
[26] A. Sher, M. van Schilfgaarde, A.-B. Chen, and W. Chen, Phys. Rev. B **36**, 4279 (1987).
[27] A. Redondo, W. A. Goddard, III, T. C. McGill, and G. T. Surratt, Solid State Commun. **20**, 733 (1976).
[28] H. D. Shih, F. Jona, D. W. Jepsen, and P. M. Marcus, Phys. Rev. Lett. **37**, 1622 (1976).
[29] S. S. Chern, H. R. Vydyanath, and F. A. Kröger, J. Solid State Chem. **14**, 33 (1975).
[30] V. Kumar and F. A. Kroger, J. Solid State Chem. **3**, 387 (1971).
[31] R. Kubo, *Statistical Mechanics* (Elsevier, New York, 1971).
[32] C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1986).
[33] P. T. Landsberg and S. G. Canagaratna, Phys. Status Solidi B **126**, 141 (1984).
[34] D. Shaw, Ed., *Atomic Diffusion in Semiconductors* (Plenum, New York, 1973).
[35] K. A. Jackson, *Liquid Metals and Solidification* (American Society for Metals, Cleveland, OH, 1957), p. 174.
[36] G. Watkins, in *Lattice Defects in Semiconductors*, edited by F. A. Huntley, Inst. Phys. Conf. Ser. No. 23 (Institute of Physics, London, 1975), p. 1.
[37] J. A. Van Vechten, in *Handbook on Semiconductors*, edited by T. S. Moss (North-Holland, Amsterdam, 1980), p. 1; J. Electronchem. Soc. **122**, 419 (1975).
[38] A. K. Ray and F. A. Kröger, J. Electrochem. Soc. **125**, 1348 (1978).
[39] H. R. Vydyanath, J. Electrochem. Soc. **128**, 2609 (1981).
[40] Both vacancies will not be neutral for a given Fermi energy in the compound. Baraff and Schluter find the neutral cation (anion) vacancy will be the lowest-energy state for the Fermi energy at the valence- (conduction)-band edge. Because the energies we have calculated correspond most closely to their neutral vacancies, we compare our results to these energies.
[41] M.-F. S. Tang and D. A. Stevenson, J. Vac. Sci. Technol. A **6**, 2650 (1988).
[42] M.-F. S. Tang (private communication).
[43] M. Brown and A. F. W. Willoughby, J. Vac. Soc. Technol. A **3**, 1641 (1983).
[44] M. F. S. Tang and D. A. Stevenson, J. Vac. Sci. Technol. A **5**, 3124 (1987).
[45] H. G. Grimmeiss, E. Meijer, R. Mach, and G. O. Muller, J. Appl. Phys. **56**, 2768 (1984).
[46] L. Samuelson, S. Nilsson, Z.-G. Wang, and H. G. Grimmeiss, Phys. Rev. Lett. **53**, 1501 (1984).
[47] M. van Schilfgaarde, A.-B. Chen, and A. Sher, Phys. Rev. Lett. **57**, 1149 (1986).
[48] G. A. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. B **21**, 5662 (1980).

Defects in ZnTe, CdTe, and HgTe: Total energy calculations

M. A. Berding, M. van Schilfgaarde, A. T. Paxton, and A. Sher
SRI International, Menlo Park, California 94025

(Received 4 October 1989; accepted 2 November 1989)

Total energies for various impurities and defects in HgTe, CdTe, and ZnTe are calculated. Calculations were done using a self-consistent linear muffin-tin orbital (LMTO) method within the local density and atomic spheres approximation. We calculate the total energy for substitution on both lattice and interstitial sites, and estimate the lattice strain energies. Estimates of the variation with the alloy predict a linear variation of the substitution energy with the local concentration. We predict that the Te antisite will be more prevalent in all three of the compounds than previously thought. The problem of cross doping during heteroepitaxy on GaAs is predicted to be greater on the cation sublattice.

I. INTRODUCTION

Controlled doping of HgCdTe and HgZnTe during bulk and epitaxial growth is an important aspect of device fabrication from these materials. Understanding the thermodynamics of the various native point defects and impurities substitution energies is a key step towards controlling the doping in these narrow-gap alloys. Equilibrium between the native defects and impurities, subject to a specified stoichiometry and external phases, must be understood in order to deduce ways to improve control of doping in these materials. Additionally, the kinetics of defect and impurity diffsion can be important. Calculation of thermodynamic properties of the various native and impurity point defects require free energies (enthalpies and entropies) of formation. The energies reported here are the appropriate enthalpies. These energies also provide a reference point for the discussion of more complicated phenomena, such as dopant incorporation during nonequilibrium epitaxial growth or ion implantation.

Accurate total energies of defects and impurities in semiconductors are difficult to calculate. Even more challenging is the theory of the energy positions of the deep levels resulting from such defects. Although tight-binding methods have been successful in studying the trends of deep levels as a function of, for example, host lattice or alloy environment, experimental guidelines are often used to specify the impurity potential itself.[1] Because the local density approximation (LDA) yields an incorrect band gap, identifying the energy levels of localized states in the band gap in first-principle calculations using the LDA is still an outstanding problem. Additionally, in the present work, the use of supercells leads to a large dispersion in the localized states, adding to difficulty in directly calculating the deep state energies, although attempts have been made to do so.[2] In the present work, we do not focus on the exact positioning of the localized levels in discussing our results, although general trends in the deep levels are discussed in some cases.

The remainder of this paper is organized as follows. In Sec. II, the theoretical approach is summarized and references to the appropriate literature are made. In Sec. III, we present results for an array of native defects and dopants in HgTe, CdTe, and ZnTe. The work is summarized and conclusions are drawn in Sec. IV.

II. THEORY

The self-consistent linear muffin-tin orbital (LMTO) method, originally developed by Andersen,[3] is used to calculate the substitution energies in HgTe, CdTe, and ZnTe. As is customary, the LDA to the exchange-correlation energy is used to construct an effective one-electron potential. Within the atomic spheres approximation, LMTO is computationally efficient and, unlike the pseudopotential method, is suitable for solids composed from any atom in the periodic table. We use a minimal (*spd*) basis set on atom- and space-centered atomic spheres. The resulting 18 functions per atom are further reduced by orbital downfolding[4] of all *d*-orbitals except those on the cation atomic spheres. Without such an optimal basis set, the large data base of calculations presented here would not have been possible.

Substitution energies are calculated from a difference in total energies of the compound with and without the defect or impurity. Large unit cells of 32 spheres with one impurity per cell are repeated periodically to form a superlattice of impurities. The sensitivity of the substitution energies to the size of the unit cell has been examined and errors of the order of 0.2 eV are estimated. The periodic array of impurities results in the dispersion of the impurity states into bands in **k** space, contributing to the difficulty in identifying the deep state energies. As a check of the present supercell approach, we have repeated the LMTO Green's function calculations done by Beeler *et al.*[5] using supercells, and we find similar conclusions for the chalcogen impurities in silicon.

Because the lattice strain about some impurities and defects can be substantial, calculation of the relaxation energy is necessary to yield an accurate estimate of the total energies. In the present calculation, Methfessel's form of the force theorem[6] is used to calculate the pressure on each supercell containing a single defect or impurity. An upper estimate of the lattice-relaxation energy is calculated assuming all bonds in the supercell respond uniformly to this pressure, characterized by the bulk modulus of the host lattice.

The atomic spheres approximation (ASA), as used in the present work, also substitutes for the true density functional one that spheridizes the output density. In consequence of this, calculation for defects with high deviations from cubic

symmetry, (e.g., defects that cause large strain) must be viewed with caution.

III. RESULTS

For the pure compounds without any defects, the LMTO-ASA was found to give cohesive energies, bond lengths, and bulk moduli in good agreement with experiment. Calculated cohesive energies per bond are − 1.28, − 1.19, and − 0.91 eV for ZnTe, CdTe, and HgTe, respectively, compared to the experimental values of − 1.20, − 1.10, and − 0.82 eV. Total energies for several group IB, III, and V dopants in the pure compounds HgTe, CdTe, and ZnTe are summarized in Table I. Also shown are the total energies for the antisite and self-interstitials. Two nonequivalent tetrahedral interstitial sites were considered, one with four anion nearest neighbors *I* (Te), and the other with four cation nearest neighbors *I*(Zn, Cd, or Hg).

The total energies given in Table I are referenced to initial and final states where the excess (or exchanged) atoms are taken to free atom states. For example, for the anion antisite defect in HgTe, the initial state is bulk HgTe with a free Te atom and the final state is bulk HgTe with one antisite defect and a Hg in the free-atom state. This serves as a convenient reference from which other energies of interest can be easily obtained. We caution against a direct interpretation of the numbers in this table, because the relative and absolute magnitudes, as well as the sign of the energies, may change when another reference state is chosen. Perhaps of more interest for some cases is when the impurities are taken from the pure solids, rather than the free-atom state. This reference state was calculated for several of the impurities, and substitution energies are given in Table II. For impurity doping during LPE and bulk growth, the liquid final state is the relevant one, and can be obtained if the energy of the atoms in the liquid phase are known.

TABLE I. Total energies of defect formation. Notation is A (B), where A is the defect and B is the defect site. When atoms are added or removed in the formation of a defect, the free atom initial and final states are used as a reference. Sign in parentheses following the energies have the following meaning: (+ +) relaxation is outward and relaxation energy is greater than 0.25 eV; (+) relaxation is outward and relaxation energy is less than 0.25 eV; (.) relaxation is negligible; (−) relaxation is inward and relaxation energy is less than 0.25 eV.

Defect	Total energy (eV) ZnTe	CdTe	HgTe
Group IB impurities			
Cu (cation)	− 1.41(+)	− 0.98(−)	− 2.21(−)
Ag (cation)	− 0.28(+)	− 0.62(−)	− 1.92(−)
Au (cation)	− 0.53(.)	− 1.02(−)	− 2.33(−)
Group III impurities			
Al (cation)	− 1.04(+)	− 1.57(.)	− 3.02(−)
Ga (cation)	− 0.25(+)	− 0.77(.)	− 2.02(.)
In (cation)	+ 0.80(+ +)	− 0.44(+)	− 1.93(+)
Group V impurities			
P (Te)	+ 1.26(−)	+ 1.39(−)	+ 1.07(−)
As (Te)	+ 1.10(−)	+ 1.23(−)	+ 0.80(−)
Sb (Te)	+ 0.95(−)	+ 1.08(.)	+ 0.48(−)
P (cation)	+ 1.45(+)	+ 0.74(.)	− 0.56(.)
Sb (cation)	+ 3.07(+ +)	+ 1.51(−)	− 0.36(+)
Impurity-interstitials			
P [*I* (cation)]	+ 0.01(+ +)	− 0.17(+)	− 0.41(+)
P [*I* (Te)]	+ 0.54(+ +)	+ 0.05(+)	− 0.90(+)
In [*I* (cation)]	···	···	− 0.83(+ +)
In [*I* (Te)]	···	···	− 0.82(+ +)
Self-interstitials			
Cation [*I* (cation)]	+ 0.43(+ +)	+ 0.87(+ +)	+ 0.98(+ +)
Cation [*I* (Te)]	+ 0.78(+ +)	+ 1.03(+ +)	+ 0.84(+ +)
Te [*I* (cation)]	+ 2.61(+ +)	+ 1.45(+ +)	+ 1.33(+ +)
Te [*I* (Te)]	+ 4.28(+ +)	+ 2.63(+ +)	+ 1.32(+ +)
Antisites			
Cation (Te)	+ 4.62(+)	+ 3.94(.)	+ 3.22(.)
Te (cation)	+ 4.59(+ +)	+ 2.52(+)	+ 0.89(+)

A. Cu, Ag, and Au impurities

We first consider a situation in which a group IB metal (Cu, Ag, or Au) is deposited as a solid on the surface of the II–VI compound of interest, and ask which has the lowest energy for substitution on the cation lattice. From Table II we find that Ag has the lowest substitution energy in HgTe and CdTe, while Cu has the lowest energy for ZnTe. Implicit in these energies is the assumption that the replaced cation escapes to the vapor and equilibrium is achieved (i.e., rates of diffusion are not relevant).

The above ease of substitution will be changed if the impurity is being introduced from the vapor. In this case, Ag will substitute most easily in ZnTe and CdTe, while Au will substitute most easily in HgTe. For ZnTe and CdTe, which have relatively wide band gaps, the zero-temperature substitution energies for the intrinsic compound may be altered as a dependence on the Fermi energy and deep-state energies enters in the substitution enthalpies.

Although the above conclusions are based on enthalpy considerations alone, the conclusions still hold if the three metals are compared for the same external conditions, i.e., temperature and Hg pressure. The substitution energies are only one aspect needed to understand the diffusion of these

TABLE II. Total energies of defect formation. Notation is A (B), where A is the defect and B is the defect site. For Te and impurity, the pure solid is used as a reference; the free-atom state is used as the final state for the Zn, Cd, or Hg cation. Energies for the elemental solids are taken from experiment.

Defect	Total energy (eV) ZnTe	CdTe	HgTe
Cu (cation)	+ 2.08	+ 2.51	+ 1.28
Ag (cation)	+ 2.67	+ 2.33	+ 1.03
Au (cation)	+ 3.27	+ 2.78	+ 1.47
Al (cation)	+ 2.35	+ 1.82	+ 0.37
Ga (cation)	+ 2.56	+ 2.04	+ 0.79
In (cation)	+ 3.32	+ 2.08	+ 0.59
P (Te)	+ 2.50	+ 2.53	+ 2.31
As (Te)	+ 1.87	+ 2.00	+ 1.57
Sb (Te)	+ 1.51	+ 1.64	+ 1.04
P (cation)	+ 4.88	+ 4.17	+ 2.87
Sb (cation)	+ 5.82	+ 4.26	+ 2.39

metals into the compounds. The rates of diffusion, as well as other chemical reactions (e.g., formation of metal tellurides) must be included for a full description of the behavior of these metals with the compounds.

B. Al, Ga, and In impurities

Again we first consider the case where Al, Ga, or In solid is deposited on the surface and ask the relative ease for substitution into the compound on the cation sublattice. From the energies in Table II, we find that Al will substitute most easily on the cation lattice for all three compounds: ZnTe, CdTe, and HgTe. In CdTe and HgTe, In—an oft-employed dopant in HgCdTe—also has a low substitution energy; in ZnTe, because of a large size mismatch, the In substitution energy is substantially larger than that of Al and Ga. Based on these results, one would expect the doping of HgZnTe with this class of impurities will be more difficult than in HgCdTe. Because all of the energies for the group III impurities in HgTe are less than the respective impurities in ZnTe and CdTe, we predict that the equilibrium doping in HgTe (and thus HgCdTe and HgZnTe) will be easier than in the pure compounds CdTe and ZnTe.

When the free-atom final states are considered, the group III elements are predicted to be stable in ZnTe, CdTe, and HgTe, i.e., their binding energies are less than zero, with the exception of the highly strained In impurity in ZnTe. For the In going into the interstitial sites in HgTe, the energies are also less than zero, indicating the relative stability of these sites in HgTe, and the mercury-rich alloys HgCdTe and HgZnTe. Thus, if the In atoms are delegated to the interstitial sites during epitaxial growth (e.g., via kinetic processes on the growth surface), the interstitial sites will be relatively stable for these atoms.

Consider one further situation in which the group III element is incorporated interstitially into the compound, either through ion implantation or during epitaxial growth. We ask then how much energy does it take to exchange the interstitial with an atom on a cation lattice site, leaving the cation in an interstitial site and the impurity on a lattice site. For example, for In in HgTe, we consider the following reaction:

$$\mathrm{In}_I + \mathrm{Hg}_{\mathrm{Hg}} \rightarrow \mathrm{In}_{\mathrm{Hg}} + \mathrm{Hg}_I. \tag{1}$$

The energy for this reaction is

$$E = +E_{\mathrm{Hg}(I)} - E_{\mathrm{In}(I)} + E_{\mathrm{In}(\mathrm{Hg})} = -0.2\ \mathrm{eV}. \tag{2}$$

Although this reaction is found to be exothermic and therefore is predicted to proceed based on total energy considerations alone, this conclusion is not within the accuracy of the present calculation. Additionally, activation barriers may be large because of the atom-swapping process. Additionally, the equilibrium with the other native defects such as vacancies and antisites must be considered before predicting if the reaction will proceed. The answer in any particular case may well depend on the detailed initial state of the solid, e.g., the vacancy concentrations, dislocation density, or carrier type.

C. P, As, and Sb impurities

Because the group V elements are slow diffusers,[7] it is just this class of impurities that is desired to achieve impurity controlled p-type doping in HgCdTe. We find that the Sb substitution energy is smallest, while P is largest in all three compounds. The differences between the substitution energies in ZnTe, CdTe, and HgTe are not as pronounced as was seen for impurity substitution on the cation site. Substitution energies for P and Sb on the cation sublattice have also been calculated. Note that in HgTe the substitution energy for P does not differ significantly on the Hg and Te sublattice. Additionally, we see from Table I that the interstitial energy of P with respect to solid P is comparable to the lattice substitution energies. This result is consistent with the amphoteric behavior of this element in HgTe, observed by Vydyanath.[8]

We also note that the substitution energies for all the Group V impurities are largest in CdTe. This suggests that doping may be more difficult in HgCdTe than in HgZnTe. This conclusion is based on preliminary calculations, which indicate that the substitution energies of this class of defects varies linearly with composition in the alloy.

D. Native defects: Accommodation of deviations from stoichiometry

Deviations from stoichiometry can be accommodated by a number of native defects: antisites, interstitials, vacancies, second phase precipitates, and extended defects (e.g., point defect complexes, dislocations). As discussed by Schaake and Tregilgas,[9] which of these defects dominates is determined both by equilibrium considerations and the rate at which equilibrium is established.

We considered the relative enthalpies for accommodation of a specified nonstoichiometry by the first four defects named above. We arbitrarily choose the reference state as that of the pure compound with excess Te accommodated by Te solid. With respect to this reference state, the energy change for the fourth mode of accommodation is zero. The relevant reaction to accommodate one excess Te by the formation of an Te antisite is

$$\mathrm{Te}_s \rightarrow \tfrac{1}{2}\mathrm{Te}_{\mathrm{Hg}}\mathrm{Te}_{\mathrm{Te}}. \tag{3}$$

The energy for this reaction is given by 0.5 ($E_{\mathrm{Te}(\mathrm{Hg})} - 2E_b + 2E_{\mathrm{Te}_s}$), where E_{Te_s} and E_b are the absolute magnitudes of the energy per atom of Te solid and energy per bond for the HgTe compound, both with respect to the free atoms. To accommodate one excess Te by the formation of a Te interstitial, i.e.,

$$\mathrm{Te}_s \rightarrow \mathrm{Te}_I, \tag{4}$$

an energy $E_{\mathrm{Te}[I(\mathrm{Te\ or\ Hg})]} + E_{\mathrm{Te}_s}$ is necessary. To calculate the energy to accommodate the nonstoichiometry by Hg vacancies, we use the experimental values for the Hg vacancy-formation energy from Vydyanath,[10] $E_{V(\mathrm{Hg})} = 2.2$ eV, and the experimental cohesive energy per bond of 0.82 eV for HgTe. The relevant equation for the above is

$$\mathrm{Te}_s \rightarrow \tfrac{1}{2}V_{\mathrm{Hg}}\mathrm{Te}_{\mathrm{Te}}, \tag{5}$$

with energy given by $E_{V(\mathrm{Hg})} - 2E_b + E_{\mathrm{Te}_s}$. The energies for the above equations, as well as those for CdTe and ZnTe, are

TABLE III. Defect energies to accommodate excess Te (or Hg, Cd or Zn deficiency). Energies are referred to the pure stoichiometric compound with the nonstoichiometry accommodated by solid Te, and are per excess Te. The two values shown for the interstitial correspond to the cation- and anion-interstitial sites.

Native defect	Total energy (eV)		
	ZnTe	CdTe	HgTe
Te antisites	+ 3.21	+ 2.29	+ 1.75
Te interstitials	+ 4.82, − 6.50	+ 3.67, + 4.84	+ 3.57, + 3.57
Cation vacancies	+ 4.90	+ 4.70	+ 2.78
Te precipitates	0.00	0.00	0.00

summarized in Table III. Vacancy-formation energies for CdTe and ZnTe were taken from experiment[11] and a tight-binding calculation,[12] respectively.

Our results indicate that the antisite has a lower enthalpy of formation than the vacancy in all three compounds, and that the energy for the interstitial is comparable to or greater than that for the vacancy. The Te antisites and interstitials both exhibit large strain energy, an estimate of which has been included in the energies in Table III. These strain contributions are estimates, and thus the conclusions reached here must be viewed as somewhat tentative. The configurational entropy contribution to the free energy for the antisites and vacancies are the comparable, and our results indicate that the antisite should be abundant. If—as expected on simple considerations—the Te antisite is a donor, it could be a major contributor to the residual donor in Hg annealed samples. If both vacancies and antisites are present, compensation will occur and a net p- or n-type material can result, depending on the defect-state structure of each material. Finally, the Te antisites may have associated with them deep levels in the band gap, most probably donorlike. These may be the deep states at 0.4 E_g, and 0.7 E_g, where E_g is the band gap energy, present in undoped HgCdTe, which appear donorlike.[13] We are currently confirming the results of the total energy calculations using a full-potential LMTO, in which the ASA is removed. For now, we conclude that the antisite is probably more abundant than previously thought, and that it may be the source of the donorlike deep states in undoped HgCdTe.

E. Heteroepitaxy on GaAs

To study the tendency for cross doping across a GaAs/II–VI epilayer interface, we have calculated the energies for interchange of cations or anions. For example, for a CdTe epitaxial layer on a GaAs substrate, we must consider the reactions

$$Cd_{Cd} + Ga_{Ga} \rightarrow Cd_{Ga} + Ga_{Cd} \quad (6)$$

and

$$Te_{Te} + As_{As} \rightarrow Te_{As} + As_{Te}. \quad (7)$$

Table IV summarizes calculated substitution energies for Zn, Cd, Hg, and Te in GaAs. The energies of the above reactions for ZnTe, CdTe, and HgTe on GaAs are also given in Table IV. All of the energies for the interchange reactions are positive, indicating that, based on the enthalpy contribution to the free energy, the reactions will be inhibited (endothermic). Thus, the enthalpies of reaction will not tend to drive cross-doping heteroepitaxy in this system. Because the entropy contributions for cation and anion lattice exchange are of the same form, and because the cation energies are smaller, we can conclude that cross doping on the cation sublattice will be more dominant than on the anion sublattice. Note that the calculated enthalpies do not include the strain expected to be present on the lattice-mismatched interface.

TABLE IV. Total energies of defect formation for Zn, Cd, Hg, and Te in GaAs. Reference is same as in Table I.

Defect/defect reaction	Total energy (eV)		
	ZnTe	CdTe	HgTe
Cation (Ga)	+ 2.14	+ 3.25	+ 4.61
Te (As)	+ 2.92	+ 2.92	+ 2.92
Anion exchange [Eq. (7)]	+ 1.89	+ 2.48	+ 2.59
Cation exchange [Eq. (6)]	+ 3.93	+ 4.15	+ 3.72

IV. CONCLUSIONS AND FUTURE WORK

We are currently developing several improvements to the present work, which should result in more accurate estimates of the total energies. First, the lattice relaxation energies were estimated using the total pressure on a supercell, assuming that all the bonds in the lattice respond uniformly, characterized by the bulk modulus. Because the strain is localized in the bonds nearest to the misfit impurity or interstitial, the bonds nearer the defect will, in fact, relax more than those far from the defect. We are currently working on this problem. Second, the identification of the energy position of any deep states produced by the defects, such as the antisites and interstitials, remains a challenge. The first improvement to the current work is the implementation of a Green's function approach to calculate the properties of truly isolated defects in the compounds. Unfortunately, until the source of the band gap problem in the LDA is resolved, at most one can expect from such calculations is trends in the deep levels, not absolute energies with respect to the band edge. Finally, the variation of substitution energy with alloy concentration needs to be investigated further. Preliminary calculations indicate that the substitution energies vary linearly with alloy composition.

Several conclusions can be drawn from the above results.

(i) In appears relatively stable in the interstitial sites, perhaps explaining the difficulty in obtaining active In during epitaxial growth, where nucleation and kinetic effects may favor interstitial incorporation of In.

(ii) Enthalpy considerations indicate that interstitital In may be favored to exchange with Hg on lattice sites.

(iii) P and Sb energies on the cation and anion sublattice and interstitial sites in HgTe are all comparable, consistent with the experimentally observed amphoteric behavior of these elements, as a function of Hg pressure.

(iv) The Te antisite is predicted to be more abundant than previously expected in ZnTe, CdTe, and HgTe. The Te antisite may be responsible for the donorlike deep states at 0.4 E_g and 0.7 E_g in the HgCdTe.

(v) The energy to remove a Hg atom from a lattice site to the vacuum is smaller than to remove a Hg atom from a lattice site to an interstitial.

(vi) Based on preliminary calculations of the alloy variation of the substitution energies, we predict that doping of the group V elements will be easier in HgZnTe than in HgCdTe.

ACKNOWLEDGMENTS

This work has been supported by NASA Contract No. NAS1-18232 and ONR Contract No. N00014-88-C-0096.

[1]See, for example, C. W. Myles, P. F. Williams, R. A. Chapman, and E. G. Bylander, J. Appl. Phys. **57**, 5279 (1985).
[2]S. G. Louie, M. Schlüter, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B **13**, 1654 (1976).
[3]O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).
[4]W. R. Lambrecht and O. K. Andersen, Phys. Rev. B **24**, 2439 (1986); A. T. Paxton and M. van Schilfgaarde (to be published).
[5]F. Beeler, M. Schiffler, O. Jepsen, and O. Gunnarsson, Phys. Rev. Lett. **54**, 2525 (1985).
[6]M. Methfessel and J. Kubler, J. Phys. F **12**, 141 (1982).
[7]M. Brown and A. F. W. Willoughby, J. Cryst. Growth **59**, 27 (1982).
[8]H. R. Vydyanath, J. Electron Mater. **16**, 13 (1987).
[9]H. F. Schaake and J. H. Tregilgas, J. Vac. Sci. Technol. A **4**, 2181 (1986).
[10]H. R. Vydyanath, J. Electrochem. Soc. **128**, 2609 (1981).
[11]S. S. Chern, H. R. Vydyanath, and F. A. Kroger, J. Solid State Chem. **14**, 33 (1975).
[12]M. A. Berding, A.-B. Chen, and A. Sher (in preparation).
[13]C. E. Jones, K. James, J. Merz, R. Braunstein, M. Burd, M. Eetemadi, S. Hutton, and J. Drumheller, J. Vac. Sci. Technol. A **3**, 131 (1985).

ENERGETICS FOR VAPOR PHASE GROWTH MODELS OF HgCdTe

August 1990

M.A. Berding, Srinivasan Krishnamurthy, and A. Sher
SRI International
Menlo Park, CA 94025

ABSTRACT

We review our recent work on the surface binding energies for HgTe and CdTe on the (111) and ($\bar{1}\bar{1}\bar{1}$) surfaces. We find that the surface binding energies are not simply proportional to the number of surface bonds times the bulk cohesive energy per bond, but rather depend on the ionicity and band gap of the compound. From the surface binding energies for an isolated atom and an atom in the middle of an island, we have deduced an effective surface-dependent pair interaction, which is found to be attractive for HgTe and repulsive for CdTe. Concequences of these interactions on the nature of the growth are discussed.

1.0 INTRODUCTION

Molecular beam epitaxial (MBE) growth of HgCdTe has established itself as an important method for the fabrication of superlattices, abrupt heterojunctions and other novel device structures. However, growth by MBE is plagued by a number of difficulties. Material growth rate, stoichiometry and quality are determined by many parameters, including the orientation of the growth surface, substrate temperatures and the beam fluxes. In addition, dopant incorporation and activity in MBE growth differ markedly from traditional bulk growth from the liquid, and may vary with substrate orientation. Optimization of growth parameters in the absence of any microscopic insight as guidance is a difficult as well as an unsatisfying task. Much insight has been gained through experimental programs,

some of which include powerful surface-analytic techniques such as RHEED. Another approach has been to understand the MBE growth through theoretical models. The goal of these models has been to capture the essential physics of this admittedly complex process and to provide guidance into the effects of variations within some experimental parameter space. An important input to these models is the binding energies of the constituent atoms on the surface. To our knowledge, all of these models assume surface energies which are calculated by multiplying the number of bonds being made to the surface times the bulk cohesive energy per bond. We have removed this assumption.

In this paper we summarize our recent work on surface energies for growth models of the constituent compounds of HgCdTe, HgTe, and CdTe, on the (111) and ($\bar{1}\bar{1}\bar{1}$) surfaces. The results presented here have been published previously[1-3] and the reader is referred there for further detail.

2.0 THEORETICAL APPROACH

We begin by defining several surface energies of interest for the growth of a compound AB. Consider the (111) surface which is terminated with cations (A) which are triply bonded to the layer below. When an isolated anion (B) bonds to this surface it makes one bond to the surface. We refer to the energy needed as the dilute surface binding energy, $E_d^B(111)$. Similarly, the binding energy of an isolated A atom on the ($\bar{1}\bar{1}\bar{1}$) surfaces is $E_d^A(111)$. Another energy of interest in the modeling of the growth process is the energy of binding, not of an isolated atom, but rather an atom in the middle of an island of like atoms. We refer to this as the binding energy in the concentrated limit. For example the binding energy of a singly bonded B atom, in the middle of a B-atom island, to the (111) surface is denoted by $E_c^B(111)$. Thus, considering these two surface concentration limits only, we calculate eight energies for each compound (including the energies for the atoms that are triply bonded to the layer below.

From the dilute and concentrated binding energies for an atom on a given surface, we can deduce the effective surface- pair interaction energy. For example, by comparing

1. Krishnamurthy, S., M.A. Berding, A. Sher, and A.-B. Chen, *Physical Review Letters* 64, 2531 (1990).
2. Krishnamurthy, S., M.A. Berding, A. Sher, and A.-B. Chen, Accepted for publication in the *Journal of Applied Physics.*
3. Berding, M.A., S.Krishnamurthy, A. Sher, and A.-B. Chen, *Journal of Applied Physics* 67, 6175 (1990).

$E_d^B(111)$ and $E_c^B(111)$, we note that both energies include the binding to the underlying layer and their difference is in the in-plane bonds. On the (111) surface, these six in-plane bonds are second-neighbor bonds, and the in-plane pair interaction energy is given by $Ep = \frac{(E_c^B - E_d^B)}{6}$. We have calculated the surface binding energies using a tight-binding Green's function method which is especially suitable for systems with planar symmetry, such as superlattices and surfaces. This method has also been used to calculate the cleavage energies, yielding good agreement with experiment[3]. The ideal surface Green's function is perturbed by the removal or the addition of an atom to the surface for the calculation of the concentrated and dilute surface energies, respectively. The change in the density of states resulting from this perturbation is used to calculate the change in electronic energy upon removal or addition of an atom. Calculational details are given elsewhere[1,2]

3.0 RESULTS

Surface binding energies for the (111) and $(\bar{1}\bar{1}\bar{1})$ surfaces are calculated for HgTe and CdTe. Results are given in Table 1. Also shown are the bulk cohesive energies per bond, E_b. From the dilute and concentrated energies an effective surface pair interaction is deduced and these energies are also summarized in Table 1.

TABLE 1 : SURFACE BINDING AND PAIR INTERACTION ENERGIES

		$E_d(111)$	$E_c(111)$	E_p	$E_d(\bar{1}\bar{1}\bar{1})$	$E_c(\bar{1}\bar{1}\bar{1})$	E_p	E_b
		eV	eV	eV	eV	eV	eV	eV
HgTe	Hg	0.3	-1.1	-0.23	0.2	-0.3	-0.08	-0.8
	Te	-1.8	-3.8	-0.33	-2.8	-2.8	0	-0.8
CdTe	Cd	-2.7	-0.6	0.35	-2.2	-1.3	0.15	-1.1
	Te	-3.9.	-4.2	-0.05	-5.2	-2.1	0.52	-1.1

We note several features of the binding energies in Table 1 and their concequences. First we see that the surface binding energies are not simply proportional to the number of bonds being made to the substrate. This is evident from the fact that, for example, $E_c^B(111) \neq E_d^B(111) \neq 1 \cdot E_b$, one being the number of bonds being made to the surface. Also, as expected, the magnitude of the surface binding energies is larger for CdTe than for HgTe, consistent with the difference in bond strengths in these compounds, and the anions

are more strongly bound to the surface than the cations. Furthermore, we note that in some cases $E_d > E_c$, implying that an isolated atom is less strongly bound to the surface than an atom in an island, which is what one might expect based on the assumption of attractive second-neighbor bonds, while in other cases, $E_d < E_c$, implying that an isolated atom is more strongly bound to the surface. The former case correspond to $E_p < 0$ and an effective attractive surface- pair interaction, and the latter case to $E_p > 0$ and an effective repulsive surface-pair interaction

The surface pair energies in general are found to be repulsive for CdTe and attractive for HgTe. These results were explained[1] from differences in the charge transfer from cation to anion surface states, which is more pronounced in systems with larger band gaps. Furthermore we have found that the repulsive energies can be associated with more ionic compounds with large differences between the anion- and cation derived surface states.

The difference in the nature of surface interactions should manifest itself in the nature of growth surfaces at low temperatures. By low temperatures we refer to temperatures less than some critical temperature, T_c, which corresponds to some surface order-disorder transition. At temperatures above T_c, entropy will dominate and for surfaces with both attractive and repulsive effective pair energies the total surface will be largely disordered. Below the critical temperature the surface will be largely ordered, consisting of islands of atoms when $E_p < 0$ and consisting of ordered arrangements of atoms minimizing the number of in-plane atom pairs when $E_p > 0$.

4.0 CONCLUSIONS

Although the above discussion leads to some insight into the different behavior of the growth surfaces expected for HgTe and CdTe for $T < T_c$, the energies must be incorporated into a more complete growth model that includes the effects of, for example, surface-atom mobility. We can extrapolate to the alloy system $Hg_{1-x}Cd_xTe$ to deduce the nature of the surface-pair interactions there. Because all infrared applications of this alloy are for low x values and thus narrow energy band gaps, we expect the surface energies to be largely attractive, implying layer growth through island formation. Some differences can be expected between the two cation species on the surfaces. We are currently investigating these differences.

This work has been supported by ONR contract N00014-88-C0096 and NASA contract NAS1-18226.

Journal of Crystal Growth 109 (1991) 88–93
North-Holland

Epitaxially grown semiconductor surfaces

Srinivasan Krishnamurthy, M.A. Berding, A. Sher
SRI International, Menlo Park, California 94025, USA

and

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

The atomic distributions on surfaces of Si, GaAs, HgTe and CdTe are studied as functions of temperature and substrate orientation in the [100], [111] and [$\bar{1}\bar{1}\bar{1}$] directions. Surface entropy is calculated within the quasichemical approximation and the pair interaction energies are obtained using the tight-binding Green's function method. In most cases considered, the interactions between the atoms are attractive and they tend to congregate into islands for submonolayer coverage at temperatures below the roughening transition temperatures. Free energy curves for double-layer growth are presented. We find that growth in this case is mostly layer by layer as is often observed in atomic layer or molecular epitaxial growth. However, in cases such as Ga terminated ($\bar{1}\bar{1}\bar{1}$) GaAs surfaces and most CdTe surfaces, the in-plane interactions are predicted to be repulsive. The calculated order–disorder transition temperatures for these cases are often much larger than usual growth temperatures and, consequently, incompletely filled surfaces are expected to have domains in which atoms and vacancies arrange themselves in superlattice patterns.

1. Introduction

Properties of semiconductor surfaces during epitaxial growth are sensitively dependent on the nature of the interactions between surface atoms. When the interaction between the atoms is attractive, the surface grows with formation of islands at temperatures below a critical temperature known as roughness transition temperature. However, when the interaction between the atoms is repulsive, the surface undergoes an order–disorder transition. At growth temperatures below the order–disorder transition temperature, surface atoms and vacancies arrange themselves in a superlattice pattern that minimizes the number of atom-atom pairs. Irrespective of the nature of interaction, the sites are occupied uniformly at growth temperatures well above the corresponding critical temperature.

Growth by low-temperature epitaxial methods such as atomic layer epitaxy (ALE) and molecular beam epitaxy (MBE) is believed to take place under nonequilibrium growth conditions [1–4]. However, when the surface relaxation rate is much larger than the arrival rate, a thermodynamic model can be expected to qualitatively describe the morphology of the growth surface. This is because surface atoms will have enough time to minimize their free energy before the arrival of the next atom. Statistical models with various degrees of approximations are frequently used to study the nature of the surface [5–14]. To our knowledge, none of these studies has been extended to consider the surface arrangement of vacancies in a zinc-blende lattice.

In this paper, a thermodynamic model in a quasichemical approximation (QCA) with pair interactions is used to study the equilibrium multilayer growth surface. Results of multilayer growth of a model cubic lattice are given. The strength of pair interaction on various ideal semiconductor surfaces is obtained from various surface sublima-

0022-0248/91/$03.50

tion energies calculated using the tight-binding Green's function method [15,16]. The pair interaction energies are then used in this thermodynamic model to obtain the minimized free energy. The qualitative nature of realistic semiconductor surfaces is discussed.

2. Thermodymic model

We assume that total enthalpy of the surface can be written as a sum of various pair energies. In the QCA, we assume that these pairs are distributed randomly [17]. Although, in most cases, this approximation is better than the Bragg–Williams approximation, it still overestimates the entropy. The roughening transition temperature obtained for a square lattice with QCA is about 25% larger than the exact value due to Onsagar [5]. In addition, because the correlation between the pairs and beyond are assumed to be zero, the surface at the critical temperature, where large scale fluctuations exist, is not expected to be described correctly. However, epitaxial growth by ALE and MBE are normally carried out at temperatures well below the transition temperature, and the approximation considered here will remain valid. Higher order approximations with corresponding cluster energies will be the subject of future work. The results obtained using QCA with pair interaction can be obtained equivalently with Bethe's approximation [8]. However, generalization to multilayer growth problems and higher-order approximations are straightforward with QCA entropy expression used here.

To calculate the free energy within the QCA, first we identify all distinct classes of pairs that contribute to free energy. Then we write the probability for the occurrence of each type of pair. Using appropriate factors for indistinguishability of pairs, the entropy for different surface layers is written with constraints relating various pairs. The free energy obtained from the entropy and the interaction energy for each class of pairs, is minimized to get most probable distribution. In our multilayer model, an atomic site in any upper layer can be occupied even if the site immediately below it is not occupied. Although these configurations are energetically less favorable, they contribute to entropy and thus they should be allowed in the calculation of free energy. In a multilayer model, we consider four types of pairs (viz., AA, AV, VA, VV) within a layer, and the same four types of pairs with the layer below, where A and V, respectively, represent atom and vacancy. The number of pairs in each of these classes in the ith layer is denoted as M_p^i and $Q_p^{i-1,i}$ where p takes values 0, 1, 2, or 3. ϵ_p and E_p are corresponding pair energies in units of kT_g, where T_g is the growth temperature. We further denote the random probability of finding respective pairs in the ith layer as $y_p^i(0)$ and $z_p^i(0)$. η_0 and η_1 are the number of intra- and inter-plane near neighbors, respectively. Then the change in free energy, Δf_s, is

$$\Delta f_s = \sum_i \left[\sum_p \left[\tfrac{1}{2}\eta_0 y_p^i \epsilon_p + \eta_1 z_p^i \right] - \frac{1}{N} \ln W_i \right], \qquad (1)$$

where

$$W_i = \frac{N!}{N_i!\,(N-N_i)!}\, M!\, Q! \left[\prod_p \frac{\left[y_p^i(0) \right] M_p^i}{M_p^i!} \right] \left[\prod_p \frac{\left[z_p^i(0) \right]^{Q_p^{i-1,i}}}{Q_p^{i-1,i}!} \right],$$

$$y_p^i = \frac{M_p^i}{M}, \quad z_p^i = \frac{Q_p^{i-1,i}}{Q}, \quad p = 0, 1, 2, 3,$$

$$M = \frac{\eta_0 N}{2}, \quad Q = \eta_1 N,$$

$$y_0^i(0) = x_i^2,\ y_1^i(0) = y_2^i(0) = x_i(1-x_i),$$

$$y_3^i(0) = (1-x_i)^2,$$

$$z_0^i(0) = x_i x_{i-1}, \quad z_1^i(0) = x_{i-1}(1-x_i),$$

$$z_2^i(0) = x_i(1-x_{i-1}),$$

$$z_3^i(0) = (1-x_i)(1-x_{i-1}),$$

with constraints

$$2y_0^i + y_1^i + y_2^i = 2x_i, \quad 2y_3^i + y_1^i + y_2^i = 2(1-x_i),$$

$$y_1^i = y_2^i,$$

$$z_0^i + z_1^i = x_{i-1}, \quad z_0^i + z_2^i = x_i, \quad z_3^i + z_1^i = 1 - x_i,$$

$$z_3^i + z_2^i = 1 - x_{i-1}.$$

and

$$\sum_i x_i = \sum_i \frac{N_i}{N} = x.$$

Note that for every layer i, we introduce nine variables (viz., M_p^i, $Q_p^{i-1,i}$ and x_i), and eight constraints. Thus, the free energy has to be minimized in one additional variable for every new layer added. As a test case, we can use this model to study the (100) surface of a mono-atomic simple cubic solid. We further assume that atom–atom pair energy is ϵ_0 and all other energies are zero. Then we find that smooth-to-rough transition occurs at the value of 0.74 for kT_g/ϵ_0. In addition, the transition temperature changes little upon inclusion of interlayer interactions. This is in agreement with that obtained by Burton and coworkers [8], who calculated the transition temperature in Bethe's approximation. Although the surface decomposition and transition temperature change very little with the addition of a layer, the finite values obtained in our calculation of the number of various pairs in different layers will provide information on surface morphology when realistic crystal structures are examined.

3. Pair interaction energies

The model described in the previous section has been generalized to the case of diamond and zinc-blende lattices. In order to use this model to study realistic semiconductor surfaces, we need the pair interaction energies. These energies for various surface orientations are obtained using the tight-binding Green's function (GF) method. Because the details of calculational procedure have been published elsewhere [15,16,18], here the discussion will be limited to offer only continuity.

We define surface sublimation energy (SSE) as the difference between the total energy of a surface plus a free atom and that of the surface with the atom bonded to it. In this convention, a positive SSE means that atoms will require energy to go from surface to vapor. The calculation of SSE proceeds in two steps. First, the GF for an ideal atomically smooth surface is calculated in terms of bulk GF, using Dyson's expansion. Then, the modified GF with an atom added to or removed from the surface is calculated in terms of surface GF. The change density of states introduced by perturbations is used to obtain the electronic energy contribution to SSE.

The perturbation to the surface that involves the removal of an atom, leaving a surface vacancy behind, is referred to as surface sublimation from *concentrated* surface and the corresponding SSE is denoted as E_c. This removal entails breaking bonds with the layer below and in-plane, and sometimes, second-neighbor surface bonds. The perturbation to the surface that involves the removal of an atom thus leaving an ideal flat surface behind, is referred to as sublimation from *dilute* surface and the corresponding SSE is denoted as E_d. Here, the removal necessitates breaking bonds only with the layer below.

In this model, it is clear that the effective atom–atom pair interaction energy, between atoms on the surface is simply $(E_c - E_d)/\eta_0$, where η_0 is

Table 1
In-plane pair interaction energy, ϵ_0 and transition temperature; all energies are in eV; number of bonds broken with the layer below are given in parentheses; the roughening transition temperature (T_c) and order–disorder transition temperature (T_c^*) are in kelvin

Orientation	Removed atoms	Si		GaAs		CdTe		HgTe	
		ϵ_0	T_c (K)	ϵ_0	T_c (K)	ϵ_0	T_c (K)	ϵ_0	T_c (K)
(111)	a(1)	0.35	3700	0.25	2600	0.05	500	0.33	3500
	c(3)	0.37	3900	0.02	200	−0.35	1400 *	0.23	2400
($\bar{1}\bar{1}\bar{1}$)	a(3)	0.37	3900	0.07	700	−0.52	2100 *	0	0
	c(1)	0.35	3700	−0.10	410 *	−0.15	600 *	0.08	900
(100)	a(2)	0.48	3200	0.33	2200	−0.60	4000 *	0.18	1200
	c(2)	0.48	3200	0.25	1700	−0.50	3300 *	0.30	2200

the number of near-neighbor surface bonds. A positive (negative) ϵ_0 implies an attractive (repulsive) interaction between surface atoms. Calculated values of ϵ_0 are given in table 1 for (111), ($\bar{1}\bar{1}\bar{1}$) and (100) surfaces of Si, GaAs, CdTe and HgTe. Notice that the (111) surface can terminate either with triply bonded cations or with singly bonded anions. The opposite bonding arrangement is found in the [$\bar{1}\bar{1}\bar{1}$] direction.

E_c is expected to be larger than E_d, because all dangling bonds in the concentrated case interact to form a partially filled band with resulting lower energy. We see that this interpretation explains the trends observed in Si, HgTe and, in most cases, GaAs. However, in all other cases we find the in-plane surface interaction to be repulsive. The mechanism that drives some semiconductor orientations to have E_c smaller than E_d has been interpreted to be the charge transfer between the surface Fermi level and the newly exposed dangling bond states [15]. Because such a transfer is large in wide-gap ionic compounds, the surface interactions CdTe and ZnTe are predicted to be repulsive.

4. Results and discussion

In ALE, grown from a free atom vapor phase, only one type of atom species (anion or cation) is present at any given time. The atom–atom pair energies calculated in the previous section can be used in a single-layer growth model to obtain minimized free energy as function of layer coverage, x. In this model, all other pair energies are assumed to be zero. When the interaction between the atoms is attractive ($\epsilon_0 > 0$), the surface will undergo a smooth-to-rough transition as the growth temperature is varied near the critical temperature, T_c. At temperatures below T_c, the growth will take place with the formation of islands and atomic concentration in those islands will be given by location multiple minima in free energy versus x curve. At temperatures well below T_c, the minima are expected near $x = 0$ and $x = 1$. In these cases, islands formed on the surface are nearly fully occupied and the reminder of the area is nearly empty. As the surface coverage is increased, the islands grow in size with a fixed vacancy concentration until the layer is fully grown.

Because the (100) surface is a square lattice, exact value of 0.57 ϵ_0/k_B given by Onsagar [5] is used to obtain T_c. k_B is the Boltzmann constant. For the hexagonal lattice, observed in (111) orientation, the exact value of T_c is 0.91 ϵ_0/k_B [19]. The calculated T_c are given in table 1 for Si, HgTe and most GaAs surfaces. Note that the T_c is much larger than the typical growth temperature for these elements and compound (except for ($\bar{1}\bar{1}\bar{1}$) Te terminated HgTe surface) ALE growth of these materials is dominated by formation of nearly perfect islands.

When the interaction between surface atoms is repulsive, as in a Ga-terminated ($\bar{1}\bar{1}\bar{1}$) GaAs surface and all CdTe surfaces, the free energy of eq. (1) always remains negative with one minimum. However, these surfaces undergo a different kind of phase transition known as order–disorder transition. At temperatures above the critical temperature, T_c^*, the surface is disordered and sites will be occupied randomly by atoms and vacancies. At temperatures below T_c^*, the atoms and vacancies on the surface will arrange themselves in domains of ordered patterns to minimize the number of atom–atom pairs. The nature of the superlattice arrangements and area of these domains depend on the growth temperature and surface coverage. For example, at temperatures well below T_c^* for a square lattice of 50% coverage, the superlattice pattern will look like a checker board, where every near-neighbor site of atom (vacancy) is occupied by vacancies (atoms). This type of arrangement can be envisioned for a mono-layer growth in (100) orientation. T_c^* calculated using Onsagar's result for (100) surfaces and that using an exact value of 0.35 ϵ_0/k_B [20] for (111) orientations are given in table 1. We see that the calculated temperatures are well above the typical ALE or MBE growth temperatures for these compounds (except for ($\bar{1}\bar{1}\bar{1}$) Ga terminated GaAs surface). Hence the partially filled growth surface is expected to exhibit superlattice arrangements with atoms and vacancies.

In MBE growth of compounds, both types of species (anion and cation) are present, and simultaneous growth of an anion–cation double layer

can be expected. We extend our model to study the double-layer growth of semiconductors with zinc-blende lattice. The intralayer pair interaction energies deduced for the single layer growth model are still valid. The inter-layer anion–cation pair interaction energies are simply E_d/η_1, where η_1 is the number of interlayer near neighbor sites. We define an anion stabilized growth to be the case where relative anion and cation fluxes are such that any exposed surface atoms are anions. When the interaction between the atoms in the upper layer and that in the lower layer is strong (compared to kT_g), the surfaces always grow in the smooth limit. (100) surfaces of HgTe and GaAs and all surfaces of Si are predicted to fall into this category. However, when the interaction between adjacent layers are not always strongly attractive, as in the case of $(\bar{1}\bar{1}\bar{1})$ HgTe and (111) GaAs surfaces, the nature of growth depends on whether it is anion or cation stabilized.

The free energy calculated in QCA, at a growth temperature of 185°C, for Te and Hg stabilized $(\bar{1}\bar{1}\bar{1})$ surfaces is plotted respectively in figs. 1a and 1b as a function of Te and Hg surface concentration. Each corner of this plot represents the surface fully covered by the element noted there. In ALE, growth takes place along the sides. Whenever the minimum occurs between the two corners, that surface is predicted to grow in the rough limit where atoms and vacancies randomly occupy the surface sites. Whenever the minima occur near the corners, the surface is predicted to grow in the smooth limit with formation of nearly perfect islands. In MBE where both species are present in double-layer stochastic growth of compounds, the above arguments are applied to the free energy curve along the diagonal line connecting the origin and upper-right corner. Accordingly, from fig. 1, we see that, while ALE growth of Hg and Te layer takes place in the smooth and rough limits respectively, MBE growth of Te stabilized surface takes place in the smooth limit. This is because the pair energy for Te(upper layer)–Hg(lower layer) is strong, 0.9 eV [15,16]. However, the Hg stabilized surface grows in the rough limit because the Hg in the upper layer is weakly attracted to Te below (nearly 0 eV [15,16]) and atoms do not congregate into islands to maximize the cation–anion pairs. Similar behavior is expected in the growth of (111) GaAs surface.

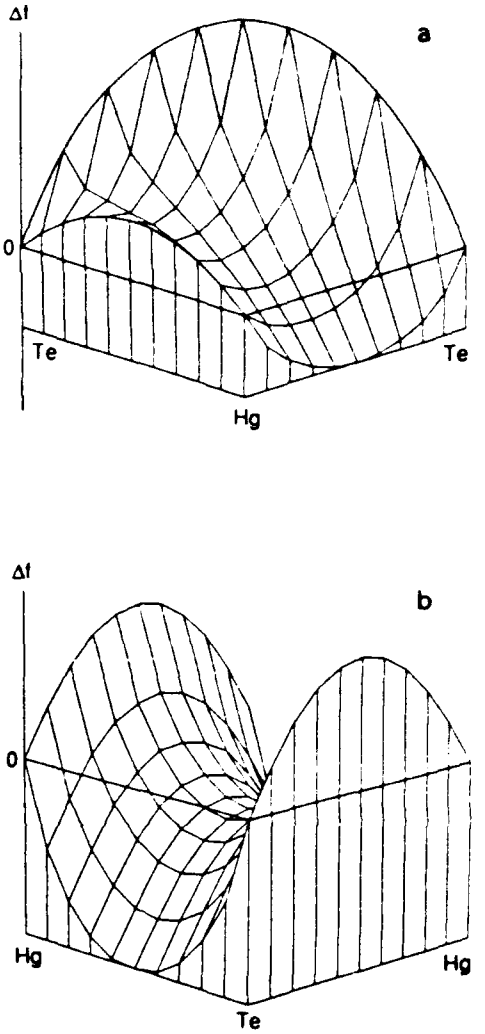


Fig. 1. Free energy for (a) Te stabilized and (b) Hg stabilized $(\bar{1}\bar{1}\bar{1})$ surface of HgTe at 185°C as a function of Hg and Te surface concentrations.

These calculations were carried out only for those cases where the interactions are attractive. For the repulsive case, it is known that QCA with pair interactions predicts incorrect phase diagrams [21]. It has been demonstrated that the smallest-sized cluster to get a correct phase diagram contains four atoms in a fcc lattice [22]. Similar

calculations with cluster variation method will have to be carried out for studies of most CdTe surfaces.

5. Conclusions

We have calculated intralayer and interlayer atom–atom pair interaction energies using a tight-binding Green's function method. These energies are then used in a thermodynamic model to study the nature of growth surfaces. As the growth temperature is varied, attractive interaction between surface atoms leads to smooth-to-rough transition, and Si, HgTe and most GaAs surfaces are calculated to fall in this category; the repulsive interactions lead to superlattice ordered-to-disordered transition, and $(\bar{1}\bar{1}\bar{1})$ GaAs and most CdTe surfaces are expected to be in this category. Smooth or rough surface growth is found to be decided by inplane interactions in ALE growth and by interplane interactions in double-layer MBE growth. In our calculations of pair energies, we considered only ideal surfaces. The critical temperatures and other conclusions in this paper, can, at best, be expected to represent the trends. Nonidealities such as reconstruction, dimerization, relaxation of adatoms and surface can lower surface energy substantially [23–26]. For quantitatively more accurate values, the effects of these mechanisms on the values of $(E_c - E_d)$ and statistical approximation that are better than QCA with pairs will have to be included.

Acknowledgments

The work was supported by NASA contract NAS1-18226, by ONR contract N00014-88-C0096, and by AFOSR contract F49620-88-K-0009.

References

[1] A. Madhukar and S. Ghaisas, CRC Critical Rev. Solid State Mater. Sci. 14 (1988) 1.
[2] R.L. Harper, J.W. Han, S. Hwang, Y. Lansari, N.C. Giles, J.W. Cook and J.F. Schetzina, J. Vacuum Sci. Technol. B7 (1988) 244.
[3] S. Sivanandan, X. Chu, J. Reno and J.P. Faurie, J. Appl Phys. 60 (1986) 1359.
[4] J. Greene, CRC Critical Rev. Solid State Mater. Sci. 11 (1983) 47.
[5] L. Onsagar, Phys. Rev. 65 (1944) 117.
[6] K.A. Jackson, Liquid Metals and solidification (American Society for Metals, Cleveland, OH, 1958) p. 174.
[7] D.E. Temkin, Soviet Phys.-Cryst. 14 (1969) 344.
[8] W.K. Burton, W. Cabrera and F.C. Frank, Phil. Trans Roy. Soc. London A243 (1951) 340.
[9] R. Kikuchi, J. Chem. Phys. 57 (1972) 4633.
[10] M. Sanchez, D. de Fontaine and W. Teitler, Phys. Rev. B26 (1982) 1465.
[11] V. Kumar and K.H. Bennemann, Phys. Rev. Letters 53 (1984) 278.
[12] H. van Beijeren, Phys. Rev. Letters 38 (1977) 993.
[13] J.D. Weeks, G.H. Gilmer and H.J. Leamy, Phys. Rev Letters 31 (1973) 549.
[14] R. Heckingbottom, J. Vacuum Sci. Technol. B3 (1985) 572.
[15] S. Krishnamurthy, M.A. Berding, A. Sher and A.-B. Chen, Phys. Rev. Letters 64 (1990) 2531.
[16] S. Krishnamurthy, M.A. Berding, A. Sher and A.-B. Chen, J. Appl. Phys., in press.
[17] A. Sher, M. van Schilfgaarde, A.-B. Chen and W. Chen, Phys Rev. B36 (1987) 4279.
[18] A.-B. Chen, Y.-M. Lai-Hsu and W. Chen, Phys. Rev. B39 (1989) 923.
[19] R.J. Baxter, Exactly Solved Models in Statistical Mechanics (Academic Press, New York, 1982) p. 308.
[20] M. Schick, J.S. Walker and M. Wortis, Phys. Rev. 1316 (1977) 2205.
[21] R. Peierls, Proc. Roy. Soc. (London) A154 (1936) 207.
[22] C.N. Yang, J. Chem. Phys. 13 (1945) 66.
[23] K.C. Pandey, Phys. Rev. Letters 49 (1982) 223.
[24] G.-X. Qian, R.M. Martin and D.J. Chadi, Phys. Rev. B38 (1988) 7649.
[25] J.E. Northrup, Phys. Rev. Letters 57 (1986) 154.
[26] D.J. Chadi, Phys. Rev. Letters 52 (1984) 1911.

Energetics of molecular-beam epitaxy models

Srinivasan Krishnamurthy, M. A. Berding, and A. Sher
SRI International, Menlo Park, California 94025

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

(Received 23 April 1990; accepted for publication 20 June 1990)

A Green's function method is used to calculate the removal energies of constituent atoms from various unreconstructed semiconductor surfaces. An efficient difference equation approach within the second-neighbor tight-binding model is used. For a compound AB, binding energies for the A and B atoms on the (111), ($\bar{1}\bar{1}\bar{1}$), (100), and (110) surfaces are calculated. Energy to remove an atom from the nearly full surface, E_c (where the removed atom leaves behind a surface vacancy), and from the nearly empty surface, E_d (where the removed atom was isolated on the surface), is obtained. Results are presented for Si, GaAs, CdTe, and HgTe. The surface sublimation energies are shown to depend on surface coverage and do not exhibit a simple linear relationship to the number of bonds broken, as is often assumed in modeling growth by molecular-beam epitaxy (MBE). Although the anion and cation extraction energies depend on surface coverage and orientation, when averaged over a double layer, they always sum to the bulk cohesive energy. Moreover, $E_c - E_d$ can be positive, implying effective attractive in-plane surface interactions, or negative, implying effective repulsive interactions. $E_c - E_d$ tends to be positive for covalent and narrow-gap semiconductors, and negative for wide-gap and more ionic semiconductors. Surface sublimation energies are important input parameters for the modeling of MBE growth; their importance is demonstrated using a simple thermodynamic growth model and results are shown to explain anomalies found in MBE growth of HgCdTe.

I. INTRODUCTION

In recent years there has been considerable advancement in the field of epitaxial growth of semiconductors, with molecular-beam epitaxy (MBE) being one of the most important techniques. Modeling of the MBE growth process involves a description of the incoming and outgoing fluxes of atoms from the surfaces, as well as in-surface migration and reactions among the constituents. An important contribution to the out-flux arises from the sublimation of the constituents from the growth surface. Attempts to determine the activation energy for surface sublimation processes have been made in a variety of experiments on GaAs[1-5] and CdTe.[6-8] Because the values obtained depend on the experimental environmental conditions,[9] the reported values differ substantially. Moreover, the experiments, in which the evaporation of many layers is permitted, measure quantities related to bulk cohesive energies rather than to the surface-orientation-dependent activation energies which are more relevant to growth modeling. Yet in the modeling of MBE growth, accurate values of the activation energies for the removal of constituent atoms from various sites on a given surface to the vapor are essential.[10,11] In this paper, we calculate the surface sublimation energy (SSE) for various classes of sites.

Several theoretical methods such as cluster methods, slab methods, and Green's function methods have been used in the literature[12-19] to study semiconductor surfaces. Both the cluster[12,13] and slab[14,15] methods approximate the semi-infinite bulk terminated at a particular surface by a finite number of atoms or two-dimensional (2D) atomic layers, respectively. While first-principles electronic structures can be incorporated into these methods, doing so is extremely time consuming, because the size of the matrix involved in these calculations is often large as a result of the range of the surface wave functions. Also, the calculations require taking differences between two large total energies. The present method differs in detail from other Green's function methods[16-19] advanced for the study of surfaces in that we take full advantage of the symmetry of the surface problem, and reduce the problem to an efficient solution of a difference equation.[20]

In this paper, we use a tight-binding (TB) Green's function (GF) method to calculate the SSE. The calculation of the SSE proceeds in two steps. First, the GF for the ideal flat surface is calculated in terms of the bulk GF using Dyson's expansion. Second, the GF for the surface with an atom either removed from or added to the ideal surface is calculated in terms of the surface GF. Unlike the cluster or slab methods, the size of the matrix equation to be solved is determined by the size of the perturbation Hamiltonian, and thus exploiting the in-plane periodicity of the unperturbed surface, the resulting matrix is comparatively small. Additionally, the electronic contribution to the total energy can be obtained from the change in the density of states (DOS) which is calculated directly from the unperturbed GF and the perturbation Hamiltonian, thus eliminating problems associated with calculating small energies by taking the differences between two large

 0021-8979/90/204020-09$03.00

energies. The DOS is calculated as a continuous function of energy, thus treating bulk and surfacelike states on an equal footing. The GF is calculated using a novel and computationally efficient difference equation approach developed recently by Chen.[20]

We calculate SSE for two classes of cations and anions for each surface. When the perturbation to the surface GF involves the removal of an atom from an ideal surface, leaving a surface vacancy behind, we refer to the process as the surface sublimation from a *concentrated* surface. The energy required to remove the atom in this limit is denoted E_c. When the perturbation to the surface GF involves the addition of an atom to the ideal surface, thus beginning a new layer of atoms, we refer to the *inverse* of this addition process as the surface sublimation from a *dilute* surface, and to the energy as E_d. Orientation-dependent E_c and E_d are energies required in modeling MBE growth.

The remainder of the paper is organized as follows. The method of obtaining the Green's functions for the bulk, the surface, and the surface with an additional atom added to or removed from the surface is described in Sec. II. In Sec. III, calculated bulk and surface DOS and surface sublimation energies of the constituent elements from (100), (111), $(\bar{1}\bar{1}\bar{1})$, and (110) surfaces of Si, GaAs, HgTe, and CdTe are discussed. To illustrate the impact of these energies on growth, they are incorporated into a simple growth model in Sec. IV. Concluding remarks are given in Sec. V.

II. GREEN'S FUNCTION METHOD

The calculation of the SSE is divided into two parts. First, the surface GF is calculated in terms of the bulk GF (which is formulated in terms of 2D Bloch states). The surface GF is calculated in terms of the bulk GF using a difference equation technique.[20] Second, the GF for the surface with an atom added to or removed from the surface is calculated in terms of the surface GF. The surface GF calculation has also been applied to the calculation of the cleavage energy,[21] and the following calculational procedure for the surface GF also appears in Ref. 21; it is included here for completeness, and to provide the groundwork for the calculation of SSE.

The bulk GF is calculated as follows. Let H_B be the Hamiltonian of the infinite periodic solid, with the corresponding eigenvalue equation

$$(H_B-E)\,|\psi\rangle=0. \tag{1}$$

For the present problem, it is convenient to decompose the three-dimensional lattice into identical 2D planar slabs, labeled by n, with their orientation determined by the particular surface to be studied. Each slab is in turn decomposed into unit cells, described by real-space lattice vectors, $\mathbf{l}_{\parallel}$ parallel to the slab interface. We expand the wave functions, $|\psi\rangle$, in terms of 2D Bloch states, $|n;\alpha\mathbf{k}_{\parallel}\rangle$:

$$|\psi\rangle=\sum_{n}\sum_{\alpha} C_n(\alpha)\,|n;\alpha\mathbf{k}_{\parallel}\rangle. \tag{2}$$

Here n sums over the planar slabs and α sums over the tight-binding basis functions within each unit cell. The $C_n(\alpha)$ are expansion coefficients to be determined. The Bloch states are defined by

$$|n;\alpha\mathbf{k}_{\parallel}\rangle=\frac{1}{\sqrt{N_s}}\sum_{l} e^{i\mathbf{k}_{\parallel}\cdot\mathbf{l}_{\parallel}}\,|n;\alpha\mathbf{l}_{\parallel}\rangle, \tag{3}$$

where the sum is over the two-dimensional lattice in the nth slab, $\mathbf{k}_{\parallel}$ is a two-dimensional Brillouin zone (BZ) wave vector, and N_s is the number of unit cells in each slab.

If the Hamiltonian contains only nearest-neighbor slab interactions, then Eqs. (1) and (2) reduce to the difference equation

$$F^{\dagger}C_{n-1}+AC_n+FC_{n+1}=0, \tag{4}$$

where the intraplanar and interplanar Hamiltonian matrices A, F, and $F^{\dagger}$ are, respectively,

$$A_{\alpha\alpha'}=\langle n;\alpha\mathbf{k}_{\parallel}\,|\,(H_B-E)\,|\,n;\alpha'\mathbf{k}_{\parallel}\rangle, \tag{5a}$$

$$F_{\alpha\alpha'}=\langle n;\alpha\mathbf{k}_{\parallel}\,|\,H_B\,|\,n+1;\alpha'\mathbf{k}_{\parallel}\rangle, \tag{5b}$$

and

$$F^{\dagger}_{\alpha\alpha'}=\langle n;\alpha\mathbf{k}_{\parallel}\,|\,H_B\,|\,n-1;\alpha'\mathbf{k}_{\parallel}\rangle. \tag{5c}$$

The general solution to Eq. (4) is given by[20]

$$C_n=\sum_{j} a_j r_j^n \beta_j, \tag{6}$$

where r_j and β_j are the jth eigenvalue and corresponding eigenvector of the following characteristic equation:

$$(F^{\dagger}+rA+r^2F)\beta=0. \tag{7}$$

The a_j of Eq. (6) are expansion coefficients determined by the boundary conditions. Letting $\gamma=r\beta$, Eq. (7) reduces to a $(2m\times 2m)$ matrix eigenvalue problem

$$\begin{pmatrix} 0 & 1 \\ -F^{-1}F^{\dagger} & -F^{-1}A \end{pmatrix}\begin{pmatrix}\gamma\\ \beta\end{pmatrix}=r\begin{pmatrix}\gamma\\ \beta\end{pmatrix}, \tag{8}$$

where m is the number of basis functions per unit cell. We use a basis set of one s and three p wave functions on each atom; thus m is four times the number of atoms in a unit cell. For example, for a zinc-blende structure with planar slabs chosen parallel to the (111) plane, the slab is chosen to contain one cation and one anion layer; thus $m=8$. The bulk GF, G^B, can be obtained similarly by solving the equation

$$(H_B-E)G^B=-I, \tag{9}$$

where H_B is the bulk Hamiltonian and I is the unit matrix. The corresponding matrix equations in the planar Bloch basis are

$$F^{\dagger}G^B_{n-1,n}+AG^B_{n,n}+FG^B_{n+1,n}=-I \tag{10a}$$

and

$$F^{\dagger}G^B_{m-1,n}+AG^B_{m,n}+FG^B_{m+1,n}=0, \tag{10b}$$

where the subscripts on G are layer indices. From the the general solution to Eq. (10), the GF of interest can be written as

$$G^B_{n,n}=[F(QXQ^{-1}-PRP^{-1})]^{-1}, \tag{11a}$$

$$G^B_{m,n}=PR^{m-n}P^{-1}G^B_{n,n},\quad m>n, \tag{11b}$$

and

$$G^B_{m,n}=QX^{m-n}Q^{-1}G^B_{n,n},\quad m<n, \tag{11c}$$

where X and R are diagonal matrices containing eigenvalues of magnitude greater than and less than unity, respectively, and Q and P are the corresponding eigenvector matrices obtained from Eq. (8). Finally, in the usual manner, the density of states (DOS) can be calculated from

$$\rho^n_B(E)=\frac{-1}{\pi}\,\mathrm{Im\ Tr}\left(\frac{1}{N_s}\sum_{\mathbf{k}_\parallel} G^B_{n,n}(\mathbf{k}_\parallel,E+i\eta)\right), \tag{12}$$

where η is vanishingly small.

From the bulk layer GF, the surface GF is obtained using Dyson's expansion. The surface is created by the bond-cutting method, in which the interactions between two semi-infinite bulks are set to zero. Hence, the Hamiltonian, H_S, for the two semi-infinite bulks separated between planes 0 and 1, is given by

$$H_S=H_B-F-F^\dagger. \tag{13}$$

Then the intraplanar surface GF, $G^S_{n,n}$, is given by

$$G^S_{n,n}=G^B_{n,n}-G^B_{n,0}F^\dagger(1+G^B_{1,0}F^\dagger)^{-1}G^B_{1,n}. \tag{14}$$

Finally, the surface DOS is given by

$$\rho^n_S(E)=\frac{-1}{\pi}\,\mathrm{Im\ Tr}\left(\frac{1}{N_s}\sum_{\mathbf{k}_\parallel} G^S_{n,n}(\mathbf{k}_\parallel,E+i\eta)\right), \tag{15}$$

and the corresponding change in DOS

$$\rho_S(E)=\sum_n(\rho^n_S-\rho^n_B). \tag{16}$$

In Eq. (16), the summation over all layers $n\geqslant l$ is found to converge rapidly away from the surface. Note that in the calculation of ρ_S both bulklike and surface-localized states, which show substantial dispersion as a result of the in-plane periodicity, are treated on an equal footing.

SSE values are calculated for both the dilute and concentrated surface coverage, as defined in Sec. I. For the dilute surface, a constituent atom is placed on the surface with bond lengths and angles corresponding to the bulk equilibrium values. Only first-neighbor interactions of the adatom with the surface atoms are included, a good approximation for covalent semiconductors. The full Hamiltonian for the surface with an isolated adatom, H_I, can then be written as

$$H_I=H_S+\sum_l(V_{0,0}+V_{0,l}+V_{l,0}), \tag{17}$$

where $V_{l,l'}$ is the matrix element between sites l' and l''. The sum goes over all nearest-neighbor sites, l, of the atom added at site 0. The change in the GF resulting from the addition of an atom to the surface is given by

$$\Delta g^n_{l,l}(E)=\sum_{l',l''} G^S_{l,l'}(E)V_{l',0}\Delta g^0_{0,0}V_{0,l''}G^S_{l'',l}(E),\quad n\neq 0, \tag{18a}$$

and

$$\Delta g^0_{0,0}(E)=\left(E-V_{0,0}\sum_{l',l''} V_{0,l'}G^S_{l',l''}(E)V_{l'',0}\right)^{-1}. \tag{18b}$$

Note that subscripts for Green's functions in Eq. (18) are now site indices for a particular layer, and n is the layer containing site l. The corresponding change in DOS is then simply given by

$$\Delta\rho_I(E)=\frac{-1}{\pi}\,\mathrm{Im\ Tr}\left(\sum_n\sum_l \Delta g^n_{l,l}\right), \tag{19}$$

where the sum is over all layers n, and is found to converge rapidly away from the surface. The sum over all sites within a layer, n, can be obtained exactly by writing the local GF in terms of Bloch Green's functions of Eq. (14) and computing the two-dimensional BZ integration. The change in DOS is given by

$$\Delta\rho_I(E)=\frac{-1}{\pi}\,\mathrm{Im\ Tr}\left\{\Delta g^0_{0,0}+\sum_n\frac{1}{N_S}\sum_{\mathbf{k}_\parallel}\left[G^S_{n,l}(\mathbf{k}_\parallel E)\times\left(\sum_{l',l''} V_{l',0}\Delta g^0_{0,0}V_{0,l''}\right)G^S_{l,n}(\mathbf{k}_\parallel,E)\right]\right\}, \tag{20}$$

where we have reintroduced the slab GF. Because we are interested in the removal of an isolated atom from the surface, the negative of Eq. (20) enters into the calculation of the SSE for a dilute surface.

In the other extreme case, in which an atom is removed from a nearly full surface (the concentrated limit), the surface sublimation energy is obtained by letting the site diagonal matrix elements of the atom at site $l=0$ go to infinity. For this perturbation, the change in the DOS is given by

$$\Delta\rho_I(E)=\frac{-1}{\pi}\,\mathrm{Im\ Tr}\left(\sum_l G^S_{l,0}(G^B_{0,0})^{-1}G^S_{0,l}\right)$$

$$=\frac{1}{\pi}\,\mathrm{Im\ Tr}\frac{\partial}{\partial E}[\ln(G^S_{0,0})]. \tag{21}$$

III. SURFACE SUBLIMATION ENERGIES

We define surface sublimation energy as the difference between the total energy of a surface plus an isolated free atom and the surface with that atom bonded to it. In the TB model, the total energy is written as the sum of the electronic energy of the occupied state and the repulsive energy V_0 between the nearest-neighbor pairs.[22] The change in the electronic energy is calculated from the change in DOS, Eqs. (20) and (21) above, via

$$\Delta E_{\mathrm{ele}}=\int_{-\infty}^{E_m}\Delta\rho_I(E)dE. \tag{22}$$

The integral over energy extends to the highest filled state, E_m, which is determined by the charge neutrality condition that

$$\sum_n = \left(\int_{-\infty}^{E_m} \rho_S^n \, dE - \int_{-\infty}^{\epsilon_f^B} \rho_B^n \, dE \right) = \pm \Delta, \qquad (23)$$

where ϵ_f^B is the highest filled bulk state (the valence-band edge in our zero-temperature calculations) and + or − sign has to be used, respectively, for cation- or anion-terminated polar surface. The quantity Δ is determined so that there is no net potential due to alternating parallel planes of charged anion and cation layers in the case of polar surfaces. Δ is $(Q_c - 3)/4$ per dangling bond per unit cell, where Q_c is the net charge on the cation in the bulk. Notice that Δ is zero for a (110) surface because the numbers of anion and cation dangling bonds on this surface are equal. E_m will in general depend on the particular surface and may differ for the complementary anion and cation termination of a polar surface. For an ideal cation-terminated surface, E_m will be at the energy level corresponding to the cation dangling bond, which is normally near the conduction-band edge. Similarly, for an ideal anion-terminated surface, E_m will be at the energy level corresponding to the anion dangling bond, which is normally near the valence-band edge. The bond length and the interatomic and intraatomic matrix elements for the atoms on the surface are assigned their bulk values, and only nonreconstructed surfaces are considered. This calculation can be easily extended to account for bond-length relaxation and surface reconstruction by using a larger unit cell. While these approximations limit our accuracy, it is the first step toward evaluating these SSE realistically.

The change in DOS due to creation of a surface is calculated for the anion- and cation-terminated (111), $(\bar{1}\,\bar{1}\,\bar{1})$, (100), and (110) surfaces for elemental silicon and for GaAs, CdTe, and HgTe compounds. In all the calculations, second-neighbor tight-binding Hamiltonians with four orbitals per atom are used to obtain the bulk electronic structures.[22–25] Slabs containing two atomic layers are defined for each surface and the resulting size of the GF matrix is 8×8 for the (111) and (100) surface orientations. In order to hold the interplanar coupling to a first-neighbor layer, the size of the basis for the (110) direction is increased to contain four atoms per unit cell, resulting in a 16×16 matrix. An analytical continuation method[26] is used to simplify the 2D BZ integration.

As an example, we present detailed results for the (100) and (110) surfaces (cleavage surfaces) of GaAs, and general results for the SSE of Si, GaAs, CdTe, and HgTe. The bulk DOS for GaAs, obtained from Eq. (12), is shown in Fig. 1. The DOS calculated using the difference-equation approach presented above is in excellent agreement with that obtained directly from the band structure.[24] The change in DOS between the surface and bulk, calculated from Eq. (16), is shown in Fig. 2 for (a) an anion-terminated and (b) a cation-terminated (100) surface of GaAs. The layer DOS was found to converge to the bulk value by the fifth slab and thus the layer sum in Eq. (16) has been carried out over only the first four slabs from the

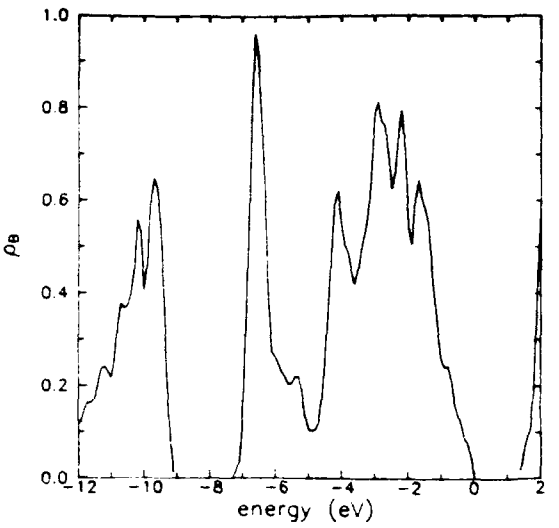


FIG. 1. Bulk DOS for GaAs in the units of states/electron volt cell. Energy is measured from the top of the valence band.

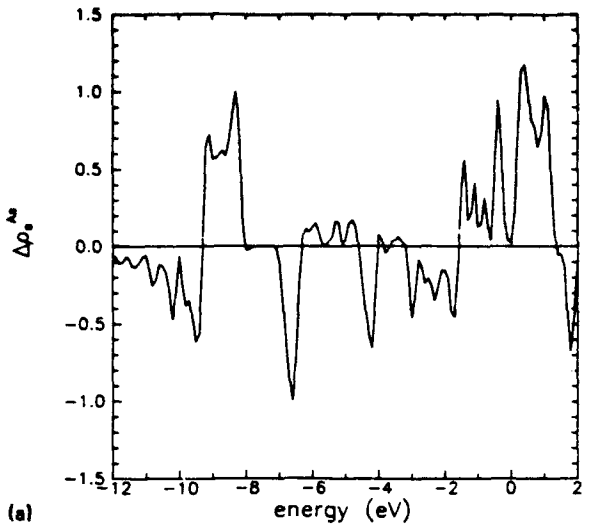
surface. In Fig. 3, the change in DOS between the surface and bulk in the vicinity of the band gap for the (110) GaAs cleavage plane is shown. Note that two states are introduced, an anionlike state at the valence-band edge and a cationlike state in the midgap region. The cleavage energies calculated from the difference between the bulk and surface DOS for the (110) surface for GaAs and the (111) surface for silicon are found to be in good agreement with experiments.[21] Finally, the change in DOS due to the removal of an isolated anion or cation from the GaAs (100) surface in the dilute limit is shown in Figs. 4 (a) and 4(b), respectively.

The change in electronic energy of the system due to the removal of an isolated atom from the surface is calculated from Eq. (22), with $\Delta\rho_l(E)$ given by Eq. (20). If the repulsive energy gained by breaking a bond is V_0, the surface sublimation energy for an atom in the dilute limit is given by

$$E_d = \Delta E_{\text{ele}} + n_b V_0 - \epsilon_{\text{atom}}. \qquad (24)$$

Here n_b is the number of bonds the adatom makes with the surface and ϵ_{atom} is the electronic energy of the free atom. V_0 is chosen so as to yield the correct bulk cohesive energy of 0.82, 1.03, 1.63, and 2.32 eV, respectively, for HgTe, CdTe, GaAs, and silicon. The SSE for an atom from the concentrated limit is calculated in a similar way, but with $\Delta\rho_l(E)$ in Eq. (22) replaced by Eq. (21).

SSEs of anions and cations from various semiconductor surfaces in silicon, GaAs, CdTe, and HgTe are given in Table I. We note several features of the SSEs: (1) the SSEs are sensitively dependent on the crystal orientation; (2) in general, they do *not* vary linearly with the number of bonds being made to the surface; (3) the SSEs for cations and anions differ and are orientation dependent; and (4) they vary considerably between the dilute and concentrated limit, even for the (111) and (100) surfaces where there are no first-neighbor in-plane bonds. The first two points are extremely important to the modeling of MBE growth,



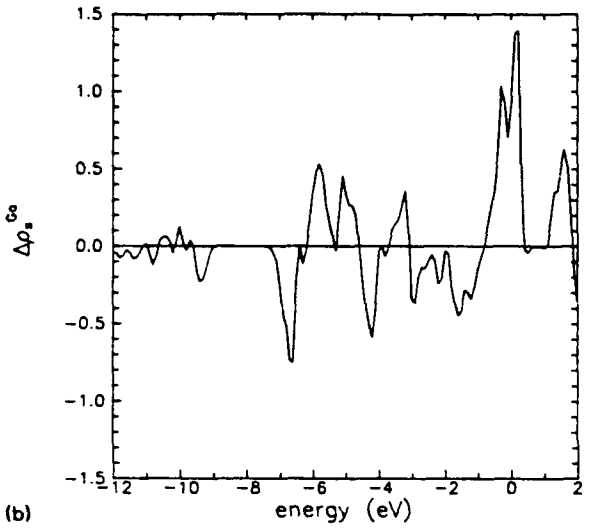


FIG. 2. (a) Difference between the surface DOS and bulk DOS for the arsenic-terminated (100) surface for GaAs; (b) difference between the surface DOS and bulk DOS for the Ga-terminated (100) surface for GaAs. In Figs. 2–4, the charge in DOS is the number of states per electron volt per cell summed over all layers.

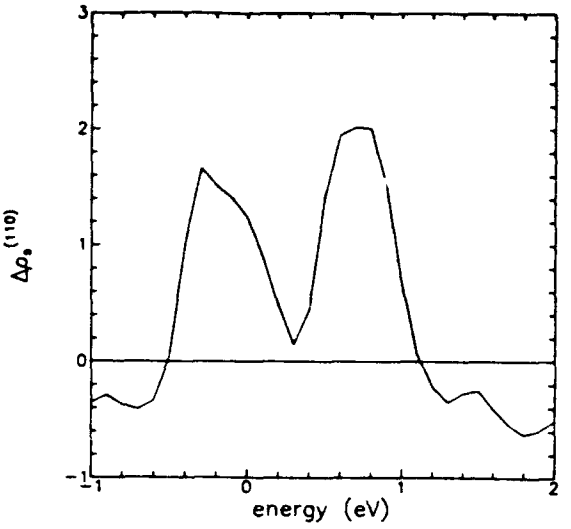


FIG. 3. Difference between the surface DOS and bulk DOS near the band gap for the (110) surface for GaAs.

where a linear dependence of the surface bonding energies with the number of bonds being made is commonly assumed.[27] Because the SSEs do not exhibit this simple linear variation with the number of bonds being made to the surface, the growth properties based on these energies are expected to differ substantially from those predicted using a linear variation (Sec. IV). Although the energy required to remove a cation or anion differs from that deduced from the bulk cohesive energy, we find that average energy per bond for the removal of an anion plus a cation layer equals the bulk cohesive energy, as it should.

Intuitively, one would expect $E_c > E_d$. The removal of an atom from the concentrated limit, in addition to breaking interlayer bonds, requires the breaking of in-plane surface bonds which are first-neighbor bonds for the (110) surface and second-neighbor bonds for the other surfaces. Even in the absence of first-neighbor bonds on the surface, the surface dangling bonds in the concentrated limit interact to form a partially filled band, which lowers energy. When a surface atom is removed, the removal of an electron from these broadened surface states will in general require more energy than if the dangling states were isolated, as in the dilute case. As seen from Table I, this conventional interpretation explains the trends observed in silicon, HgTe, and most cases in GaAs. However, for other cases, we find that E_c is smaller than E_d.

In order to understand the mechanisms that drive some semiconductors to have $E_c < E_d$, we examine the nature of the anion and cation surface states. Notice that $E_c < E_d$ occurs only for polar surfaces in semiconductors with a large band gap.

As an illustration, we choose the $(\bar{1}\bar{1}\bar{1})$ surface of CdTe where E_c is less than E_d. The surface density of states for two ideal surfaces, the Cd- and Te-terminated $(\bar{1}\bar{1}\bar{1})$, are shown in Fig. 5. For an unreconstructed surface, some of the surface states normally lie in the fundamental gap. Compounds have both cation- and anion-derived hybrid surface states, where the cation-derived most often lie higher in energy. The energy separation between the peaks in the DOS in the band-gap region is related to the energy difference between hybrid states, and thus the peak separation is larger in the II-VI than the III-V compounds.

Now let us consider the removal of a Cd atom from a $(\bar{1}\bar{1}\bar{1})$ surface in the concentrated limit. The highest occupied surface level E_m is at the cation surface state energy, 1.25 eV above the valence-band maximum. When a singly bonded cation is removed from the surface, a surface state from the anion previously bonded to the removed cation is created, and three cation surface states are destroyed. The surface density of states for the isolated anion surface state created in this process is similar to that shown by dashed line in Fig. 5 for the pure anion-terminated surface, with only minor differences in the widths and heights in the peaks. Because this anion surface state lies near the top of the valence band, the electrons from the cation surface states will transfer into this level. This process reduces the energy of the final configuration, and E_c for removal of a

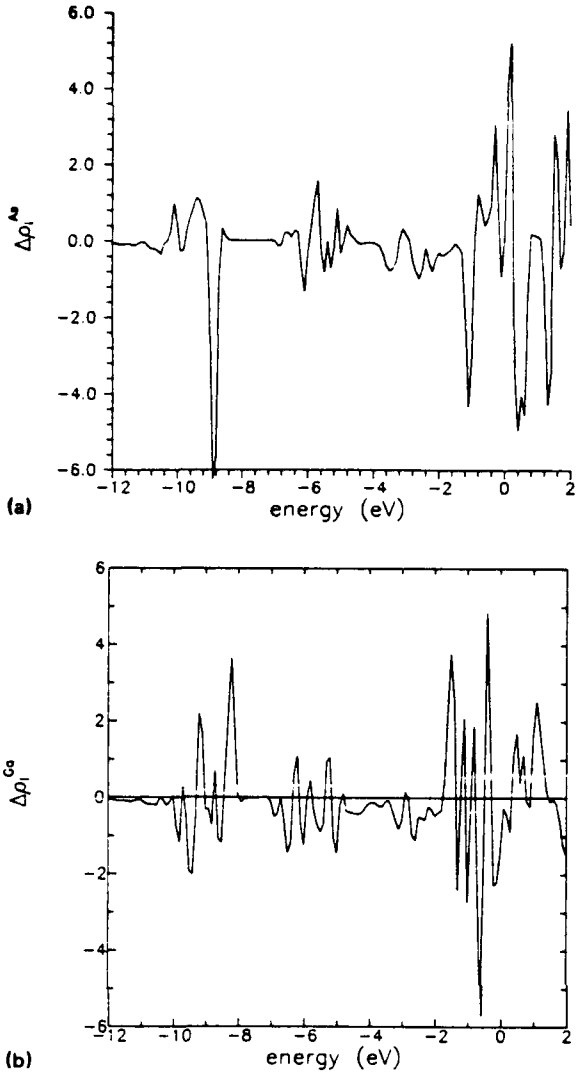


FIG. 4. (a) Change in the DOS due to the removal of an isolated gallium atom from an arsenic (100) surface. (b) Change in the DOS due to the removal of an isolated arsenic atom from a gallium (100) surface.

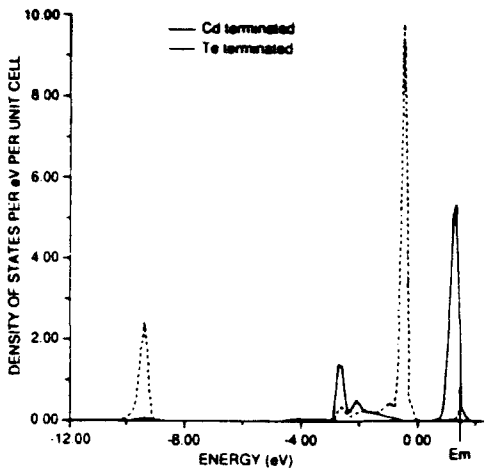


FIG. 5. Density of dangling hybrid states of anions (dashed) and cations (solid) for (111)B-oriented CdTe. E_m is the energy of the highest occupied level for a cadmium-terminated (111)B surface.

cation in this concentrated limit is correspondingly reduced. In the present case, the difference in the reduction in energies due to this charge transfer effect outweighs the surface state broadening effect, resulting in $E_c < E_d$. However, for an isolated cation removed from the $(\bar{1}\bar{1}\bar{1})$ anion-terminated surface, E_m is at the top of the valence band, which is very close to the newly exposed anion surface state. Thus, in contrast to the concentrated limit, there is a negligible reduction in E_d due to the charge transfer effect. Note that these charge transfer processes are reversed for surface sublimation of anions from the $(\bar{1}\bar{1}\bar{1})$ surface. Since three bonds are broken, more electrons are involved in the reversed spillover and the difference between E_d and E_c is correspondingly larger.

We conclude that the charge transfer from the cation to anion surface states will increase E_d and decrease E_c, and thus always reduces the difference, $E_c - E_d$. The magnitude of the charge-transfer effect depends on the amount of charge transferred and the separation between the relevant energy levels. Whenever this effect is substantial, E_c can become smaller than E_d.

In an elemental semiconductor like silicon, no transfer is expected and E_c is always expected to be larger than E_d, as is observed. In GaAs, although a charge transfer is present, the energy difference between the E_m and the newly exposed dangling bond state is very small, resulting in a reduced charge transfer effect. As a consequence, $E_c - E_d$ is reduced and, in general, remains positive. Owing to an increased ionicity in II-VI compounds, the energy separation between the dangling states is large. As illustrated above, the effect of the charge transfer is substantial in CdTe and in most cases causes E_c to be less than E_d. Although HgTe is a II-VI compound, the charge transfer does not occur because it is a semimetal with no forbidden

TABLE I. Orientation-dependent SSE and the highest occupied level (measured with respect to the top of the valence band). All energies are in electron volts. η is the number of bonds with the layer below that are broken.

Orientation	Removed atoms	η	Si			GaAs			CdTe			HgTe		
			E_m	E_d	E_c	E_m	E_d	E_c	E_m	E_d	E_c	E_m	E_d	E_c
(111)	a	1	0.0	2.5	4.6	0.4	2.3	3.8	0.7	3.9	4.2	0.0	1.8	3.8
	c	3	0.0	4.2	6.4	0.55	3.2	3.3	1.6	2.7	0.6	0.0	−0.3	1.1
$(\bar{1}\bar{1}\bar{1})$	a	3	0.0	4.2	6.4	0.0	4.6	5.0	0.0	5.2	2.1	0.0	2.8	2.8
	c	1	0.0	2.5	4.6	1.3	3.3	2.7	1.6	2.2	1.3	0.0	−0.2	0.3
(100)	a	2	0.0	3.7	5.6	0.3	3.4	4.7	0.7	5.0	2.6	0.0	2.4	3.1
	c	2	0.0	3.7	5.6	0.1	2.6	3.6	1.6	2.1	0.1	0.0	−0.2	1.0

gap. The dangling states for HgTe are resonant in the band and E_m is always at the conduction-band edge. Because the cation surface states reside in the conduction band, they are never occupied and consequently a charge-transfer effect is not present. Thus for HgTe, E_c is always larger than E_d. In general, the reduction in the difference of E_c and E_d due to the charge transfer is expected to be large for all wide-gap II-VI compounds.

The relative magnitude of the surface energies can have profound consequences on the growth properties of the semiconductor crystals. From the difference between E_c and E_d, an effective surface pair interaction can be calculated. When $E_c < E_d$, the effective surface interatom pair interaction is repulsive and the atoms can decrease their enthalpy by maximizing their interaction separation. When the effective surface interatom pair interaction is attractive, i.e., $E_c > E_d$, the isolated atoms can lower their enthalpy by coalescing and forming islands. In both cases, the actual configuration of atoms in equilibrium on the surface will be determined by total free energy. Two types of phase transitions can occur depending on whether $E_c > E_d$ or $E_c < E_d$. When $E_c > E_d$, the phase transition corresponds to the spinodal decomposition of surface atoms and surface vacancies. When $E_c < E_d$, the phase transition corresponds to an order-disorder transition, where the ordered structure is a superlattice between the atoms and vacancies on the surface. The particular surface superlattice arrangement will depend on the growth temperature, growth surface orientation, and surface atom concentration. For a special case that includes relaxation, it has been demonstrated by Chadi,[28] in agreement with experiment,[29] that a superlattice will form at $\frac{3}{4}$ Ga coverage of a (111) surface. We predict small attractive interactions in this case, but relaxation could change the sign to agree with Chadi's result.

From Table I we conclude that silicon, HgTe, and most GaAs surfaces will exhibit spinodal decomposition while most of the CdTe surfaces and the singly bonded gallium-terminated (111) surface will undergo an order-disorder transition. The actual nature of a surface during growth will be determined by the growth temperature with respect to the critical temperature. We are currently calculating the critical temperature for the various order-disorder transitions.

IV. DISCUSSION

It is difficult to compare the calculated surface sublimation energies directly with experiment. Experimental values of the sublimation energy of gallium from the (100) GaAs surface[1-5] vary from 2.9 to 4.8 eV and for the cadmium and tellurium from the CdTe (100) surface, values vary between 5 and 4.7 eV and 0.4 and 2.1 eV, respectively.[6-8] These values are normally obtained from the slope of an Arrhenius plot of excess partial pressure of one constituent versus $1/T$. The apparent discrepancy in the experimental values can be explained even using the bulk cohesive energy in the equilibrium partial pressure expression. When the cations vaporize as atoms and the anions as diatomic molecules, the product of equilibrium partial pressures of constituent materials can be obtained in terms of temperature T, atomic (molecular) mass m, vibrational energy $\hbar\omega_c$, rotational energy B_c, and cohesive energy E_b. It is given by

$$P_c^e(P_a^e)^{1/2} = (kT)^{3/2}\left(\frac{2\pi m_c kT}{h^2}\right)^{3/2}\left(\frac{4\pi m_a kT}{h^2}\right)^{3/4} \times \left(\frac{kT\,kT}{\hbar\omega_c B_e}\right)^{1/2} e^{-E_b/kT}, \tag{25}$$

where subscript $a(c)$ denotes anion (cation) and superscript e denotes equillibrium value. For known values of applied partial pressures, P_a^A and P_c^A, the excess pressure can be calculated as function of $1/T$. The slope of Arrhenius plots calculated using this analysis falls within the range of experiments. But, because the slopes thus measured are dependent on the uncontrolled overpressures of the constituents, as pointed out by Heckingbottom,[9] these slopes are not a basic property of the system. Therefore, activation energies deduced from these experiments depend strongly on particulars of the experiments and, hence, the SSE calculated in this paper cannot be compared to existing experimental values. Yet, the pressure and temperature-independent activation enthalpies calculated here are required input for MBE growth models.

To illustrate the use of the surface sublimation energies, we examine the MBE growth of $Hg_{1-x}Cd_xTe$ for small x, and demonstrate the failure of the cohesive energy in explaining one important experimental result. We do this using a simple model for MBE growth which encompasses only the necessary features to demonstrate the use of the SSEs. Our calculations predict small mercury SSE for all HgTe surfaces, in agreement with the experimental observation that $Hg_{1-x}Cd_xTe$ at low x is especially difficult to grow and requires a large flux of mercury. The calculated values of the SSEs are nearly zero for the dilute surfaces, indicating that the nucleation of new mercury layers is especially difficult. It has been experimentally observed from MBE growth at 185 °C that, for a given growth rate, the (111) surface requires an order of magnitude more mercury than the $(\bar{1}\bar{1}\bar{1})$ surface.[30] We use our calculated values of the SSE in a simple growth model to study the corresponding trends. The growth model to be described here is similar to that proposed by Galliard.[31] An improvement over the previous work is the inclusion of the change in surface entropy due to the evaporation of an atom from the surface. In this model, the growth rate R is a difference between incoming and outgoing terms:

$$(1-x_a)J^a_{Hg} - x_a^2\Theta^a_{Hg}/(1-x_a) = R,$$

$$x_a J^a_{Te} - (1-x_a^2)\Theta^a_{Te}/x_a = R, \tag{26}$$

where α is A or B corresponding to the (111) or the $(\bar{1}\bar{1}\bar{1})$ surface, x_a is the fraction of the surface covered by mercury, J_{Hg} and J_{Te} are mercury and tellurium fluxes, respectively, and Θ_{Hg} and Θ_{Te} are, respectively, the mercury and tellurium evaporation rates. This set of equations can be

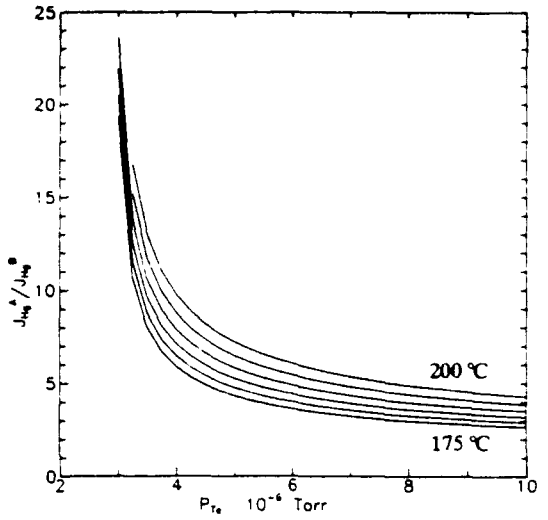


FIG. 6. Ratio of mercury vapor pressures, required for same growth rate in (111)*A* and (111)*B* directions, plotted as a function of tellurium partial pressures and the growth temperatures in steps of 5 °C.

solved for x_α and J^{α}_{Hg} for a given growth rate and various values of the tellurium overpressure and growth temperatures.

The evaporation rate Θ is assumed to be equal to $\nu e^{-E/kT}$, where the escape frequency is calculated from the spring constant of HgTe. The growth rate is assumed to be 1 μm per hour. The orientation and coverage-dependent SSE are used in the evaluation of the ratio of J^{A}_{Hg}/J^{B}_{Hg} for various values of tellurium overpressure and growth temperature. The calculated ratio J^{A}_{Hg}/J^{B}_{Hg} is plotted in Fig. 6. It can be seen that the ratio changes by an order of magnitude when the tellurium pressure is increased from 3 to 10×10^{-6} Torr. Also, note that the ratio is sensitive to growth temperature, particularly at low values of tellurium partial pressures; these conclusions fall within the range of the experiment.[30] In the only reported experiment,[30] the tellurium partial pressure, while held constant, was not recorded, so the measured ratio of 9 (at $T=185$ °C) cannot be compared directly with our calculations. However, as can be seen from Fig. 6, a factor of 9 for this ratio occurs in a reasonable range of the tellurium partial pressure. It would not be possible to obtain this result assuming a linear dependence on the cohesive energies for the SSEs. In addition, our calculations suggest that mercury incorporation can be increased by growing at low temperatures with higher tellurium partial pressures. Limits will be set by surface migration rates, here assumed to be fast compared to layer growth rates.

V. CONCLUSIONS

We have used the second-neighbor TB Hamiltonian and GF method to study the electronic structure of the ideal surfaces and those with an isolated atom added to or removed from the surface. A computationally efficient calculation is made possible by using a difference equation approach and analytical continuation technique. We have obtained the SSE for various surfaces of HgTe, CdTe, GaAs, and silicon. The calculated values imply that the SSEs do not increase linearly with the number of broken bonds. Also, the SSE depends on the surface coverage. In some cases, even when there are no nearest-neighbor surface bonds, we find it easier to remove an isolated atom than to remove an atom from a large island, which provides an additional reason to form islands. In other cases, however, we find that less energy is required to remove an atom from a concentrated surface than is needed when the atom is isolated. We have used the calculated SSEs in a simple model to study the ratio of mercury partial pressues of (111) to $(\bar{1}\bar{1}\bar{1})$ surfaces for a given growth rate. This ratio is sensitively dependent on tellurium flux and growth temperature, but our predictions fall within the range of experimental values. We further predict that better incorporation of mercury is possible at low temperature with increased tellurium flux.

Although the perturbation due to surface and adatom is treated exactly in this GF method, the accuracy of the calculated values should be improved by considering several neglected effects, including self-consistency and surface and adatom relaxation and reconstruction. It is known from detailed calculations that reconstruction,[32-35] dimerization,[34] and relaxation of adatom and surface-layer lower surface energy.[28,35] Although there is no simple relationship between the surface energy and SSE, the values calculated here will nevertheless change when these nonidealities are considered. For example, Northrup[35] has shown that the preferred position for a silicon adatom on the Si (111) surface (to which it is, ideally, to be single bonded) is the T_4 site where it bonds to three silicon atoms. This rearrangement from the ideal to the singly bonded site seems to increase SSE by 0.7 eV or so. If the surface grows reconstructed, these nonideality-limited SSE are the appropriate ones to use in modeling. While the removal of the approximations mentioned earlier will refine our reported SSEs, the underlying mechanism of charge transfer will always be present. The magnitude of effective surface interaction energies causes estimated surface order-disorder transition temperatures in excess of typical MBE growth temperatures, and consequently must impact the nature of the surface and its growth considerably.

ACKNOWLEDGMENTS

This work has been supported by National Aeronautic Space Administration Contract No. NAS1-18226, by Office of Naval Research Contract No. N00014-88-c0096, and by Air Force Office of Scientific Research Contract No. F49620-88-K-0009.

[1] T. Kojima, N. J. Kawai, T. Nakagawa, K. Ohta, T. Sakamoto, and M. Kawashima, Appl. Phys. Lett. 47, 266 (1985).
[2] J. M. Van Hove and P. I. Cohen, Appl. Phys. Lett. 47, 726 (1985).
[3] K. R. Evans, C. E. Stutz, D. K. Lorance, and R. L. Jones, J. Cryst. Growth (in press).
[4] A. J. SpringThorpe and P. Mandeville, J. Vac. Sci. Technol. B 6, 754 (1988).
[5] J. R. Arthur, J. Appl. Phys. 39, 4032 (1968).

[6]J. D. Benson, B. K. Wagner, A. Torabi, and C. J. Summers, Appl. Phys. Lett. **49**, 1034 (1986); J. D. Benson and C. J. Summers, J. Appl. Phys. **66**, 5367 (1989).
[7]J. D. Benson, D. Rajavel, B. K. Wagner, R. Benz II, and C. J. Summers, J. Cryst. Growth (in press).
[8]J. J. Dubowski, J. M. Wrobel, and D. F. Williams, Appl. Phys. Lett. **53**, 660 (1988).
[9]R. Heckingbottom, J. Vac. Sci. Technol. B **3**, 572 (1985).
[10]J. Singh and A. Madhukar, J. Vac. Sci. Technol. B **1**, 305 (1983).
[11]M. Thomson and A. Madhukar, J. Cryst. Growth **80**, 275 (1987).
[12]L. P. Batra and O. Roubaux, Surf. Sci. **49**, 653 (1975).
[13]J. Harris and G. S. Painter, Phys. Rev. Lett. **36**, 151 (1976).
[14]M. Schluter, J. R. Chelikowsky, S. G. Louie, and M. L. Cohen, Phys. Rev. B **12**, 4200 (1975).
[15]G. X. Qian and D. J. Chadi, Phys. Rev. B **35**, 1288 (1987).
[16]J. Pollmann and S. T. Pantelides, Phys. Rev. B **18**, 5524 (1978).
[17]S. L. Weng, Phys. Rev. Lett. **38**, 434 (1977).
[18]J. Ladik, Phys. Rev. B **17**, 1663 (1978).
[19]R. E. Allen, Phys. Rev. B **20**, 1454 (1979).
[20]A.-B. Chen, Y.-M. Lai-Hsu, and W. Chen, Phys. Rev. B **39**, 923 (1989).
[21]M. A. Berding, S. Krishnamurthy, A. Sher, and A.-B. Chen, J. Appl. Phys. **67**, 6175 (1990).
[22]W. A. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980), pp. 172–178.
[23]S. Krishnamurthy and J. A. Moriarty, Phys. Rev. B **32**, 1027 (1985).
[24]D. N. Talwar and C. S. Ting, Phys. Rev. B **25**, 2660 (1982).
[25]K. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 1088 (1983).
[26]K. Hass, B. Velicky, and H. Ehrenreich, Phys. Rev. B **29**, 3697 (1984).
[27]For example, T. Kawamura, A. Kobayashi, and S. Das Sarma, Phys. Rev. B **39**, 12723 (1989).
[28]D. J. Chadi, Phys. Rev. Lett. **52**, 1911 (1984).
[29]S. Y. Yong, G. Xu, and W. N. Mei, Phys. Rev. Lett. **52**, 1693 (1984).
[30]S. Sivanandan, X. Chu, J. Reno, and J. P. Fauri, J. Appl. Phys. **60**, 1359 (1986).
[31]J. P. Gailliard, Rev. Phys. Appl. **22**, 457 (1987).
[32]R. J. Jaccodine, J. Electrochem. Soc. **110**, 524 (1963).
[33]K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982).
[34]G.-X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988).
[35]J. E. Northrup, Phys. Rev. Lett. **57**, 154 (1986).

Semiconductor Surface Sublimation Energies and Atom-Atom Interactions

Srinivasan Krishnamurthy, M. A. Berding, and A. Sher
SRI International, Menlo Park, California 94025

A.-B. Chen
Department of Physics, Auburn University, Auburn, Alabama 36849
(Received 7 December 1989)

The energy required to remove an atom from semiconductor surfaces is calculated using a Green's-function approach. Contrary to intuition, we find that, in some cases, less energy is needed to remove an atom from the nearly full surface than from a nearly empty surface. The results are explained in terms of the relative energies of anion and cation dangling bonds, and the charge transfers between them. The deducted effective pair-interaction energies and their effects on surface morphology and growth perfection are discussed.

PACS numbers: 68.55.Nq, 68.35.Md

In recent years, considerable advances have been made in the field of epitaxial growth of semiconductors, with molecular-beam epitaxy (MBE) being one of the most important techniques. Modeling of the MBE-growth process entails a description of the incoming and outgoing fluxes of atoms from the surfaces, as well as the in-surface migration and reactions of the constituents. An important contribution to the outflux arises from the sublimation of the constituents from the growth surface. Attempts to determine the activation energy for surface sublimation processes have been made in a variety of experiments on GaAs (Refs. 1–3) and CdTe.[4,5] Because the values obtained depend on the experimental environmental conditions,[6] the reported values differ substantially. Moreover, the experiments, which permit evaporation of many layers, measure quantities related to bulk cohesive energies rather than to the activation energies that are more relevant to growth and that depend on surface orientation. Yet in the modeling of MBE growth, it is essential to have accurate values of the activation energies needed by the constituent atoms to escape from sites on a given surface to the vapor.[7]

In this Letter, we present the energies calculated with a Green's-function (GF) approach for atoms removed from isolated positions on the surface or from a nearly full surface. The difference in sublimation energies in these two limits is explained in terms of the energy difference between the cation and anion dangling bonds, and the charge transfers between those states. The energies affect surface order-disorder transitions and growth perfection, as discussed.

We use a second-neighbor, tight-binding Hamiltonian with four orbitals per atom to obtain the bulk electronic structures.[8–10] A tight-binding GF method is used to calculate the surface sublimation energies (SSE). The GF is calculated using a novel and computationally efficient difference-equation approach developed recently.[11] The bond length and the interatomic and intra-atomic Hamiltonian matrix elements for the atoms on the surface are assigned their bulk values, and only non-reconstructed surfaces are considered.

The calculation of these SSE proceeds in two steps. First, the GF for the ideal flat surface is calculated in terms of the bulk GF using Dyson's expansion. Bulklike and surfacelike states are treated on an equal footing, and the density of states is calculated as a continuous function of energy. The highest occupied state E_m, which can differ from the valence-band maximum, is identified by applying the condition that no net field exists in the ideal stoichiometric crystal. Second, the modified GF for the surface with an atom either removed from or added to the ideal surface is calculated in terms of the ideal surface GF. The electronic contribution to the total energy can be obtained from the change in the density of states, which is calculated directly from the unperturbed GF and the perturbed Hamiltonian. This procedure avoids problems associated with calculating small energies by taking the difference between two large energies.

The surface sublimation energy is defined as the difference between the total energy of a surface plus a free atom and that of the surface with the atom bonded to it. In this convention, the atoms with positive SSE require energy to go from the surface to vapor, and those whose SSE is less than or equal to zero will give up energy in going from the surface to vapor. In the tight-binding model, the total energy is written as the sum of the electronic energy of the occupied state and the repulsive energy between the nearest-neighbor pairs. We demand that our energy parameters produce correct bond length, cohesive energy per bond (0.82, 1.03, 1.63, and 2.32 eV, respectively, for Te, CdTe, GaAs, and Si), and band gap for the bulk. We calculate SSE for two classes of cations and anions for each surface.

When the perturbation to the surface involves the removal of an atom from the ideal surface, leaving a sur-

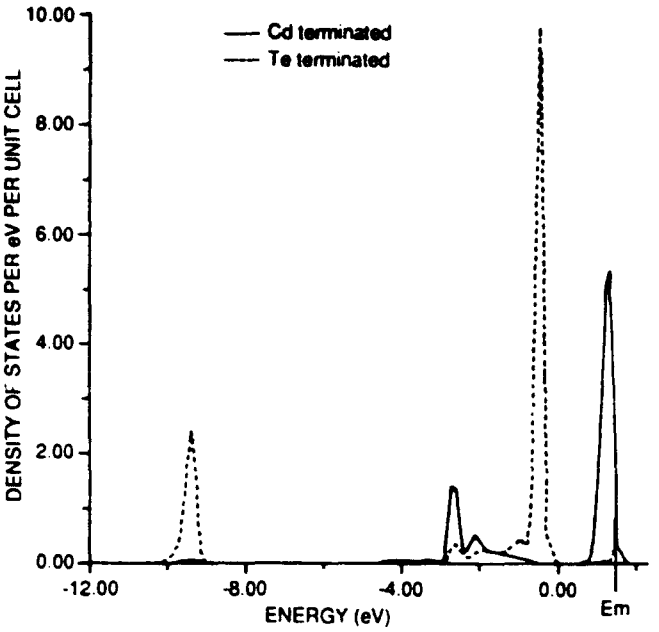


FIG. 1. Density of dangling surface hybrid states of anions (dashed line) and cations (solid line) for $(1\bar{1}1)$-oriented CdTe. E_m is the energy of the highest occupied level for a Cd-terminated $(111)B$ surface.

bonded to the removed cation is created. The surface density of states for the isolated anion surface state created is similar to that shown in Fig. 1 for the pure anion-terminated surface, with only minor differences in the widths and heights in the peaks. Because this anion surface state lies near the top of the valence band, the electrons from the cation surface states will transfer into this level.

As this process reduces the energy of the final configuration, the SSE for removal of a cation in this concentrated limit is also reduced. However, for an isolated cation removed from the underlying $(1\bar{1}1)$ anion surface, E_m is at the top of the valence band, which is very close to the newly exposed anion surface states. The reduction in the SSE from charge transfer is smaller than that in the concentrated case just considered. Apparently, the difference in the charge-transfer energies between E_d and E_c outweighs the band-broadening energy gained in E_c and results in $E_d > E_c$. The charge-transfer processes are reversed for surface sublimation of anion from the $(1\bar{1}1)$ surface. Since three bonds are broken, more electrons are involved in the reversed charge transfer, and the difference between E_d and E_c is also larger.

We conclude that the charge transfer between the cation and anion dangling states will always increase E_d and decrease E_c. The magnitude of this effect depends on the amount of charge transferred and the separation between the relevant energy levels. Whenever the effect is substantial, E_c becomes smaller than E_d and an effective repulsive surface-pair interaction results.

Because silicon contains two identical atoms in the unit cell, no charge transfer is expected and E_c is always larger than E_d. The surface-pair interactions are always attractive. Although charge transfer occurs in GaAs, the energy difference between the maximum occupied level and the newly exposed dangling-bond state is very small. The value of $E_c - E_d$ is reduced to make the surface-pair interactions weak and still attractive in most cases, but the interaction is repulsive for the Ga-terminated $(1\bar{1}1)$ surface. Because of an increased energy difference in the hybrid states in II-VI compounds, the energy separation between the dangling-bond states is large. In CdTe, for example, the effect of the charge transfer is substantial and causes E_c to be less than E_d for every case in Table I except for anions on the $(111)A$ surface. Although HgTe is a II-VI compound, charge transfer does not occur because HgTe is a semimetal with no forbidden gap. In HgTe, the dangling states are resonant in the band and the conduction-band edge is always the maximum occupied level. The cation dangling states are never occupied and, consequently, charge transfer does not occur and E_c is always larger than E_d. The reduction in the difference of E_c and E_d due to the charge transfer is expected to be large for all wide-gap II-VI compounds.

The relative magnitudes of the surface energies have profound consequences on the growth habits of these crystals. When E_c is larger than E_d, the effective interaction between the surface atoms is attractive and going from below the critical temperature to above it, the surface will undergo a phase transition from smooth to rough. The smooth surface will have islands while in the rough limit, the atoms will occupy the sites randomly. Well below the critical temperature, the islands contain few vacancies.

When E_c is smaller than E_d, however, the effective surface-pair interaction is repulsive. At growth temperatures below the order-disorder transition temperature, the atoms on the surface will arrange themselves to maximize the number of atom-vacancy bonds, resulting in superlattice arrangements. The preferred order superlattice arrangements depend on surface composition. At temperatures above the transition temperature, the arrangement of atoms will be nearly random. We conclude from Table I that Si, HgTe, and most GaAs surfaces grow in the smooth-to-rough transition region, while most CdTe and singly bonded Ga-terminated $(1\bar{1}1)$ surfaces grow in the superlattice-ordered-to-random transition region. The surface phase is determined by the growth temperature. Hence, the surface morphology is critically dependent on the nature and the magnitude of the interactions between the surface atoms.

In this Letter, we considered only ideal surfaces and illustrated that the charge-transfer mechanism will always reduce the effective atom-atom interaction on the surface. However, it is known from detailed calculations

that reconstruction,[13-15] dimerization,[14] and relaxation of adatom and the surface layer lower the surface energy substantially.[15,16] Although the effective interaction depends on the difference of $E_c E_d$, effect due to these mechanisms will have to be included for quantitatively more accurate results. While the removal of the approximations will refine our reported values, the underlying mechanism of charge transfer will always be present. For one special case, Chaid,[16] in agreement with an experiment,[17] has demonstrated that a superlattice will form at the relaxed $\frac{3}{4}$ Ga-filled (111) surface of GaAs. His predicted long-range order is mainly driven by the same kind of charge shifts as those we find in our GF method. The magnitude of effective surface energies causes the estimated order-disorder transition temperatures to be in excess of typical MBE-growth temperatures and, consequently, must impact on the nature of the surface and its growth considerably.

The financial support provided for this work through ONR Contract No. N00014-88-C0096, NASA Contract No. NAS1-18226, and AFOSR Contract No. F49620-88-K-0009 is gratefully acknowledged.

[1]T. Kojima, N. J. Kawai, T. Nakagawa, K. Ohta, T. Sakamoto, and M. Kawashima, Appl. Phys. Lett. **47**, 266 (1985).

[2]J. M. Van Hove and P. I. Cohen, Appl. Phys. Lett. **47**, 726 (1985).

[3]A. J. Springthorpe and P. Mandeville, J. Vac. Sci. Technol. B **6**, 754 (1988).

[4]J. D. Benson, B. K. Wagner, A. Torabi, and C. J. Summers, Appl. Phys. Lett. **49**, 1034 (1986).

[5]J. J. Dubowski, J. M. Wrobel, and D. F. Williams, Appl. Phys. Lett. **53**, 660 (1988).

[6]R. Heckingbottom, J. Vac. Sci. Technol. B **3**, 572 (1985).

[7]For a review, see A. Madhukar and S. V. Ghaisas, CRC Crit. Rev. Solid State Mater. Sci. **14**, 1 (1988).

[8]D. N. Talwar and C. S. Ting, Phys. Rev. B **25**, 2660 (1982).

[9]K. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 1088 (1983).

[10]S. Krishnamurthy and J. A. Moriarty, Phys. Rev. B **32**, 1027 (1985).

[11]A.-B. Chen, Y.-M. Lai-Hsu, and W. Chen, Phys. Rev. B **39**, 923 (1989).

[12]M. A. Berding, S. Krishnamurthy, A. Sher, and A.-B. Chen (to be published).

[13]K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982).

[14]G.-X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988).

[15]J. E. Northrup, Phys. Rev. Lett. **57**, 154 (1986).

[16]D. J. Chadi, Phys. Rev. Lett. **52**, 1911 (1984).

[17]S. Y. Tong, G. Xu, and W. N. Mei, Phys. Rev. Lett. **52**, 1693 (1984).

SURFACE ROUGHNESS THEORY AND LOW-TEMPERATURE EPITAXY

SRINIVASAN KRISHNAMURTHY*, M.A. BERDING*, A. SHER*, AND A.-B. CHEN**
*SRI International, Menlo Park, CA 94025
**Department of Physics, Auburn University, Auburn, AL 36839.

ABSTRACT

A single-and multilayer growth model is presented. Surface order-disorder transitions are studied with the entropy calculated in the Bragg-William approximation and in the quasi-chemical approximation. A plausible explanation for high-quality growth obtained with energy-assistance is given. The model has been extended to study low-temperature epitaxial growth of HgTe and CdTe on different surfaces. The relevant surface energies are evaluated in a Green's function approach.

INTRODUCTION

The theory of surface order of solids is essential to an understanding of their growth properties. The mechanisms dominating surface order depend on a number of factors, including the crystal orientation, substrate temperature, and growth method. The atoms arriving at a growing surface interact both with the underlying surface and with one another. Their mutual interactions may cause them to undergo an order-disorder transition from a "rough" to a "smooth" (or ordered surface) or from a "rough" to a "superlattice" (SL) surface. In the smooth limit, the newly arriving atoms tend to cluster into islands; in the disordered state, the atoms tend to be randomly distributed [1], and in the SL-ordered state, the atoms and vacancies on the surface form long-ranged ordered patterns. For a half-filled cubic surface, for example, one SL pattern is a checkerboard arrangement of atoms on the available sites. The temperature that characterizes the transition from smooth to rough surface is called *surface roughness transition temperature*, T_c. When the growth temperature T is smaller than T_c, the surface will be smooth. The growth rates and crystal quality are expected to differ in these two limits.

The rough-to-smooth transition occurs when the effective interaction between surface atoms is attractive, while the rough-to-SL transition is a consequence of repulsive surface atom-atom interactions. Because repulsive interactions among surface atoms have not been previously considered to be realistic, this type of order-disorder transition on the surface has not been treated extensively in the literature.

In this paper, we first study the surface roughness for a cubic lattice by obtaining the surface entropy contribution to the change in free energy in a random approximation [also called Bragg-Williams approximation (BWA)], and in the quasichemical approximation (QCA) with pair interactions. Based on the difference between their predictions, we suggest a plausible explanation for better quality growth obtained with energy assistance. The surface roughness model is then extended to study the surfaces of realistic semiconductors—HgTe and CdTe—grown by atomic layer epitaxy (ALE) and by molecular beam epitaxy (MBE). The transition temperatures and surface roughness for growth in various orientations are given.

SURFACE ROUGHNESS THEORY

Because of space limitations, the approach will be discussed in brief; the details can be found in the literature for the single-layer growth model with BWA [1] and for multilayer growth

Mat. Res. Soc. Symp. Proc. Vol. 161. ©1990 Materials Research Society

model with QCA [2]. We begin by considering a monatomic cubic solid grown from a (100)-oriented seed with a smooth crystal surface. The change in enthalpy and entropy are calculated to obtain the change in free energy. The most probable configuration is obtained by minimizing the free energy. The atoms arriving at the surface make one bond each with the layer below and above, and four in-plane surface bonds.

In the multilayer growth model with QCA, we first identify all distinct classes of pairs that contribute to the free energy. Secondly, we write the probability for occurence of each type of pair. Then, using appropriate factors for indistinguishability of pairs, the entropy for different layers is written with appropriate constraints. The free energy is then minimized to obtain the most probable pair distribution. In this multilayer model, an atomic site in any upper layer is allowed to be occupied, even if the site immediately below it is not occupied. In other words, "over-hangs" are permitted. Although these configurations are energetically less favorable, they contribute to the entropy, and thus should be included in the free energy. In this model, we have four types of pairs: atom-atom, atom-vacancy, vacancy-atom, and vacancy-vacancy. For each layer, there are nine variables (four interlayer pairs, four intralayer pairs, and one layer concentration) and six constraints relating them. The minimization of free energy in two of the remaining three variables can be carried out analytically. The numerical minimization in the remaining one variable is then carried out. The temperature at which the second derivative of the free energy vanishes for the surface coverage of 0.5 gives the transition temperature, T_c.

The calculated layer concentrations in the three-layer growth model are shown in Figure 1 as a function of surface coverage averaged over three layers for three growth temperatures T_1, T_c, and T_2 ($T_1 < T_c < T_2$) by thick-dashed, dashed, and solid lines, respectively. The index on the curves represent the layer number. At $\bar{x} = 1$, 3N atoms have been added to the surface, where N is the number of sites per layer. In this convention, Layer 1 is immediately above the seed surface. We note that at growth temperatures near and above T_c, all three layers grow simultaneously, but with $x_1 > x_2 > x_3$. However, at temperatures below T_c, the upper layers do not grow until the lower layers are nearly full. In this limit, layer-by-layer growth takes place.

At temperatures well below T_c, the atoms arriving at the surface cannot move and imperfect growth results. The quality of the growth can be characterized by the fraction, y_0, of atom-atom pairs to the total number of pairs in a layer. y_0 has been calculated in the BWA and QCA. Figure 2 shows that less perfect growth occurs in BWA case. However, if there is energy assistance, where photons or ions provide enough energy to permit the atoms to move on the surface, then the QCA entropy is more appropriate and the perfection improves (y_0 increases). The reason for the difference between y_0 in these cases is evident: The BWA entropy is larger, and therefore it emphasizes the drive toward complexity. This results in layers at a given coverage with more imperfections. QCA, where the entropy is smaller, predicts that layers are more nearly perfect.

When the substrate temperature is sufficiently low, the characteristic time constant for surface migration is larger than the the equilibration time for interaction of the surface with beam and gas (thought of as a heat bath). Under these conditions, BWA is more appropriate. However, if the surface mobility is enhanced with an energy assist, the effective time constant for the migration is reduced (so that surface equilibration can take place faster than equilibration with beamor gas bath), then the QCA is more appropriate. This is a plausible explanation for high quality obtained in energy-assisted epitaxial experiments [3,4]. The energy assist in these experiments is sufficient to allow the needed surface mobility.

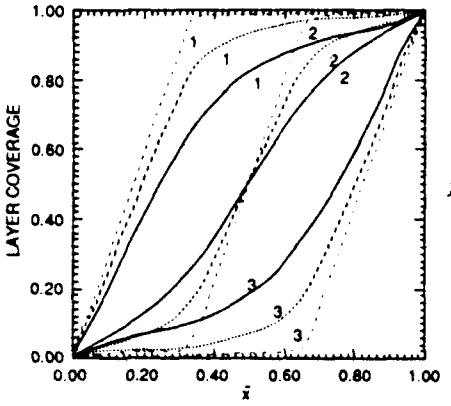


Fig. 1. Layer coverage as a function of average coverage.

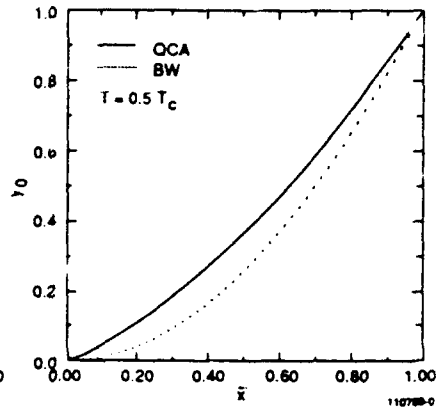


Fig. 2. Variation of y_0 with average coverage.

ENERGETICS AND GROWTH MODEL FOR SEMICONDUCTORS

The model described in the previous section has been generalized to the case of diamond and zinc blende lattices. Also, the energies of interaction between pairs on the surface and that between layers has been evaluated, using a Green's function approach. We calculate the energy to remove the atoms from nearly empty surfaces, referred as the *dilute case*, and that to remove the atoms from nearly full surfaces, referred as *concentrated case* [5]. The respective energies are denoted E_d and E_c. The calculations were carried out for the removal of anions and cations from various orientations of Si, GaAs, CdTe, and HgTe. The details of these calculations will be published elsewhere [5]. We report only HgTe and CdTe results here.

In the single-layer growth model, we consider three kinds of pairs: namely, atom-atom, atom-vacancy, and vacancy-vacancy and their respective energies E_{aa}, E_{av}, and E_{vv}. The number of effective bonds an atom makes to the substrate, η_0, and the number of surface bonds, η_1, must also be specified. The free energy and the surface-pair populations are obtained from the effective energy E_{eff}, given by $(E_{av} - (E_{aa}+E_{vv})/2)$. Without calculating E_{aa}, E_{av}, or E_{vv}, we can find E_{eff} from the difference of calculated E_c and E_d, and η_1 from Table I. In our convention, a negative E_{eff} indicates an attractive interaction between surface atoms.

In atomic layer epitaxy, grown from a free atom vapor phase, only one type of atom species (anion or cation) is present at a given time, and the growth is necessarily layer by layer. The single-layer model developed in the previous section can be applied to study HgTe and CdTe with the calculated pair energies. QCA is is used in the evaluation of surface entropies. When the pair interaction is attractive, the smooth-to-rough transition takes place, as the growth temperature is varied near T_c. However, when the pair interaction is repulsive, the surface is either rough or the atoms are in a SL arrangement with vacancies. Then the order-disorder transition takes place near a different T_c^*. The preferred ordered SL arrangements are composition-dependent; however, in regions of ordering, the surface will be arranged to maximize the number of atom-vacancy pairs. Calculated values of T_c are given for growth of HgTe and CdTe in (111) and (100) directions in Table I.

Table I. Extraction energies and critical temperatures.

Surface	Atom/ Layer	η_0	η_1	CdTe E_d [eV]	CdTe E_c [eV]	CdTe T_c [K]	HgTe E_d [eV]	HgTe E_c [eV]	HgTe T_c [K]
(111)A	a	1	6	3.9	4.2	710	1.8	3.8	4700
	c	3	6	2.7	0.6	5000*	–0.3	1.1	3300
(111)B	a	3	6	5.2	2.1	7400*	2.8	2.8	0
	c	1	6	2.2	1.3	2100*	–0.2	0.3	1200
(100)	a	2	4	5.0	2.6	5000*	2.4	3.1	1500
	c	2	4	2.1	0.1	4200*	–0.2	1.0	2500

In our calculation, the pair interaction is always attractive in HgTe and the temperatures shown in Table I indicate T_c for the rough-to-smooth transition. However, the pair interaction is repulsive for CdTe, except for the growth of a singly bonded anion layer on the (111)B surface. Temperatures T_c^* denote the rough-SL transition temperatures. Because the T_c values are very large, ALE growth of HgTe occurs mostly in the smooth domain limit and that of CdTe occurs mostly in the SL domain limit.

In layers grown by MBE, both anion and cation atoms impinge on the surface at the same time and both layers can grow simultaneously. The way in which the layer grows depends on the flux rates of the two constituents and their respective sticking coefficients. Hence, an appropriately generalized multilayer model to that described briefly in the previous section is applicable. For simplicity, we assume only double-layer growth. The interlayer pair interaction energies are obtained from E_d and η_0. The intralayer pair energies for each layer are obtained from E_c, E_d, and η_1. With the surface entropy obtained in QCA, the free energy is calculated for various values of anion and cation layer coverage. The effect of anti-site defects is not included.

Because of the repulsive intralayer pair energies, CdTe surface is expected to grow in SL domains. Hence, in addition to random distributions considered here, one must consider all ordered configurations for a given concentration. This case is being studied in detail and will not be presented here. However, the interlayer and intralayer pair interactions in HgTe are attractive, so the double-layer growth model applies to this system.

When the minima in the free energy at a given temperature occur near 0 and 1, islands are formed on the surface that are nearly fully occupied and the remainder of the area is nearly empty. As the surface coverage is increased, the islands grow in size with a fixed vacancy concentration until the layer is fully grown. The vacancies fill only at the end of the layer growth; then the formation of the next layer takes place. These events are calculated by examining a sequence of equilibrium arrangements. Realistically, kinetic effects will prevent all vacancies from filling in one layer before the start of the next layer. Thus, we expect that island growth—where the vacancy concentrations are low—is likely to correlate with high-quality material.

Sequential double-layer growth is predicted to be the case for HgTe, for example, in (111)A direction. In this case, at 185°C we predict smooth growth in ALE for both Hg and Te layers, and in either Hg or Te stabilized MBE growth [6] we also predict smooth layer growth. At this same temperature on the (111)B surface in ALE we find smooth growth for the Hg layer, but rough growth for the Te. Moreover, in this particular case, we find Te-stabilized MBE growth to be smooth, and Hg-stabilized MBE growth to be rough. In principle, it is possible to find situations and temperatures where in ALE one layer grows smooth and the other rough, and in MBE the layers grow rough for both stabilizations. This provides a rationale for the sensitivity of material quality to stabilization type, and temperature.

CONCLUSION

We have calculated the excess free energy of the surfaces for addition of several layers of atoms. The surface energies that enter the statistical mechanics are found in a Green's function approach. The entropy calculation employs a surface modification of QCA. While the calculation only finds the equilibrium arrangement of atoms for specified surface coverages, given reasonable speculations about kinetic effects, we can appreciate how energy assist functions, in principle, to improve the quality of epitaxial layers; the phenomenology dominating HgTe growth for different orientations; and reasons to expect differences between the MBE growth habits of CdTe and HgTe.

ACKNOWLEDGMENT

This work was supported in part by NASA under Contract NAS1-18226 and ONR N00014-88-C0096.

REFERENCES

1. K.A. Jackson, D.R. Uhlmann, and J. D. Hunt, J. Crys. Growth, *1*, 1 (1967).

2. S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, J. Crys. Growth (submitted).

3. R.L. Harper, J.W. Han, S. Hwang, Y. Lansari, N.C. Giles, J.W. Cook, and J.F. Schetzina, J. Vac. Sci. Tech., *B7*, 244 (1988).

4. J. Green, CRC Critical Reviews and Solidstate and Materials Sciences, *11*, 47 (1984).

5. S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, J. Appl. Phys. (to be submitted).

6. We define Hg-stabilized growth to be the case where the relative Te and Hg fluxes are such that any exposed surface atoms are Hg; in Te-stabilized growth, the converse is true.

Review of the status of computational solid-state physics

A. Sher, M. van Schilfgaarde, and M. A. Berding
SRI International, Menlo Park, California 94025

(Received 5 November 1990; accepted 18 December 1990)

The current status of *ab initio* methods to study the properties of solids is reviewed. During the past ten years, a wealth of calculations have established that many properties of solids can be accurately calculated from first principles. Recent advances in methods now make it possible to treat a number of practical problems in materials science within tolerable times and costs.

Practically every solid-state physics text begins by pointing out the difficulty of solving the Schrödinger equation for $\sim 10^{22}$ particles, and that numerous approximations are required to make progress. Twenty-five years ago, Hohenberg and Kohn[1] established the framework for modern first-principles calculations by proving that the total energy of a system was a functional only of the electron density. This functional, however, is unknown. A subsequent "local" approximation[2] made it possible to cast the problem as a set of independent particles moving in an effective "one-electron" potential for which explicit expressions could be obtained. This approach, now known as density functional theory (DFT) and the local density approximation (LDA), has evolved into a parameter-free theory capable of predicting structural and electronic properties with reasonably high precision.[3] However, until recently, the computational methods required to actually solve these equations were so slow that even on large machines calculations have been restricted to simple systems, e.g., silicon and GaAs.[4] There has been a significant advance in computational methods in the last few years, making it possible now to solve problems much more efficiently than previously.[5,6] This constitutes an impressive advance in science, but, perhaps more important, a revolution in technology. Now it is practical in systems as complicated as HgTe, CdTe, and soon the alloys $Hg_{1-x}Cd_xTe$, to compute, for example, native defect energies,[7,8] impurity substitution energies,[7,8] impurity interstitial energies,[9] and diffusion coefficients,[9] as well as the simpler bond lengths, cohesive energies, and elastic constants.[10] Accurate calculations of defect electrical activity and transport properties based on these parameter-free theories are within reach. This capability should have a dramatic impact on device processing and performance design.

The purpose of this paper is to review the status of this advance and to provide the reader with some insight into its prospects. The nature of the analysis is illustrated schematically in Fig. 1. It begins with a trial density n_{in}, often a superposition of free-atomic densities. Using the LDA (see Fig. 1), this density generates an effective one-electron potential V_{in} and a Hamiltonian, which in turn, generates an electron density n_{out}. n_{out} is mixed with n_{in} and the cycle is repeated until self-consistency is reached, i.e., $n_{out} = n_{in}$. Coulomb and magnetic interactions as well as relativistic terms are included. The Schrödinger equation must be solved in terms of a basis set. Two leading, generally applicable basis sets are the linearized augmented plane wave (LAPW)[11] and linear muffin-tin orbital (LMTO)[3] methods. For any basis set, the computational time for a solution using conventional techniques is proportional to N^3, where N is the number of basis states per unit cell. The output is the wave functions $\Psi(r)$, and the corresponding electron density $n_{out}(r)$.

The exact functional that generates V_{in} from n_{in} is not known, and in practice it is possible to do this only within the LDA, for which explicit expressions can be calculated. The LDA has been shown to produce good valence band states[3] in semiconductors and good structural properties in most materials, e.g., atomic volume, cohesive energies, and elastic constants.[10] The band gap and conduction band states are less accurately predicted in LDA. However, it has recently been shown that these properties can be accurately calculated using a technique called the "GW" approximation.[12] GW is the first term in a perturbation expansion of the exact many-body Schrödinger equation. Very good optical properties have been obtained in semiconductors when the LDA is used as a starting point to calculate GW, the Green's function G and screen Coulomb interaction W.

Among the basis sets used to solve the Schrödinger equation, the pseudopotential method is the simplest because it uses plane waves. Plane waves (PW) are suitable only for smooth potentials as they are eigenfunctions of free electrons, but they are unsuitable for systems with d electrons, such as CdTe and HgTe. Even in a simple semiconductor such as Si, 100 to 1000 PW are needed per atom, depending on the precision sought. Because the computation time varies as N^3, the direct application of this method is limited to problems with relatively few atoms per unit cell. Car and Parrinello[5] devised an improvement to this method in which the Hamiltonian itself is allowed to evolve while iterating for the eigenstates. The computation time for one iteration in their method is proportional to NM^2, where M is the number of occupied states. This advance permits much larger systems to be treated, since $M \ll N$ in that method.

A second important method is LMTO. It employs a much more efficient basis set with only ~ 20 orbitals/atom required to reach convergence. The LMTO method, until re-

$$H(V_{in}) \xrightarrow{\text{Schrodinger Eqn}} \psi(\vec{r}), n_{out}(\vec{r}) \xrightarrow[\text{DFT}]{\text{Poisson Eqn.}} V_{out}(\vec{r})$$

$$\text{Until } V_{out} = V_{in}$$

FIG. 1. Self-consistency loop.

cently, employed the atomic spheres approximation (ASA) in which within each iteration of the self-consistency loop the potential is approximated by its spherical average about the ions and some interstitial positions (empty spheres).[3] This approximation limited the applicability of the method to problems of high symmetry; it excluded problems like substitutional impurity atoms, or surfaces from those for which the precision was limited by LDA. Two years ago, a full-potential LMTO (FP-LMTO) was devised[6] that removed the ASA, thereby opening a whole new range of applications. The Car and Parrinello technique still works in FP-LMTO but does not yield much speed enhancement because the basis set is already small. There is another advance of value, however: the Harris-Foulks functional.[13,14] This is a technique for selecting a density functional that is insensitive to details of the electron density and obviates the need for the self-consistency loop. The technique increases the speed of calculations with precision still limited by LDA.

A concrete example of the speed of current calculations is presented in Table I. The example selected is a calculation of the tellurium antisite defect in CdTe using a supercell of 64 atoms. The quantities calculated are the substitution energy, the impurity levels, and strain distributions around the relaxed antisite.

In this approach, the crystal is treated as a periodic arrangement of supercells with one antisite in each cell. When the cell size is large enough so the antisites no longer interact with one another, then the calculated properties converge to those of an isolated antisite. We find that a cell size of 32 atoms is often sufficient but the numbers in Table I are for 64 atom cells.

The estimated computation times required to solve this problem on an Apollo DN10,000 computer, using a single CPU are listed in Table I. This machine is small enough to sit in a closet, costs $< \$100{,}000$, and is run by our professional staff. Its clock time is roughly 1/6 that of a Cray, but because the Cray is a vector processor, it runs this kind of problem about 30 times faster than the Apollo. Quoted times and precisions are approximate, and four iterations to self-consistency are assumed.

As you can see from the table, conventional pseudopotential methods are out of the question for a problem with 64 atoms per unit cell. However, the Car-Parrinello (for smaller problems) and LMTO methods are quite practical. A time of 100 h, or about four days, for an ASA-LMTO solution, on a machine that sits in a closet operating in the background, is perfectly acceptable. Our machine actually has three CPUs and can accommodate four, so if one person is doing a production run it does not seriously affect the rest of the group. If the Harris-Foulks functional is used, a factor of four is gained. A time of 500 h for FP-LMTO requires some patience, but even this is acceptable. Using an optimally vectorized and parallelized code can greatly reduce this time:[15] indeed, we are now testing an experimental version of the ASA program that does this calculation in about 10 h, using the three CPUs in parallel. We wish to emphasize that problems involving 64 atoms per unit cell are well outside the capabilities of earlier methods and remain time consuming even using FP-LMTO. However, a problem dealing with host material properties with only two atoms per cell, e.g., the band structure or cohesive energy $r_{in}(64/2)^3 = 3\times 10^4$ times faster, so instead of 500 h the results are obtained in about one minute.

In the past, solid-state theory was able only to calculate general trends of phenomena from first principles, or with more accuracy predict interrelations between various observables, i.e., parametrize potentials. Parametrized potentials generally can be trusted only in circumstances where the local atom arrangement deviates little from the one for which the parameters were chosen. Hence, for example, if parameters are chosen to fit bulk properties, quantitative predictions of surface reconstruction may not be trustworthy, although the symmetry may be properly given. The new *ab initio* methods reviewed here do not suffer from this kind of uncertainty and, as a consequence, can be trusted to within a given precision. This ability should be an invaluable help in interpreting observations, devising means to circumvent deleterious effects, and designing manufacturing processes.

Acknowledgments: We thank ONR (Contract Nos. N00014-89-K-0132 and N00014-88-C0096), AFOSR (Contract No. F4920-88-0009), and NASA (Contract No. NAS1-18226) for support of our program.

TABLE I. Computational effort to calculate an antisite defect in CdTe using a supercell approach (64 atoms/cell). Time is for an Apollo DN10,000, single CPU. Times and precisions are approximate. Four iterations to self-consistency are assumed.

Method	Orbitals/atom	Time (h)	Precision (eV)
Pseudopotential[a]	90	10 000	0.2
	300	$< 10^3$	LDA[b]
LAPW	50	5 000	0.2
	100	40 000	LDA[b]
Car-Parrinello[c]	300	1 000	LDA[b]
LMTO-ASA[c]	13	100	0.5
LMTO-FP	22	500	LDA[b]

[a] Pseudopotential form unsuitable for *d*-band materials.
[b] LDA error is unknown but is expected to be about 0.1 eV.
[c] Unsuitable for calculating lattice relaxations.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3] O. K. Anderson, O. Jepsen, and D. Glötzel, in *Highlights of Condensed Matter Theory*, edited by F. Bassani, P. Tosi, and P. Fimi (North Holland, Amsterdam, 1985).
[4] G. A. Baraff, M. Schlüter, and G. Allen, Phys. Rev. Lett. **50**, 739 (1988).
[5] R. Car and M. Parrinello, Phys. Rev. **55**, 2471 (1985).
[6] M. Methfessel and M. van Schilfgaarde, Phys. Rev. (to be submitted).
[7] M. Berding, M. van Schilfgaarde, A. T. Paxton, and A. Sher, J. Vac. Sci. Technol. A **8**, 1103 (1990).
[8] M. J. Caldas, J. Dabrowski, A. Fazzio, and M. Scheffler, Phys. Rev. Lett. **65**, 2046 (1990).
[9] Y. Bar-Yam and J. D. Joannopoulos, Phys. Rev. Lett. **52**, 1129 (1984).
[10] M. van Schilfgaarde, M. Methfessel, and A. T. Paxton (in preparation).
[11] S. H. Wei and H. Krakauer, Phys. Rev. Lett. **55**, 1200 (1985), and references therein.
[12] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **58**, 1551 (1987).
[13] J. Harris, Phys. Rev. B **31**, 1770 (1985).
[14] M. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).
[15] M. van Schilfgaarde and A. T. Paxton (in preparation).

Semicond. Sci. Technol. 6 (1991) C59–C70. Printed in the UK

HgCdTe status review with emphasis on correlations, native defects and diffusion†

A Sher‡, M A Berding‡, M van Schilfgaarde‡ and An-Ban Chen§

‡ SRI ernational, Menlo Park, CA, 94025 USA
§ Aub 1 University, Auburn, AL, 36849 USA

Abstract. We review the current status of knowledge of fundamental properties of the alloy $Hg_{1-x}Cd_xTe$. The most vexing questions are about its correlation state. Several different experiments now suggest it is highly correlated, but no theory predicts this result. We also discuss other properties, including dislocations at interfaces, the residual donor, worms, surface segregation and its impact on passivation, and concentration fluctuations. The forces driving these phenomena, where they are known, will be presented. Most of the paper focuses on the following: correlations; native defects, formation enthalpies and entropies; native defect equilibria with mercury gas and with tellurium inclusions; and self-diffusion coefficient activation energies including its contribution from migration energies. We will take advantage of new first-principles, high-accuracy calculations to help explain the experimental situation. The calculations predict that the main native defects found in alloys equilibrated at low Hg pressures are Hg vacancies, while at high Hg pressures they are Hg interstitials, and, surprisingly, Hg antisites.

1. Introduction

As we have been told in this conference [1], sophisticated focal plane arrays (FPAs) of sizes as large as 256 × 256 elements have been demonstrated on $Hg_{1-x}Cd_xTe$ materials grown epitaxially by several means on various substrates. Control of the material and its processing is reaching a stage where practical yields of arrays with adequate performance for some applications are nearly at hand, so these array costs are becoming acceptable. While the community is now well into a manufacturing phase, there remain many unanswered questions about the nature of $Hg_{1-x}Cd_xTe$ alloys, questions whose resolution would improve the performance, yield and stability of focal plane arrays. The purpose of this paper is to review the current status of our knowledge of this alloy system [2, 3], with the aim of suggesting potential mechanisms for the uncertain properties and additional experiments to test these hypotheses. We also will demonstrate the merit of recently developed first-principles computational methods [4] to help settle outstanding problems associated with the material which will eventually lead to accurate processing and performance models. Figure 1 depicts schematically the major issues that will be addressed. They were selected because of their potential impact on performance of FPAs. The figure shows a passivated HgCdTe sample on a substrate. On the right is a graph of a vertical composition profile through one region of the sample, and on the left is a blow-up of a 'microcluster'. The major topics identified are point defects [2], correlations [3], 'worms' [5], misfit dioslocations [6, 7], substrates, extended defects [1], inclusions [7], bulk concentration fluctuations [8] and surface segregation [9]. Other topics such as metallization [10], particularly to p-type material, could have also been included in this list, but we limit our discussion to those mentioned above. We begin by summarizing below the most important aspects of each of these topics. A selected set will be greatly expanded upon in later sections of the paper.

1.1. Point defects

Most $Hg_{1-x}Cd_xTe$ alloys are Te rich as-grown [1, 2]. The excess Te is accommodated by Hg vacancies (designated V_{Hg}) or as Te solid inclusions (Te_s) [7, 11]. The Hg vacancies are shallow acceptors (possibly negative-U-centre double acceptors) [12]. With the exception of some MBE or MOCVD material, as-grown material is p-type, both at room temperature and at 77 K [13, 14]. This material can usually be annealed at low temperature ($\sim 250\,^{\circ}C$) in an Hg overpressure to fill the V_{Hg} sites and to dissolve the Te inclusions [7]. When this is done the material converts to n-type with carrier concentrations that are remarkably low, from a few times $10^{14}\ cm^{-3}$ to a few times $10^{15}\ cm^{-3}$ in high-purity material [14]. The n dopant(s) (called the residual donors) are unknown. There is evidence [14] that the principal residual donor is an impurity rather than a native defect, but because it is so pervasive the evidence in not completely convincing. The residual donor concentration is unknown, uncon-

† This work was supported in part by NASA Contract NAS1-18226, ONR Contract N00014-88-C-0096, and AFOSR Contract 49620-88-K-0009.

0268-1242/91/120C59 + 12 $03.50

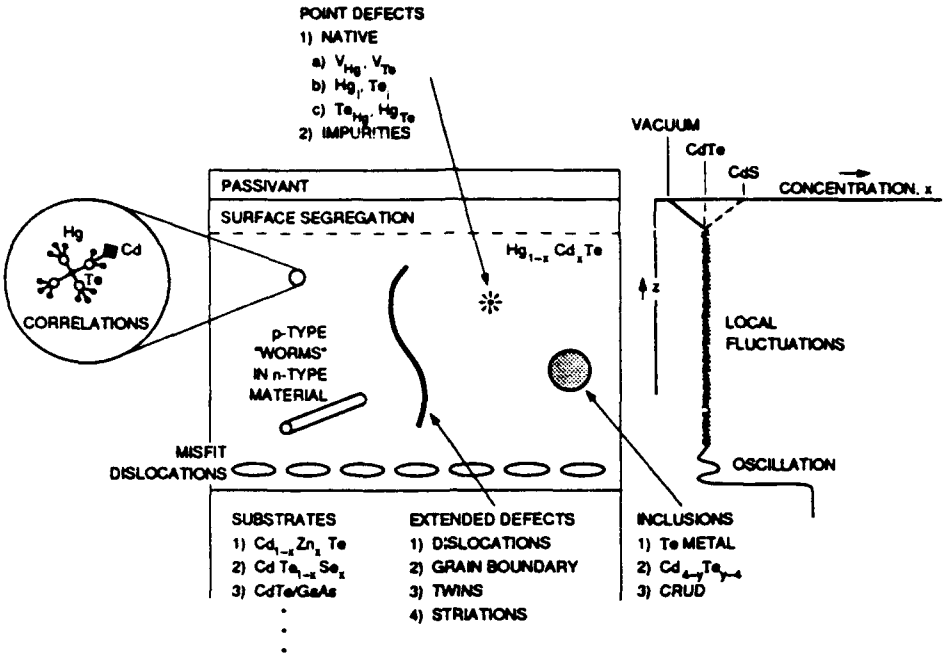


Figure 1. *Schematic features of HgCdTe alloys.*

trolled, and sets FPA limits for most operating modes. Identification and control of the residual donor would undoubtedly be beneficial. It has been demonstrated by diffusion measurements that samples annealed in higher Hg pressures contain Hg interstitials (Hg_i) [15, 16], but while they are likely to be donors, the electrical state of an Hg_i has never been directly identified. It has been speculated that Te interstitials (Te_i) and antisites (Te_{Hg}) [17] are active recombination centres, but direct evidence on their deep states is not available. The Hg antisite (Hg_{Te}) has not been previously suggested as a high-concentration defect, but we will present arguments in a later section that it may be present in concentrations sufficient to affect devices [2].

So far we have mentioned only native defects, but impurity doping is now being brought under control in HgCdTe alloys. It has proved to be difficult to gain control of both donor and acceptors, but methods are now in hand. Well behaved impurities are In [18] on cation sites serving as donors, and As on Te sites serving as acceptors. They are inserted into the material during epitaxial growth. Ion implantation and diffusion doping have not been successful enough to be used in any manufacturing process. Partial explanations for these observations have been published, but refinements are needed to make them quantitative. Many devices are still made utilizing V_{Hg} acceptors and residual donors [19].

1.2. Correlations

It is now well recognized that semiconductor alloys of the form $A_{1-x}B_xC$ are rarely truly random. The A and B atoms do not occupy their sublattice in the zinc blende structure at random but are somewhat correlated, so there is a degree of short-range order, and in some cases even long-range order [20, 21]. The principal driving terms for correlations are bond length and chemical difference between the constituents AC and BC. In $Hg_{1-x}Cd_xTe$ the bond lengths of HgTe and CdTe are nearly equal, and the differences in their chemical terms are also small, so that correlations are predicted to be small. However, a high degree of correlation has been deduced using five different experimental techniques (Te nuclear magnetic resonance (NMR) [22], Raman spectroscopy [23, 24], infrared (IR) reflection spectroscopy [25] and x-ray diffraction [26]). In the concentration range near $x = 0.25$, the correlations are such that the material is tending toward an ordered structure rather than spinodal decomposition. Those findings may have important consequences for devices, in particular if the correlation state varies spatially. Such correlation fluctuations may be a source of spatial variability of material properties that occurs even if the composition is uniform, with the variation affecting the uniformity of the bandgap, native defects and impurity concentration, mobilities, etc. The magnitudes of such variations may differ greatly from one property to another. For example, one would expect relatively slight bandgap variations, but rather more substantial differences in vacancy concentrations and diffusion coefficients between regions of different correlation states. This occurs because band structures are properties that depend on site occupations averaged over many lattice spacings, but vacancy formation energies depend sensitively on the local atom arrangement around the vacant site. Both the experimental and theoretical understanding of the correlation state of HgCdTe need to be clarified. A more detailed exposition of this problem is presented in section 2.

1.3. 'Worms'

Material grown by LPE and MOVPE and annealed to n type often exhibits low carrier concentrations $\sim 10^{15}$ cm^{-3} but with electron mobilities far below those of bulk or MBE-grown materials. If these materials are intentionally impurity doped to be slightly more n type, their mobilities return to the high values obtained by other growth methods. The currently accepted, although unconfirmed, explanation for this phenomenon is that after the Hg overpressure anneal, there are 'woodworm-like' domains in the n-type material that remain p type [5]. These regions produce p-n junctions, and mobilities measured in a Hall measurement exhibit a mixed effective mobility that appears low. Impurity doping converts these p-type worms to n type, so while there is still a spatial variation of the doping concentrations, all the material is n type so the measured Hall mobility is higher. From the viewpoint of device limits, such impurity doping has two potential deleterious effects. First the minimum carrier concentration that can be used is increased, and second there is a spatial variation of the carrier concentration.

The worms are an appealing explanation for the observed transport properties, but if the goal is to improve devices made from LPE material then we must understand why the worms form and how to eliminate them during growth, or in subsequent processing. Do the worms form because they correspond to regions of low residual donors, or to regions where the V_{Hg} formation energies are small so that they do not anneal compietely? Are the worms regions where the correlation state differs from the remaining material? Why are worms more prevalent in LPE or MOVPE material than in higher-growth-temperature bulk material or lower-growth-temperature MBE material? These and other questions are still unanswered.

1.4. Misfit dislocations

Substantial misfit dislocation densities are alv ys found at the interface between $Hg_{1-x}Cd_xTe$ and substrate materials, e.g. $Cd_{1-y}Zn_yTe$, even when the two materials are perfectly lattice matched [6, 7]. This situation also occurs to a lesser but still significant extent at heterojunctions. The reasons for these anomalous dislocation densities remain to be determined. Because dislocations have been demonstrated to have a deleterious impact on minority carrier lifetimes, their effect is of particular concern at heterojunctions where they degrade R_0A products [27-29].

1.5. Substrates

This topic has been extensively reviewed in other conference papers [1]. The ideal substrate material would be a large-area, IR-transparent insulator that is matched to the active layer by lattice constant, thermal expansion coefficient and chemical potential, and on which signal processing circuitry could be integrated. This idealization does not exist for $Hg_{1-x}Cd_xTe$, so compromises must be made. The materials CdZnTe, and GaAs or Si with buffer layers, are the current leading contenders.

1.6. Extended defects

A variety of extended defects, grain boundaries, antiphase boundaries, twin planes, threading dislocations, striations, etc have all been directly observed in transmission electron microscope (TEM) studies or their existence deduced from indirect evidence. Their impact on device characteristics is not well established, but is certainly never helpful. Because large-area FPAs are being made, extended defects obviously are relatively dilute or ineffective in current device-grade material. However, improvements are still possible.

1.7. Inclusions

We have already mentioned Te inclusions, and they will be analysed more extensively in later sections of this paper. They always have an array of dislocations around them [7], are thought to serve as getters for some impurities, and, as we suggest later, are likely to have a non-equilibrium atmosphere of V_{Hg} and Te_{Hg} around them [2]. When they are dissolved in the Hg anneal, some of their associated dislocations climb to surfaces [7] and are eliminated, although some remain; the gettered impurities may remain as inclusions, and the larger Te_s inclusions may not be completely dissolved. Means to avoid Te_s inclusion formation in as-grown material are most desirable.

1.8. Bulk concentration fluctuations

Even if an alloy is random, any segment in space has a Bernoulli probability concentration distribution [30]. Thus one expects small-scale concentration fluctuations if there is nothing to suppress them. Because of the near lattice match between HgTe and CdTe, a region of space with a concentration that differs significantly from the average produces no long-range strain field and is therefore not strain suppressed. The same effect has been found in the lattice-matched alloy $Al_{1-x}Ga_xAs$ [31]. However, in a lattice-mismatched material like $Hg_{1-x}Zn_xTe$, such short-range concentration fluctuations are suppressed. One consequence of this is that in HgCdTe the exciton line is very broad, while in HgZnTe it is narrow [8]. This effect should be particularly troublesome in VLWIR material where the bandgap is very small and fluctuations could produce semi-metal shorting domains.

1.9. Surface segregation

In general, an alloy in equilibrium will not have a uniform concentration near an interface [9]. This is most easily understood at a vacuum surface where there are dangling bonds. The material can minimize its free energy by having the alloy species that makes the weakest bonds concentrate at the surface. Also in lattice-

mismatched alloys the low-concentration species will be driven to the surface by 'strain release' because they experience less strain energy at the surface than in the bulk. In the near-lattice-matched $Hg_{1-x}Cd_xTe$ alloy, strain release plays little part, and therefore surfaces are Hg rich, because Hg makes weaker bonds to Te than does Cd. Detailed predictions are that for $x < 0.4$ the surface layer will have concentrations in the semi-metal range returning to the bulk concentration within a few atomic layers. Thus, in equilibrium, MWIR and LWIR HgCdTe alloys will have Hg-rich semi-metal surfaces. A CdTe-passivated HgCdTe alloy will have a uniform concentration up to the interface because, both in the bulk and at the interface, cations will be surrounded by four Te atoms, so there is no driving force for segregation. There is also little dr ving force for segregation at a ZnTe interface with HgCdTe. These conclusions are true only if interdiffusion is prohibited across the interface.

Because there is less tendency for interdiffusion between ZnTe and $Hg_{1-x}Cd_xTe$ than between CdTe and the alloy, we have suggested that a few atomic layers of ZnTe followed by a thicker CdTe layer should be a superior passivant.

A CdS/HgCdTe interface should be Cd rich because the CdS bond strength exceeds that of HgS by more than that between CdTe and HgTe. A similar argument applies to oxides, except that the lattice-constant mismatch becomes so large that dislocations and grain boundaries begin to exert a much larger influence on the net result.

The rapid return of the interface concentration to its bulk value (a few atomic layers) is predicted to occur only above an alloy's order-disorder transition temperatures [3, 32]. Below this temperature, long-range large-excursion concentration oscillations should occur. Because of its expected low order-disorder temperature, one would tend to discount such phenomena from consideration in HgCdTe. However, the observations of large correlation effects makes us pause. If these effects are confirmed, then there may be comparatively long-range concentration fluctuations adjacent to an epitaxial layer-substrate interface, which may help to explain the anomalous misfit dislocation densities that are observed. This suggestion (first made by Spicer *et al* [10], but without a mechanism identified) is highly speculative, and while concentration oscillations driven by surface segregation and interface boundary conditions are a natural consequence of a higher than expected order-disorder transition temperature, there is no direct evidence to support its occurrence at a HgCdTe/CdTe interface. Relevant experiments would be helpful.

The remainder of this paper will be devoted to an exposition of our current understanding of correlations, native defects and their diffusion.

2. Review of correlations

We [21] and others [33-35] have demonstrated that there are always correlations in alloy semiconductors. Here we summarize our previous work, focusing attention on those aspects that may help us to identify what may be missing from the interactions driving correlations in HgCdTe. We consider an average population $\bar{x}_j$ of five-atom anion-centred clusters including 16 bonds (most others treat four bonds only). The population of five-atom clusters of the form $A_{4-n_j}B_{n_j}C$ can be shown to be grand canonical ensembles of the form

$$\bar{x}_j = g_j \exp[(\mu_B n_j - \varepsilon_j)/kT]/q(\{\varepsilon[j]\}, \mu_B) \qquad (1)$$

where the partition function is

$$q(\{\varepsilon_j\}, \mu_B) \equiv \sum_{j=0}^{J} g_j \exp[(\mu_B n_j - \varepsilon_j)/kT] \qquad (2)$$

and g_j is the degeneracy of cluster $j = 0, 1, \ldots, J$ ($4 \leq J \leq 2^4$), n_j is the number of B atoms on cluster j, ε_j is the excess energy of cluster j relative to the virtual crystal approximation (VCA) average $\bar{\varepsilon}_{VCA} = (1-x)\varepsilon_A + x\varepsilon_B$ (called Δ_j by Sher *et al* [21]), ε_A and ε_B are cluster energies of the pure AC and BC compounds, k is the Boltzmann constant and T is the absolute temperature.

If the clusters do not have their normal counting degeneracies

$$g_j^0 = \binom{4}{n_j}$$

split, then we have $g_j = g_j^0$ the binomial coefficient, and $J = 4$. The chemical potential is set by insisting that the composition x is correct

$$4x = \sum_j n_j \bar{x}_j. \qquad (3)$$

If one makes a transformation to the reduced excess-energies representation

$$\Delta_j = \varepsilon_j - \left(1 - \frac{n_j}{4}\right)\varepsilon_0 - \frac{n_j}{4}\varepsilon_j \qquad (4)$$

(called Δ'_j by Sher *et al* [21]) then $\bar{x}_j$ becomes

$$\bar{x}_j = g_j \exp[(\mu'_j n_B - \Delta_j)kT]/q(\{\Delta j\}, \mu'_B) \qquad (5)$$

where

$$\mu'_B = \mu_B - (\varepsilon_J - \varepsilon_0)/4. \qquad (6)$$

Note that we have $\Delta_0 = \Delta_J = 0$, so at least two members of the set $\{\Delta_j\}$ vanish. More importantly, only the reduced excess energies actually drive the cluster populations. We shall see shortly that the energies of the $\{\Delta_j\}$ set are often five to ten times smaller than those of the $\{\varepsilon_j\}$ set. Most of the strain contribution to the $\{\varepsilon_j\}$ set subtracts out and contributes little to the $\bar{x}_j$ values.

There is an additional effect. It can best be appreciated when stated as a theorem: If

$$\Delta_j = \Delta_0 + n_j\delta \qquad \text{and} \qquad g_j = \binom{4}{n_j}$$

then $J = 4$ and

$$\bar{x}_j = x_j^0 \equiv g_j^0(1-x)^{(4-n_j)}x^{n_j} \qquad (7)$$

where x_j^0 are the populations of a random alloy. Thus, no matter how large the constant δ may be, as long as the

energies Δ_j vary linearly with the number of B atoms on the cluster n_j and the counting degeneracy is not split, the alloy is random. Strain energies from bond-length differences between the constituents, chemical energies and electron-electron interactions all have some nonlinear variations with n_j. The degeneracy g_j^0 is split by coherent strains, temperature gradients, or anything that establishes preferred directions for the locations of A and B atoms on a cluster. These splittings always establish a preference for one type of cluster in a particular orientation [e.g. for a strain in the (111) direction, A_3B with the B oriented along the (111) axis], and therefore drive deviations of the $\{\bar{x}_j\}$ set from $\{x_j^0\}$ toward compound formation, and, if the splittings are large enough, toward long-range order. The effect of the energies $\{\Delta_j\}$ can drive the $\{\bar{x}_j\}$ set toward either compound formation or spinodal decomposition, depending on details.

The excess free energy of an $A_{1-x}B_xC$ alloy can be written as

$$\Delta F = \Delta E - T\Delta S \tag{8}$$

where the excess energy is

$$\Delta E = M \sum_j \varepsilon_j \bar{x}_j \tag{9}$$

with M being the number of clusters. The excess entropy is

$$\Delta S = k \ln\left(\frac{N!}{N_A! N_B!} \prod_{j=0}^{J} \frac{M!}{M_j!} [g_j (1-x)^{4-n_j} x^{n_j}]^{M_j}\right) \tag{10}$$

with $M_j = x_j M$ (see Sher *et al* [21] for the justification). The total number of clusters M is related to the number of Bravais lattice sites N in the crystal by $M = N/4$ for a 16-bond cluster (and $M = N$ for a four-bond cluster). Equations (9) and (10) can be rewritten for the 16-bond cluster as

$$\Delta E = \frac{N}{4}\left((1-x)\,\varepsilon_0 + x\varepsilon_J + \sum_j \Delta_j \bar{x}_j\right) \tag{11}$$

and

$$\begin{aligned}\Delta S = &- kN\Big((1-x)\ln(1-x) + x \ln x \\ &+ \frac{M}{N}\sum_j (\bar{x}_j \ln \bar{x}_j - \bar{x}_j \ln x_j^0)\Big) \\ = &\frac{N}{4T}\sum_j \Delta_j \bar{x}_j - N\frac{\mu_B}{T}x + \frac{N}{4}\ln q(\{\Delta_j\}, \mu_B'). \end{aligned} \tag{12}$$

Note in the first equality in equation (12) that M/N is 1 for a four-bond and 1/4 for a 16-bond cluster. In the 16-bond cluster case, there is, as a consequence, an exact cancellation of the $(1-x)\ln(1-x) + x \ln x$ term by a portion of the third term in the bracket. This occurs physically because, in the 16-bond cluster, all the bonds associated with a given substituted atom are in the cluster. In the four-bond case, however, bonds from each substituted atom contribute to four different clusters. Several features of equations (11) and (12) are noteworthy. First, ε_0 and ε_J are functions of x but are temperature independent. For bond-length-mismatched alloys, they vary with x roughly as $\varepsilon_0 \approx \varepsilon_0^0 x^2$ and $\varepsilon_J \approx \varepsilon_J^0 (1-x)^2$, so that we have

$$\Delta E \approx N\left(\Omega\, x(1-x) + \sum_j \Delta_j \bar{x}_j\right) \tag{13}$$

where the mixing enthalpy parameter, Ω is approximately

$$\Omega \approx [\varepsilon_J^0 + (\varepsilon_0^0 - \varepsilon_J^0)x]/4 \tag{14}$$

and nearly x independent if $\varepsilon_0^0 \approx \varepsilon_J^0$, as is often the case. The second term in equation (13) is usually small compared with the first and contains all the temperature dependence. However, when equations (5), (12) and (13) are inserted into equation (8) to obtain ΔF, a term in $T\Delta S$ exactly cancels the second term in equation (13). Thus, it is impossible to determine the temperature variation of ΔE from a measurement of ΔF. The expression for ΔF becomes

$$\Delta F = N[\Omega\, x(1-x) + x\mu_B' - \tfrac{1}{4}kT \ln q(\{\Delta_j\}, \mu_B')]. \tag{15}$$

The chemical potential μ_B' is determined from equation (3), and both μ_B' and q are generally temperature dependent.

We have calculated $\{\varepsilon_j\}$, $\{\Delta_j\}$, μ_B', $\{\bar{x}_j - x_j^0\}$, $\Delta E(T)$ and $\Delta F(T)$ for $Hg_{1-x}Cd_xTe$ (figure 2) and $Hg_{1-x}Zn_xTe$ (figure 3) alloys. The energies ε_j for these figures were calculated as discussed in detail in [21]. The 16-bond cluster was attached to a rigid medium at the virtual crystal positions in the third shell from the cluster centre. The atom positions inside the third shell were adjusted to minimize the strain and chemical energies. The energy ε_j is the total energy of the 16 bonds with the atoms in their minimum-energy positions. The variations of atom bond lengths in different alloys are well predicted by this method. The ΔF curves for $Hg_{1-x}Cd_xTe$ show no tendency for an order-disorder transition in the temperature range studied, but $\bar{x}_j - x_j^0$ values do deviate from zero. $Hg_{1-x}Zn_xTe$ does have an order-disorder spinodal transition with critical [36] temperature $T_c \cong 380$ K. Although $Hg_{1-x}Cd_xTe$ shows significant deviations of the $\bar{x}_j$ from random alloy values for a sample equilibrated at 300 K, they are still not as large as those measured in the NMR experiments, meaning that a major interaction may be missing. However, at 300 K diffusion may be fast enough in HgCdTe for the material to equilibrate even at this low temperature (see the discussion in the next section). The slower diffusion observed in HgZnTe may be a distinct advantage, one that counters its above-room-temperature critical temperature.

We have shown that only the small $\{\Delta_j\}$ drive deviations in the populations $\bar{x}_j$ away from randomness. Moreover, if $\Delta_j = n_j\delta$ is linear in n_j the number of B atoms, then despite the size of δ, the populations are still random. Thus, effects that would otherwise be considered small may compete with the larger energies retained if they have the proper nonlinearity. We are examining several possibilities in $Hg_{1-x}Cd_xTe$. These include effects caused by screening in the composition range where the alloy is a semi-metal, direct second-neighbour chemical

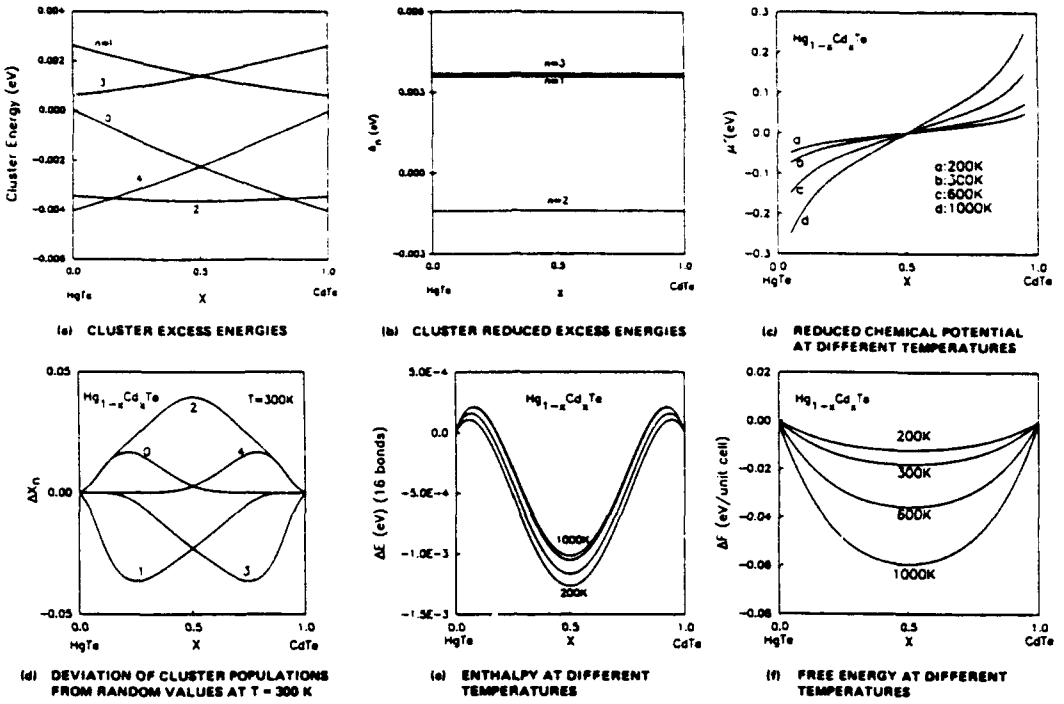


Figure 2. Composition variation of correlation-state-dependent quantities for $Hg_{1-x}Cd_xTe$.

interactions treated in the context of Harrison's model [37], electron-electron Coulomb interactions driven by polarity differences between the constituents, and composition variations of the elastic constants.

3. Native defects and diffusion

3.1. Data fitting

There have been several measurements [15, 16] of the Hg vacancy formation energy and Hg, Cd and Te tracer diffusion coefficients. These measurements are always fitted to a functional form

$$G = G_\infty(P)\exp(-E'/kT) \tag{16}$$

where E' is interpreted as the activation energy and $G_\infty(P)$ is the infinite temperature limit of G. For vacancy formation, G represents the vacancy concentration, and for diffusion the infinite temperature diffusion coefficient. It is assumed in the fitting process that $G_\infty(P)$ can be viewed as a function only of the partial pressure P above the sample of the species under investigation. However, the pre-exponential coefficient always has a power-law temperature contribution, i.e. a T^m multiplicative component. Because the measurements extend over a limited temperature range, $T_1 \le T \le T_2$, usually have some scatter, and the T^m term varies more slowly than the exponential, reasonable fits to the data are obtained. However, when E' values are quoted to more than one significant figure, and G_∞ to the proper order of magnitude, it is important to include the T^m factor in the fitting procedure. Because we do not have access to the original data we have corrected published numbers by assuming the straight-line fits of ln G versus $1/T$ were made through the points T_3 and T_4 given by

$$T_3^{-1} = \tfrac{1}{3}(2T_1^{-1} + T_2^{-1})$$
$$T_4^{-1} = \tfrac{1}{3}(T_1^{-1} + 2T_2^{-1}) \tag{17}$$

Then the correct expression

$$G = C(P)T^m \exp(-E/kT) \tag{18}$$

parameters are related to the fitted parameters by the relations

$$E = E' - mk\tau \tag{19}$$

and

$$\frac{1}{m}\ln\frac{G_x}{CT_4^m} = \frac{\tau}{T_4} \tag{20}$$

where

$$\tau \equiv (T_3^{-1} + T_4^{-1})^{-1}\ln(T_4/T_3) \cong T_4. \tag{21}$$

The approximate expression to τ in equation (21) is accurate if $T_4 - T_3 \ll T_4$. Note that if m is positive $E < E'$ and if m is negative $E > E'$.

To illustrate the effect of these corrections, we have examined the case of V_{Hg} formation and diffusion and the interstitial Hg_i diffusion in $Hg_{0.78}Cd_{0.22}Te$. The mass action equation for formation of V_{Hg}

$$HgTe \Leftrightarrow V_{Hg}Te + Hg_g \tag{22}$$

where Hg represents a Hg atom on a Hg site, and Hg_g represents Hg in the vapour phase, has been analysed

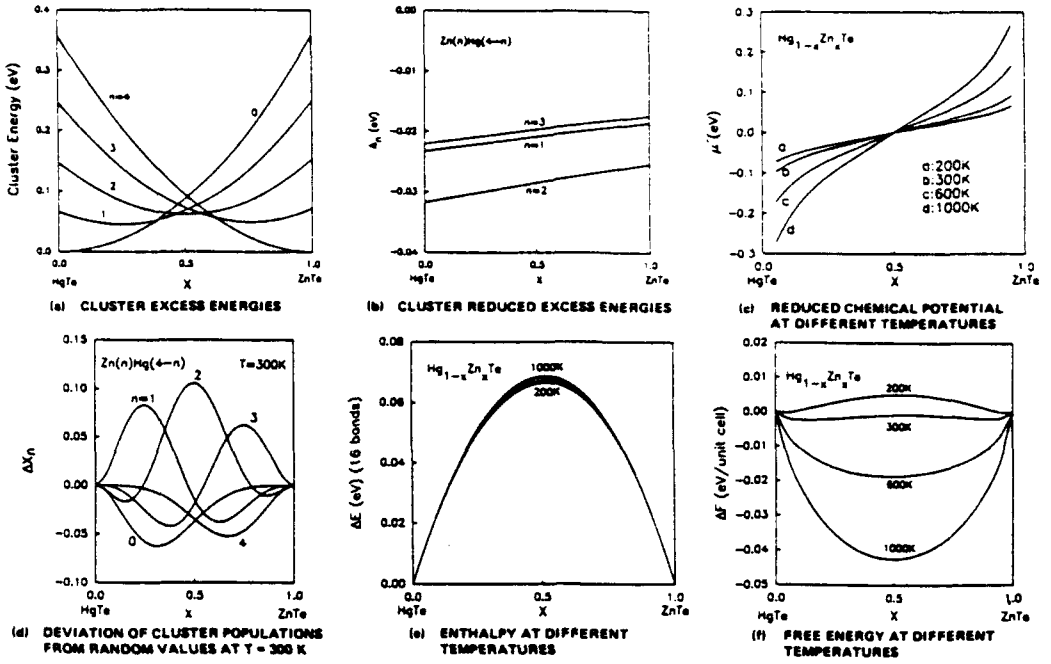


Figure 3. Composition variation of correlation-state-dependent quantities for $Hg_{1-x}Zn_xTe$.

using standard results in statistical mechanics. The diffusion coefficient D is given by

$$D = \frac{1}{6}\frac{a^2}{\tau_c}\left(\frac{n}{N}\right) \tag{23}$$

where

$$\frac{1}{\tau_c} = \omega_p \tfrac{1}{2} f_{I'} \tag{24}$$

and where a is the hop distance, τ_c is the mean free time between hops, n/N is the fraction of lattice sites N occupied by vacancies, ω_p is the presentation frequency of an atom at the barrier between it and an adjacent vacant site and $f_{I'}$ is the probability that an adjacent atom reaches the intermediate position at the peak of the barrier between it and an adjacent vacancy. (The prime is inserted to remind us that this is not a normal interstitial because it lies between two vacant sites.)

The quantity $f_{I'}$ is calculated roughly from the mass action equation for the effective reaction

$$HgTe + V_{Hg_I} \Leftrightarrow V_{Hg}Te + Hg_{I'} \tag{25}$$

again remembering that the HgTe is adjacent to a V_{Hg} in the initial state.

Then, the assumption that the $Hg_{I'}$ in the intermediate position does not remain long enough for its vibrational states to thermally equilibrate leads to the power-law m values quoted in table 1.

There are several noteworthy features to the information in table 1. First, the corrections to the activation energies are 10 to 20%, which is important if the second and third significant figures quoted are to be taken seriously. Second, note that for both vacancy and interstitial diffusion in the Chen [16] and in the Tang and Stevenson [15] work a larger activation energy correlates to a larger pre-exponential factor. This occurs

Table 1. Corrections to experimental activation energies and pre-exponential coefficients.

Experiment	E'(eV)	G_∞ (cm^2 s^{-1})	T_1 (K)	T_2 (K)	τ	E(eV)	C	m
Vacancy formation								
Vydyanath [13, 14]	2.24		400	650	495	2.01		11/2
Vacancy diffusion								
Chen [16]	2.40	$4.7 \times 10^5 P_{Hg}^{-1}$	400	490	441	2.08	$3.2 \times 10^{-21} P_{Hg}^{-1}$†	17/2
Tang and Stevenson [15]	2.10	$4.87 \times 10^4 P_{Hg}^{-1}$	350	500	412	1.80	$5.8 \times 10^{-22} P_{Hg}^{-1}$†	17/2
Interstitial diffusion								
Chen [16]	0.54	$1.1 \times 10^{-7} P_{Hg}$	400	490	441	0.64	$5.5 \times 10^{-13} P_{Hg}$‡	−5/2
Tang and Stevenson [15]	0.61	$5.5 \times 10^{-7} P_{Hg}$	350	500	412	0.70	$2.4 \times 10^{-13} P_{Hg}$‡	−5/2

† (cm^2 s^{-1}) − $T^{17/2}$
‡ cm − $T^{5/2}$ s^{-1}.

because of the T^m dependence and the fact that the two experiments were conducted over different temperature intervals T_1 and T_2.

The corrected numbers for the two groups are closer together than the uncorrected ones in both cases, though still not in perfect agreement. If we average the corrected Hg diffusion activation energies of the two experimental groups we get

$$E_{D,V_{Hg}} = 1.94 \pm 0.2\,eV$$
$$E_{D,Hg_i} = 0.67 \pm 0.04\,eV.$$

3.2. Native defects

Much of the following discussion has appeared previously in [2]. As mentioned earlier the doubly ionized cation vacancy is believed to be the dominant native defect in HgCdTe that dictates the electrical behaviour of the undoped material [13, 14]. We have recently reported the results of calculations of native and impurity defect total energies in HgTe, CdTe and ZnTe [18]. The energies for the formation of various native point defects in HgTe are summarized in table 2; the HgTe solid and the Hg in the vapour are used as reference energies†. These calculations used the linearized muffin-tin orbital (LMTO) method within the local density approximation (LDA) to the exchange correlation energy. Large supercells containing one defect per cell were repeated periodically, and from the difference in total energies per cell, with and without the defect, the defect formation energy was calculated. To expedite the calculations, the energies were calculated within the atomic spheres approximation (ASA) with a small (spd) basis set. In the ASA, an approximation to the exact density functional is evaluated; as a result, an error is introduced which is larger than in other LDA methods [38], and relaxation energies cannot be accurately calculated. Thus only those differences in energies that are > 0.5 eV here should be viewed as significant for these calculations.

An appropriate set of mass action constants for the neutral defect reactions is also given in table 2. The notation in that table is as follows. A_B corresponds to an A species occupying a B site, where I corresponds to an interstitial and V to a vacancy. No subscript on a species indicates that it is occupying the correct lattice site, e.g. $Hg = Hg_{Hg}$. Square brackets [ ] refer to concentrations. A subscript 'g' indicates the species in the gaseous, or vapour, phase, and P_{Hg} is the Hg vapour pressure. Most of the reactions in table 2 involve the creation or destruction of one or more unit cells. Because the resulting change in volume is accommodated at the surface, the change in the number of unit cells will enter into the determination of the defect equilibrium through the surface entropy. Additionally, surface preparation and orientation will affect the surface free energy. We have assumed for the present that such surface effects are negligible, i.e. that the volume expansions and contractions can occur with negligible changes in the surface properties.

To complete the defect equilibrium determination correctly we must include the equilibration of the electronic charges of the system. To do so we must have knowledge of the dominant charge states of the defects and their activation energies with respect to the neutral defect. Such calculations are complicated by the fact that most *ab initio* calculations of the electronic band structure of semiconductors predict an incorrect bandgap, E_G, a shortcoming of the local density approximation (LDA). Therefore, we shall focus on the neutral native defects here, and the established or expected charge states of these defects.

In wide-bandgap materials the equilibration of defects can be substantially affected by the Fermi level; for example the formation energy of a donor will decrease when the Fermi energy is near the valence band edge, since the donor electron can drop into a vacant state near the valence band, thereby lowering the energy by $\sim E_G$. Because we are discussing HgCdTe with a narrow bandgap, we expect the Fermi effects to be small, but not insignificant at high temperatures. Because of the small conduction band effective mass, in n-type material the filling of the conduction band states by electrons can shift the Fermi energy significantly. Combined with the increase in the bandgap for the high temperatures at which most defect studies are done, the effective bandgap can be substantially larger than the usual 77 K bandgap associated with a given concentration of HgCdTe.

† A different reference is used here than was used in table III in [18]. An error appears in that table owing to the incorrect use of an energy of $2E_b$ per unit cell rather than $4E_b$. The defect energies in table I in [18], from which the energies in table III were derived, are correct.

Table 2. Defect reactions and formation energies.

Defect reaction	Defection concentration	Energy (eV)
$E_{V_{Hg}} + HgTe \leftrightarrow V_{Hg}Te + Hg_g$	$[V_{Hg}] = P_{Hg}^{-1}K^0_{V_{Hg}} \exp(-E_{V_{Hg}}/kT)$	2.01†
$E_{Te_{Hg}} + 2HgTe \leftrightarrow Te_{Hg}Te + 2Hg_g$	$[Te_{Hg}] = P_{Hg}^{-2}K^0_{Te_{Hg}} \exp(-E_{Te_{Hg}}/kT)$	4.53
$E_{Te_i} + HgTe \leftrightarrow Te_iTe + Hg_g$	$[Te_i] = P_{Hg}^{-1}K^0_{Te_i} \exp(-E_{Te_i}/kT)$	4.96
$E_{V_{Te}} + Hg_g \leftrightarrow HgV_{Te}$	$[V_{Te}] = P_{Hg}K^0_{V_{Te}} \exp(-E_{V_{Te}}/kT)$	3.12‡
$E_{Hg_{Te}} + 2Hg_g \leftrightarrow HgHg_{Te}$	$[Hg_{Te}] = P^2_{Hg}K^0_{Hg_{Te}} \exp(-E_{Hg_{Te}}/kT)$	−0.42
$E_{Hg_i} + Hg_g \leftrightarrow Hg_i$	$[Hg_i] = P_{Hg}K^0_{Hg_i} \exp(-E_{Hg_i}/kT)$	0.84, 0.98

† Corrected experimental number from Vydyanath [13, 14]
‡ Calculated using a tight-binding Hamiltonian [17].

First we consider the defects which accommodate excess tellurium—the first three defects in table 2. The mass action constants are given by a product of the form

$$K_i = K_i^0 \exp(-\Delta E_i/kT). \quad (26)$$

For the first three equations the K_i^0 are given by

$$K^0_{V_{Hg}} = C_0(kT)^{5/2}(2\pi m_{Hg})^{3/2}h^{-3}\exp(\Delta S_{V_{Hg}}/k) \quad (27)$$

$$K^0_{Te_{Hg}} = C_0(kT)^{5}(2\pi m_{Hg})^{3}h^{-6}\exp(\Delta S_{Te_{Hg}}/k) \quad (28)$$

and

$$K^0_{Te_i} = C_0(kT)^{5/2}(2\pi m_{Hg})^{3/2}h^{-3}\exp(\Delta S_{Te_i}/k) \quad (29)$$

Here T is the temperature in kelvin, k is Boltzmann's constant, m_{Hg} is the mass of the mercury vapour atoms, h is Planck's constant, ΔS_i is the change in vibrational entropy upon formation of the defect and C_0 converts from site fraction to volume concentrations. Estimates, valid at high temperature, of the temperature variation of the ΔS_i terms were included in the pre-exponential dependence T^m of the reaction constants in table 1. Because two unit cells of HgTe are destroyed when a tellurium antisite is created, compared with one unit cell when a mercury vacancy is created, we do not expect that $\exp[(\Delta S_{Te_{Hg}} - \Delta S_{V_{Hg}})/kT] \approx 1$. While we have not completed the evaluation of these entropy terms, our preliminary estimates indicate that this ratio is $\sim 10^4$. For the tellurium interstitial and the mercury vacancy, we expect that $\exp[(\Delta S_{Te_i} - \Delta S_{V_{Hg}})/kT] \approx 1$ will be correct within a factor of 10. Evaluating the numerical constants we find

$$\frac{[Te_{Hg}]}{[V_{Hg}]} \simeq 10^{-10} \quad (30)$$

$$\frac{[Te_i]}{[V_{Hg}]} \simeq 10^{-18} \quad (31)$$

for $T = 500\,^{\circ}C$ and $P_{Hg} = 1$ atm. The conclusion from equations (30) and (31) that the mercury vacancy is the dominant native defect is consistent with experimental observation. This conclusion is unchanged if we include the possibility that the species may be ionized at the equilibration temperature where the material is expected to be intrinsic. Although the tellurium antisite density decreases more rapidly with decreasing Hg pressure than does the mercury vacancy density, the point at which the concentrations are comparable is at less than $P_{Hg} \cong 10^{-10}$ atm, and certainly the HgTe phase boundary is reached before such low Hg pressures can be achieved. This is also consistent with the fact that no p-to-n conversion is observed in isothermal anneals for low mercury pressures [13, 14, 38], as would be expected if tellurium antisites became the majority native defect. Because the pressure dependences of the tellurium interstitial and the mercury vacancy concentrations are the same, the above conclusions will hold independent of the mercury pressure.

We have checked the sensitivity of the calculated concentration ratios to the magnitude of the reaction enthalpy. Because the enthalpies enter the exponents, small changes in the enthalpies will result in large changes in the predicted defect concentrations. For example, let us assume that our calculated antisite formation enthalpy is in error by 0.5 eV: in this case the ratio of antisite-to-vacancy concentrations (at 500 °C) will be reduced to $\sim 1 \times 10^{-11}$. For an antisite formation enthalpy in error by 1.0 eV, this ratio is reduced to $\sim 1 \times 10^{-13}$. We do not expect the ASA errors to exceed 0.5 eV [5].

If the HgCdTe is not completely annealed, and tellurium precipitates are still present, the defect equilibrium will not be that predicted by the mass action equations given in table 2. For example, near the inclusions we can assume that the defects will be nearly in equilibrium with the tellurium solid; thus

$$E'_{V_{Hg}} + Te_s \leftrightarrow V_{Hg}Te \quad (32)$$

and

$$E'_{Te_{Hg}} + 2Te_s \leftrightarrow Te_{Hg}Te \quad (33)$$

will be the appropriate reactions. The formation energies for a tellurium antisite and an Hg vacancy from the tellurium solid are calculated to be 1.63 eV and 1.15 eV, respectively. Although the difference in the formation energies is less than when both defects are referenced to the mercury vapour (~ 0.5 eV compared with ~ 2 eV), the gas phase entropy factor does not enter into the ratio of the defect concentrations. Using the same estimate of the entropy ratio, the defect concentration ratio using tellurium solid as the reference state is

$$\frac{[Te_{Hg}]}{[V_{Hg}]} \simeq 10^{-8} \quad (34)$$

Thus, near the inclusions we expect higher relative concentration of tellurium antisites, as compared with the rest of the material equilibrated with the Hg vapour. Additionally, the absolute $[V_{Hg}]$ defect concentrations may differ substantially in the two regions of the material. At present we expect $[V_{Hg}]$ to be higher in an 'atmosphere' surrounding a Te inclusion. A better calculation of the vibrational entropy is needed before we can predict these absolute defect concentrations and their spatial extent. Differences in the defect concentrations arising from different equilibration conditions are a possible source of spatial variability of the HgCdTe material. If the material is not fully annealed to equilibrium, for example because of an abundance of tellurium precipitates, this history may affect subsequent processing.

In the above we have discussed the defect energies for HgTe and applied them directly to the small-x $Hg_{1-x}Cd_xTe$ system. Because we are dealing with the native defects of an alloy, we expect a number of complexities to affect the above analysis. First the variation of the defect formation energies for vacancies is sensitive to the alloy environment, in particular for the vacancies of the non-substituted species, such as tellurium in HgCdTe [39]. Even for vacancies of the substituted species, we have found that the formation energies may vary by several tenths of an electron volt. Because of this variation in the formation energy, the fraction of defective sites

will vary by as much as a factor of 100 from one class of sites to the next. Consider various classes of Hg sites in ideal HgCdTe, which can be distinguished to first order by specifying the number of Hg and Cd atoms in the second-neighbour shell (the four first neighbours are always tellurium), $Hg_{12-j}Cd_j$, with a concentration given by $[j]$. The total vacancy concentration is given by

$$[V_{Hg}] = \sum_{j=1}^{12} [j] P_{Hg}^{-1} K_j^0 \exp(-E_j/kT) \qquad (35)$$

where E_j is the vacancy formation energy for the jth cluster. The populations of vacancies in each class of cluster, j, can be expected to differ because of differences in the cluster populations and the formation energies. Additionally, the defect energy levels may differ in the various classes of sites, possibly leading to different ionization states for vacancies in different classes of sites. If the cations in the alloy are randomly arranged, such differences may be difficult to infer experimentally. If, however, the cations are correlated, exhibiting short-range order, more complex behaviour may be present. As discussed earlier, such short-range order has been found in HgCdTe. In these cases, the contribution to the vacancy densities from the dominant class of clusters will be increased. Because the studies finding short-range order focus on the tellurium-centred five-atom clusters of the form $Hg_{4-n}Cd_n$ rather than on cation-centred clusters of the form $Hg_{12-n}Cd_n$, higher-level five-atom cluster-cluster correlations must be known to predict the effects on the vacancy populations.

Next we examine the defects which accommodate excess Hg in the solid. The existence region for HgCdTe is always tellurium rich, and thus the native defects which accommodate excess tellurium are expected to dominate. For these equations in table 2, K_i^0 is given by

$$K_{V_{Te}}^0 = C_0^{-1}(kT)^{-5/2}(2\pi m_{Hg})^{-3/2}h^3 \exp(\Delta S_{V_{Te}}/k) \quad (36)$$

$$K_{Hg_{Te}}^0 = C_0^{-1}(kT)^{-5}(2\pi m_{Hg})^{-3}h^6 \exp(\Delta S_{Hg_{Te}}/k) \quad (37)$$

and

$$K_{Hg_i}^0 = C_0(kT)^{-5/2}(2\pi m_{Hg})^{-3/2}h^3 \exp(\Delta S_{Hg_i}/k). \quad (38)$$

If we assume the change in entropy is comparable for all three defects, we find

$$\frac{[Hg_{Te}]}{[V_{Te}]} \simeq 10^{+14} \qquad (39)$$

and

$$\frac{[Hg_i]}{[V_{Te}]} \simeq 10^{+14} \qquad (40)$$

for $T = 500\,^{\circ}C$ and $P_{Hg} = 1$ atm. From equations (39) and (40) we see that the tellurium vacancy is a minority defect species. For the pressure and temperature considered, the density of Hg antisites is predicted to be comparable to the density of Hg interstitials. Because the ratio of $[Hg_{Te}]$ to $[Hg_i]$ is nearly unity, any errors in the calculation of the activation energy could push the balance toward one side or the other. Thus we must depend on the next generation of calculations, with the ASA removed and full relaxation included, plus a quantitative comparison of the entropy differences between the mercury antisite and the tellurium interstitial, to determine the dominant defect in this class.

As mentioned above, we have shown that the tellurium vacancy formation energy varies significantly with the alloy environment. Because the tellurium vacancy is not expected to be a dominant defect in HgCdTe, and the tellurium diffuses by an interstitial mechanism, we do not expect any measurable manifestation of this variation. On the other hand, the Hg antisite may be the dominant Hg-excess defect, and its formation energy may vary significantly with the alloy environment. We are currently calculating the magnitude of this variation.

3.3. Diffusion

HgCdTe exhibits a complex tracer difusion profile, with both a fast and a slow branch. The fast branch is attributed to a mechanism with vacancy and interstitial diffusion in parallel where the dominant diffuser is determined by the pressure and temperature, while the slow component fits a mechanism with vacancy and interstitial diffusion in series [15]. The activation energies for the fast branch as discussed previously are 1.90 eV and 0.67 eV for the vacancy and interstitial mechanisms, respectively. Our calculated formation energy for the mercury interstitials are 0.84 and 0.98 eV for the anion- and cation-surrounded tetrahedral interstitial sites, respectively, and the experimental formation energy for the mercury vacancy is 2.01 eV. Comparing these energies with the experimental activation energies we find close agreement, indicating that the migration energy contribution to the diffusion activation energies are small for both interstitials and vacancies.

In the vacancy diffusion case we are comparing two experimental numbers for the diffusion and formation energies. Because the diffusion energy (1.94 eV) is the sum of the formation and migration energy, the fact that the vacancy formation energy (2.01 eV) is larger but should be smaller is an indication that something is amiss, and the experiments should be repeated with the goal of attaining higher accuracy.

There is a similar discrepancy in the interstitial diffusion case, but now we are comparing the experimental diffusion activation energy 0.67 eV with a theoretical formation energy 0.89 eV. We know that refinements to the theory will lower the predicted value. These refinements need to be done before more definitive conclusions can be made.

The difference between the two numbers obtained for cation- and anion-surrounded interstitital sites (0.84 eV and 0.98 eV) sets a lower bound on the interstitial migration energy. Most of the interstitial Hg will sit on the lower-energy anion-surrounded site, and migrate through the intermediate-energy cation-surrounded sites. Examination of the lattice arrangement between these sites leads us to believe that the potential profile is

unlikely to have a large barrier between the two classes of site, and, as a consequence, the interstitial migration barrier should be quite low.

In a recent experiment on mercury diffusion in ion-implantation-damaged HgCdTe, an activation energy of several tenths of an electron volt was measured [40]. The disparate results can be interpreted as a measure of only the defect migration contribution to the diffusion activation energy, since defects in excess of the equilbirium concentration were probably formed during implantation. It is not evident that the measured activation energy corresponds to the vacancy or the interstitial mechanism. The conclusion that the diffusion activation energies are largely defect formation energies, with the migration energies being much smaller, is in agreement with the above interpretation of the Richter and Kalish [40] experiment.

4. Conclusions

We have incorporated our calculated defect energies into the mass action equations for the neutral defects in HgCdTe. In agreement with experiment, we find the mercury vacancy to be the dominant native defect in tellurium-rich material. We also find the mercury antisite and interstitial defect densities to be comparable, although a better calculation of the vibrational entropy is needed to confirm this result. Comparing the defect formation energies of Vydyanath [13, 14] and our theory to the diffusion measurements by Chen [16] and by Tang and Stevenson [15], we find agreement with their diffusion activation energies for both the vacancy and the interstitial mechanisms, if we assume that the migration energy is small in both cases. Small migration energies are consistent with the measurement of Richter and Kalish [40] and our theory. Further work is in progress to incorporate the defect charge states into the calculation, and also to calculate the fully relaxed defect energies with the full potential LMTO.

If the cation migration energies are as small as we are suggesting, then some cation motion will be present even at room temperature. This implies that the crystal can equilibrate to its room temperature order-disorder correlation state. Most semiconductor alloys at room temperature are in metastable correlation states frozen to the temperature where diffusion stopped as they were cooled. How this affects devices remains to be determined.

We predict that near a Te inclusion there will be an 'atmosphere' of defects differing from the one the bulk solid has in equilibrium with a Hg vapour. In this atmosphere the ratio of the concentrations of the tellurium antisites to Hg vacancies will be increased. The tellurium antisite concentration may still be too small to affect any device properties, but if the Hg vacancy concentration is greater near the inclusions than in the bulk, this will be a source of spatial variability in the material. This is a potential source of the worms in LPE material. We are cautious about this suggestion because TEM studies of annealed LPE material indicate that there are few, if any, Te inclusions remaining [7].

First-principles theories are now fast enough and accurate enough for practical use in process and performance modelling. An effort to capture the results of many experiments into consistent models, and thereby increase their reliability, has, as we have tried to demonstrate in this paper, become a realistic goal.

Acknowledgment

This work was supported by NASA contract NAS1-18226, by ONR contract N00014-88-C0096 and by AFOSR contract F49620-88-K-0009.

References

[1] Balcerak R 1991 *Semicond Sci. Technol.* **6** C1
[2] Berding M A, Sher A and Chen A-B 1991 *Mater. Res. Soc. Symp. Proc.* **216** 3
[3] Sher A, Berding M A, Chen A-B and Patrick R 1988 *Advanced Research Workshop NATO International Science Exchange Programme, 5–9 Sept 1988*
[4] Sher A, van Schilfgaarde M and Berding M A 1991 *J. Vac. Sci. Technol.* B **9** 1738
[5] Chen M C, Parker S G and Weirauch D F 1985 *J. Appl. Phys.* **58** 3150
[6] Yoshikawa M, Maruyama K, Saito T, Maekawa T and Takigawa H 1987 *J. Vac. Sci. Technol.* A**5** 3052
[7] Schaake H F, Tregilgas J H, Lewis A J and Everett P M 1983 *J. Vac. Sci. Technol.* A**1** 1625; Schaake H F, Tregilgas J H, Beck J D, Kinth M A and Gnade B E 1985 *J. Vac. Sci. Technol.* A**3** 143
[8] Nguyen C, Triboulet R and Lemasson P 1987 *Seventh American Conf. on Crystal Growth in Conjunction with II-VI-87, 12-17 July, Monterey CA*
[9] Patrick R S, Chen A-B, Sher A and Berding MA 1989 *Phys. Rev.* B **39** 5980
[10] Spicer W E, Silberman J A, Lindau I, Chen A-B, Sher A and Wilson J A 1983 *J. Vac. Sci. Technol.* **1** 1735
[11] Cole S, Carey G P, Silberman, J A, Spicer W E and Wilson J A 1985 *J. Vac. Sci. Technol.* A **3** 206
[12] Cooper D E and Harrison W A 1990 *J. Vac. Sci. Technol.* A **8** 1112
[13] Vydyanath H R 1981 *J. Electrochem Soc.* **128**, 2609
[14] Vydyanath H R, Donovan J C and Nelson D A 1981 *J. Electrochem Soc.* **128** 2625
[15] Tang M-F S and Stevenson D A 1990 *J. Vac. Sci. Technol.* 7 544 and references therein
[16] Chen J S 1985 *PhD Thesis* University of Southern California
[17] Jones C E, James K, Merz, J, Barunstein R, Burd M, Eetemadi M, Hutton S and Drumheller J 1985 *J. Vac. Sci. Technol.* A**3** 131
[18] Berding M A, van Schilfgaarde M, Paxton A T and Sher A 1990 *J. Vac. Sci. Technol.* A **8** 1103
[19] Destefanis G L 1991 *Semicond. Sci. Technol.* **6** C88
[20] Sher A, Chen A-B and van Schilfgaarde M 1986 *J. Vac. Sci. Technol.* A **4** 1965
[21] Sher A, van Schilfgaarde M, Chen A-B and Chen W 1987 *Phys. Rev.* B **36** 4279
[22] Zamir D, Beshah K, Becla P, Wolff P A, Griffin R G, Zax D, Vega S and Yellin N 1988 *J. Vac. Sci. Technol.* **6** 2612

[23] Amirtharaj P M and Pollak F H 1984 *Appl. Phys. Lett.* **45** 789
[24] Compaan A, Bowman R C and Cooper D E 1990 *Appl. Phys. Lett.* **56** 1055
[25] Vodopyanov L K, Kozyrev S P, Aleshchenko Y A, Triboulet R and Y Marfaing 1990 *Appl. Phys. Lett.* **56** 1057
[26] Quintana J P 1990 private communication (Mater. Res. Soc. Boston)
[27] Shin S, Arias J, Zandian M, Pasko J and DeWames R private communication
[28] Johnson S M, Rhiger D R, Rosbeck J P, Peterson J M, Taylor S M and Boyd M E, private communication
[29] Chandra D private communication;
Chandra D, Tregilgas J H and Goodwin M W 1991 *J. Vac. Sci. Technol.* B **9**, 1852
[30] Muller M W 1984 *Phys. Rev.* B **30** 6196
[31] Parayanthal P and Pollak F H 1984 *Phys. Rev. Lett.* **52** 1822
[32] Sher A, Berding M A and Chen A-B 1989 *J. Cryst. Growth* **98** 27
[33] Wei S-H and Zunger A 1988 *J. Vac. Sci. Technol.* A **6** 2597
[34] Czyzyk M T, Podgorny M, Balzarotti A, Letardi P, Motta N, Kisiel A and Zemmal-Slarnavska M Z 1986 *Z. Phys.* B **62** 3850
[35] Ichimina M and Sasaki A 1986 *J. Appl. Phys.* **60** 3850
[36] Chen A-B and Sher A 1985 *Phys. Rev.* B **32** 3695
[37] Harrison W A 1980 *Electronic Structure and the Properties of Solids* (San Francisco: Freeman)
[38] Jones C L, Quelch M J T, Capper P and Gosney J J 1982 *J. Appl. Phys.* **53** 9080
[39] Berding M A, Sher A and Chen A-B 1990 *J. Appl. Phys.* **68** 5064
[40] Richter V and Kalish R 1990 *J. Appl. Phys.* **67** 6578

Formation energies, bond lengths, and bulk moduli of ordered semiconductor alloys from tight-binding calculations

Chin-Yu Yeh and A.-B. Chen
Department of Physics, Auburn University, Auburn, Alabama 36849

A. Sher
SRI International, Menlo Park, California 94025
(Received 20 September 1990)

A tight-binding model is used to calculate the formation energies, bond lengths, and bulk moduli of a number of ordered III-V and II-VI semiconductor alloys. The parameters in the model are adjusted so that the bond lengths, cohesive energies, bulk moduli, and shear elastic constants for the constituent compounds are described correctly. The model is then applied to alloys without further adjustment. Based on the calculated excess energies, we conclude that none of the ordered alloys found experimentally is in its stable bulk equilibrium state at the growth temperatures. Although the alloy excess energies can be negative, if the reference constituent compounds are constrained to match the substrate lattice used in epitaxial growth, their magnitudes are not large enough to account for the observed ordering. A possible explanation of the observed states in terms of a barrier that prevents the metastable ordered alloy from decomposing into separate phases is presented. However, this explanation only applies to alloys with lattice-mismatched constituents. Detailed results on the bond lengths and bulk moduli are also discussed.

I. INTRODUCTION

The bulk semiconductor alloys $A_{1-x}B_xC$ were long thought to be ideal pseudobinary compounds, in which the C atoms sit in a fcc sublattice while the substituting atoms A and B randomly occupy the sites of the other fcc sublattice. However, several recent findings in the last few years revealed quite a different picture. First, extended x-ray-absorption fine-structure (EXAFS) experiments[1,2] clearly showed a bimodal distribution for the nearest-neighbor bond lengths in these alloys, which implies that the equilibrium atomic positions are not the virtual-crystal sites of the zinc-blende crystal. Recent experiments[3,4] and theories[5,6] also indicated that the arrangement of the alloying atoms in these systems is not completely random. Most surprising of all, however, has been the discovery of long-range ordering (LRO) in these alloys grown epitaxially.[7–19] Essentially, all the III-V alloys grown by molecular-beam epitaxy (MBE) or metal organic chemical vapor deposition (MOCVD) under some special growth conditions are found to be ordered. While a great majority of these ordered alloys form the ABC_2 compounds in one or more of the following three crystal structures: the CuPt (CP), CuAu I (CA), and chalcopyrite (CH), a few alloys are ordered in the form of A_3BC_4 with the famatinite or luzonite structure. These crystal structures have been well described by Wei *et al.*[20] Table I is a partial list of the LRO alloys that have been grown, together with the growth conditions and ordered structure found. For later emphasis, we note that the substrate temperatures for the ordering to occur range from 400 °C to 800 °C, and the ordering directions are not necessarily the same as the growth direction.

Finding LRO is surprising, because it is at variance with the well-established conventional picture for the bulk semiconductor alloys having simple phase diagrams with miscibility gaps[21–24] driven by strain energy. The question is, are these ordered alloys thermodynamic equilibrium states? This question can be answered if accurate values for the alloy excess energies can be determined. To be specific, we shall only consider three important structures, CP, CA, and CH, for the ABC_2 alloys, and define an excess energy ΔE as

$$\Delta E = E(ABC_2) - [E(AC) + E(BC)] , \quad (1)$$

where $E(ABC_2)$ is the energy per molecule, or per four atoms, in the alloy ABC_2, and $E(AC)$ and $E(BC)$ are the energies per pair of atoms in the AC and BC zinc-blende compounds, respectively. If these ordered alloys are in their thermodynamically stable states at the growth temperature, ΔE has to be negative and must have a magnitude considerably larger than 200 meV on the present scale.[25] This would be in contradiction with the positive values of ΔE previously reported for the bulk semiconductor alloys.[21–24,26] However, the ordered semiconductor alloys found from the epitaxial growth may be in a constrained equilibrium state, where the constraint is imposed by the substrate strain. We need to know the energetics of the various states involved besides the bulk access energy before we can start understanding the ordering and stability of these alloys. The calculation of some of these energies along with an accuracy analysis

TABLE I. A list of the ordered III-V semiconductor alloys identified experimentally.

Alloys	Structure	Growth method	Substrate	Temperature (°C)	Reference
$AlGaAs_2$	CA	MOCVD and MBE	GaAs(100) and GaAs(110)	600–800	Ref. 7
$AlInP_2$	CP	MOCVD	GaAs(001)	650–700	Ref. 13
$AlInAs_2$	CP	MOCVD	InP(001)	600	Ref. 9
$GaInP_2$	CP	MOCVD	GaAs(001)	650	Ref. 14
	CP	MOCVD	GaAs(001)	640	Ref. 15
	CP	MOCVD	GaAs(001)	650–700	Ref. 13
	CP	MOCVD	GaAs(001)	600–630	Ref. 16
	CP	MOCVD	GaAs(001)	600–700	Ref. 17
$GaInAs_2$	famatinite	LPE	InP(110)	630	Ref. 8
	luzonite	MBE	InP(001)	400	Ref. 10
	CA	MBE	InP(110)	500	Ref. 11
	CP	VLE	InP(001)	650–660	Ref. 12
Ga_2AsSb	CA	MOCVD	InP(100)	550–680	Ref. 18
	CH	MOCVD	InP(100)	600	Ref. 18
	CP	MBE		540	Ref. 19

is the main purpose of this paper.

The excess energies for a number of ordered semiconductor alloys have been calculated from local-density-functional (LDF) theory using various band-structure methods.[6,20,27–32] Although the LDF error for the cohesive energy of a III-V compound is typically several tenths of one eV per pair of atoms, it is generally believed that the excess energy based on the same technique is accurate to several meV, because the errors in LDF cancel in Eq. (1). However, the ordered alloys that we are considering are open structures containing several atoms per unit cell. The atomic positions in these alloys are usually distorted away from the regular zinc-blende sites. Only the most sophisticated band-structure theories which are capable of treating shear distortion, such as the full-potential linear combination of muffin-tin orbitals (FP-LMTO),[33] full potential augmented plane waves (FP-APW),[34] and the fully converged plane-wave pseudopotential method,[35] can be expected to yield precise results within LDF. Even with present-day computers, it is still too expensive to use these methods to perform calculations over a wide range of semiconductor alloys. On the other hand, although the valence-force-field (VFF) model[36,37] is simple and is effective in treating the strain energy, it cannot account for the chemical energy.[26] These considerations have motivated us to use the empirical tight-binding (ETB) method. The ETB not only can treat both the strain and chemical energies but also allows for precise and systematic computations. To eliminate the propagation of errors from constituent compounds to alloys, the parameters in ETB are adjusted to produce the experimental values for the cohesive energy, bond length d, bulk modulus B, and the shear elastic constant $C_{11}-C_{12}$ for each constituent crystal. These parameters are then used in the alloy calculation without further adjustment. Our method thus corresponds to an interpolation scheme for the alloys between the constituent compounds. Such an approach is particularly appropriate for the present study, because the alloys and the constituent crystals have very similar structures and local bonding.

The rest of this paper is arranged as follows: Section II describes the ETB model, the way the parameters are determined, and the results for the structural properties of the constituent crystals. Section III briefly describes the structural parameters and the energy-minimization procedure for the three alloy structures in both their bulk equilibrium states and in states constrained to match substrates. The calculated excess energies are summarized in Sec. IV and are compared with those from LDF and VFF. To provide more detailed structural information, the calculated bond lengths and bulk moduli of the alloys are also presented. The final section, Sec. V, contains a summary and discussion.

II. TIGHT-BINDING MODEL

The tight-binding (TB) model that we are going to use is very similar to that used by Chadi[38] and Harrison.[39,40] The total energy of a semiconductor crystal is assumed to be the sum of the electron energies $\varepsilon_v(\mathbf{k})$ in the valence bands and the pair repulsive energies u_{ij} between the nearest-neighbor atoms:

$$E_T = E_{BS} + U_r = \sum_v \sum_k \varepsilon_v(\mathbf{k}) + \sum_{i>j} \sum u_{ij} \ . \tag{2}$$

Furthermore, the band energies are calculated using a minimum-basis TB Hamiltonian which includes one s and three p orbitals per atom. The interaction parameters needed from the Hamiltonian are the term values ε_s

and ε_p of each atom and the nearest-neighbor two-center interactions $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$, and $V_{pp\pi}$. For a given crystal structure, the Hamiltonian $H(\mathbf{k})$ associated with a given wave vector $\mathbf{k}$ within the first Brillouin zone has a dimension of $4m$, where m is the number of atoms per unit cell. The diagonal part of $H(\mathbf{k})$ consists of the term values, and the off-diagonal elements are computed as

$$H_{\gamma\gamma'}(\mathbf{k})=\sum_{d} e^{i\mathbf{k}\cdot\mathbf{d}_{\gamma\gamma'}} h_{\gamma\gamma'}(\mathbf{d}_{\gamma\gamma'}) , \tag{3}$$

where the sum runs over the first-neighbor bond displacements $\mathbf{d}_{\gamma\gamma'}$ that point from the orbitals denoted γ to γ'. The γ stands for the s, p_x, p_y, or p_z orbital of a particular atom in a given unit cell. The matrix elements $h_{\gamma\gamma'}$ are related to the two-center interactions by the Slater-Koster relations,[41]

$$h_{ss}=V_{ss\sigma} , \tag{4}$$

$$h_{sx}=\alpha_1 V_{sp\sigma} , \tag{5}$$

$$h_{xx}=\alpha_1^2 V_{pp\sigma}+(1-\alpha_1^2)V_{pp\pi} , \tag{6}$$

$$h_{xy}=\alpha_1\alpha_2(V_{pp\sigma}-V_{pp\pi}) , \tag{7}$$

where $\alpha_i=x_i/d$ are the direction cosines of $\mathbf{d}$, while all the V's depend only on the length d.

The first task is to determine the forms for the interactions $V_{\alpha\alpha'}$ and the repulsive energy u from the constituent compounds. Since the strain energy plays a very important role in the alloy formation energy, we shall make sure that our model produces the correct elastic constants. To keep the model close to Harrison's[39,40] original form, but to free it from his rigid $1/d^2$ and $1/d^4$ scaling rules for $V_{\alpha\alpha'}$ and u, respectively, we assume the following forms:

$$V_{\alpha\alpha'}(d)=V^{(0)}_{\alpha\alpha'}(d_0/d)^n \tag{8}$$

and

$$u(d)=u_0(d_0/d)^m , \tag{9}$$

where the superscript and subscript 0 indicate the values evaluated at the equilibrium bond length d_0. The values of $V^{(0)}_{\alpha\alpha'}$ are taken to be Harrison's universal forms scaled by a factor f,

$$V^{(0)}_{\alpha\alpha'}=fV^{\text{Harrison}}_{\alpha\alpha'} . \tag{10}$$

Note that Harrison's[40] universal two-center interactions take the form

$$V^{\text{Harrison}}_{\alpha\alpha'}=\eta_{\alpha\alpha'}\hbar^2/(md_0^2) ,$$

where m is the free-electron mass, $\hbar$ is Planck's constant, and the η's take the following values: $\eta_{ss\sigma}=-1.32$, $\eta_{sp\sigma}=1.42$, $\eta_{pp\sigma}=2.22$, and $\eta_{pp\pi}=-0.63$.

Thus there are four adjustable parameters for each system: the scaling parameter f, the powers n and m, and the value u_0. These parameters are determined by requiring that the model produce the correct experimental values for the bond energy E_{bond}, d_0, B, and the shear elastic constant $C_{11}-C_{12}$. Since $C_{11}-C_{12}$ is governed only by $V^{(0)}_{\alpha\alpha'}$ in the present model, it alone determines the scaling factor f. Then the bond energy E_{bond} is used to determine u_0. The requirement that the first derivative of E_T be zero at d_0 then determines the ratio of the powers n/m, which couples with the equation for the bulk modulus to yield the values for n and m. One can then use these sets of parameters to calculate other quantities that are not employed in the fitting, e.g., another shear elastic constant C_{44}, the Kleinman internal-displacement parameter[42] ζ, and the optical-phonon frequencies ω at the zone center, to check the validity of the model.

In the actual calculations we used the term values tabulated by Chen and Sher.[43] Table II lists the experimen-

TABLE II. Values of bond length d, bond energy E_{bond}, bulk modulus B, and shear coefficient $C=C_{11}-C_{12}$ used to determine the parameters in Table III. Also listed are the experimental values of C_{44} and the TO-phonon mode ω at Γ to be compared with the calculations. All the elastic constants are in units of 10^{11} dyn/cm^2, d in Å, E_{bond} in eV, and ω_{TO} is given in terms of wave numbers in cm^{-1}. The sources of these values are discussed in Ref. 47. Also listed are the force constants α and β (N/m) for the valence-force-field model in Eq. (13).

	d	E_{bond}	B	C	C_{44}	ω	α	β
AlP	2.367	−2.13	8.600	6.900	6.150	440	43.867	9.429
AlAs	2.451	−1.89	7.727	7.160	5.420	361	40.360	10.132
AlSb	2.656	−1.76	5.817	4.428	4.076	366	33.417	6.790
GaP	2.360	−1.78	9.143	7.870	7.143	367	46.257	10.723
GaAs	2.448	−1.63	7.690	6.630	6.040	269	40.351	9.371
GaSb	2.640	−1.48	5.792	4.946	4.440	231	32.800	7.539
InP	2.541	−1.74	7.247	4.460	4.600	304	40.346	6.543
InAs	2.622	−1.55	5.794	3.803	3.959	219	33.165	5.757
InSb	2.805	−1.40	4.831	3.130	3.132	185	29.605	5.069
ZnTe	2.637	−1.20	5.090	3.060	3.120	177	29.445	4.659
CdTe	2.806	−1.10	4.210	1.680	2.040	141	26.374	2.722
HgTe	2.798	−0.81	4.759	1.817	2.259	116	29.773	2.935

tal values[44-47] for the d_0, E_{bond}, B, and the $C_{11}-C_{12}$ used to fit the parameters, and the values of C_{44} and transverse phonon frequency ω used for the consistency check.

The elastic constants can be calculated directly from perturbation theory. First the Hamiltonian $H(\mathbf{k})$ is expanded in powers of the infinitesimal strain parameters e up to second order:

$$H = H_0 + H_1 e + \tfrac{1}{2} H_2 e^2 , \tag{11}$$

where H_0 is the strain-free Hamiltonian and H_1 and H_2 are, respectively, the first and second derivatives with respect to e evaluated at $e=0$. The band-energy contribution to the strain coefficient then comes from the second derivative of E_{BS} with respect to e, denoted by

$$\frac{\partial^2 E_{\text{BS}}}{\partial e^2} = \sum_v \sum_{\mathbf{k}} \langle v,\mathbf{k}|H_2|v,\mathbf{k}\rangle + 2\sum_v \sum_c \sum_{\mathbf{k}} \frac{|\langle v,\mathbf{k}|H_1|c,\mathbf{k}\rangle|^2}{\varepsilon_v(\mathbf{k})-\varepsilon_c(\mathbf{k})} , \tag{12}$$

where $\varepsilon_c(\mathbf{k})$ and $|c,\mathbf{k}\rangle$ are, respectively, the eigenenergies and eigenvectors of H_0 for the conduction bands. Similarly, $v,\mathbf{k}$ indexes the valence bands. Note that the inter-valence-band contributions in the second-order perturbation sum cancel exactly and so they are not needed in Eq. (12). The matrix elements of H_1 and H_2 needed here can be expressed from Eqs. (3)–(7) in terms of the first and the second strain derivatives of the two-center interactions $V_{\alpha\alpha'}$ and the direction cosines α_i.

Table III shows the results for f, n, m, and u_0 obtained for the constituent compounds, and the corresponding values of C_{44}, ζ, and ω calculated as a consistency check. The scaling factor f ranges from 1.1 to 1.5 and tends to decrease with an increase in polarity. In the power-law dependence $V_{\alpha\alpha'} \propto (d_0/d)^n$, n ranges from 3.3 to 4.2, which is consistently larger than the $n=2$ used in Harrison's universal TB parameters. For the repulsive pair energy $u = u_0(d_0/d)^m$, the power m ranges from 5.4 to 7.1. The ratio m/n falls in the range from 1.5 to 1.9, which is smaller than the $m/n=2$ used by Harrison.[39] The calculated values of C_{44} for most systems agree with experiment to 10% or better. The calculated TO-phonon modes at Γ for most III-V systems also agree with experiments to 10% or better. The discrepancies for the II-VI systems are larger (about 15%). Reliable results for ζ from both experiments and first-principles calculations are only available for a limited number of semiconductors. The calculated ζ, C_{44}, and ω from the TB model agrees very well with those results, as shown in Table IV.

The results in Tables III and IV are based on a particular set of term values and TB parameters. It is useful to examine how the predictions are influenced by these parameters and the fitting procedure. Table V shows a result based on Chadi's procedure[38,48] in which the TB matrix elements $V_{\alpha\alpha'}$ are scaled as $1/d^2$, and the repulsive pair energy is taken to be

$$u = u_0 + u_1(d-d_0) + u_2(d-d_0)^2 .$$

The parameter u_0 is set to produce the correct bond energy, u_1 is determined by requiring the correct equilibrium bond length, and u_2 is fixed by the bulk modulus. Two sets of TB parameters are tabulated for each system: One of them is the set used by Chadi[38] and the other (labeled present work) is the set obtained by multiplying Harrison's $V_{\alpha\alpha'}$ by the scaling factor f listed in Table III. For convenient comparison, the zero of the term values is set equal to the anion s energy. Despite considerable differences in these two sets of TB parameters, the results of the predictions from the two sets are very similar and also very similar to those predicted from the other procedure given in Table IV. The only noticeable difference

TABLE III. The results for the parameters f, n, m, and u_0 obtained from the fitting of the bond energy, bond length, bulk modulus, and shear coefficient $C_{11}-C_{12}$ of Table II using the full band-structure calculation. Also listed are the calculated C_{44}, internal-displacement parameter ζ, and the TO-phonon mode ω at Γ. All the elastic constants are in units of 10^{11} dyn/cm^2, u_0 is in eV, and ω are given in terms of wave numbers in cm^{-1}.

	f	n	m	u_0	C_{44}	ζ	ω
AlP	1.294	3.530	5.598	6.435	5.827	0.516	447
AlAs	1.464	3.524	5.430	7.089	5.598	0.459	384
AlSb	1.337	3.268	5.668	4.838	3.944	0.564	354
GaP	1.395	3.705	5.683	7.285	6.857	0.501	382
GaAs	1.397	3.633	5.716	6.530	5.791	0.500	292
GaSb	1.431	3.471	5.717	5.519	4.515	0.536	256
InP	1.323	4.240	6.633	5.603	4.260	0.584	304
InAs	1.300	3.997	6.427	4.962	3.564	0.552	220
InSb	1.353	3.773	6.399	4.350	3.092	0.602	200
ZnTe	1.284	3.306	5.828	4.285	2.813	0.590	205
CdTe	1.171	3.656	6.761	3.092	1.701	0.694	156
HgTe	1.173	3.760	7.074	3.080	2.040	0.716	152

between the predictions in Table IV and Table V is that in the latter the phonon frequencies are slightly larger and C_{44} are slightly smaller. These comparisons show that the model with parameters given in Table II not only contains the correct structure properties inserted through the fitting procedure, but also predicts other mechanical properties with reasonable accuracy. This model should serve as a good basis for alloy calculations.

III. ALLOY CALCULATION

A. CuAu I structure (CA)

The ABC_2 semiconductor alloy in the CA structure forms a layer structure $ACBCACBC\cdots$ along the $\langle 001\rangle$ direction. The basic lattice vectors can be chosen as $\mathbf{a}_1=(\frac{1}{2},\frac{1}{2},0)a$, $\mathbf{a}_2=(\frac{1}{2},-\frac{1}{2},0)a$, and $\mathbf{a}_3=(0,0,\beta)a$, where β is the c/a ratio with a and c being the lattice parameters. Note that the ideal β value for a zinc-blende structure is 1. There are four atoms per unit cell: one A atom at (0,0,0), one B atom at $(\frac{1}{2},0,\beta/2)a$, and two C atoms at $(1,1,\beta+\delta)a/4$ and $(3,1,3\beta-\delta)a/4$, where δ is the internal distortion parameter for the C atoms. There are only two different nearest-neighbor bond lengths in the alloy:

$$d_{AC}=(a/4)[2+(\beta+\delta)^2]^{1/2}$$

and

$$d_{BC}=(a/4)[2+(\beta-\delta)^2]^{1/2}.$$

For a set of values for a, β, and δ, the total energy is calculated following the general description in Sec. II. The interactions $V_{\alpha\alpha'}$ and the repulsive energy are scaled by the bond lengths according to Eqs. (8) and (9), respectively. The Hamiltonian $H(\mathbf{k})$ is now a 16×16 matrix. The total energy is then minimized by varying the three parameters α, β, and δ. If the alloy is constrained to match a (001) substrate, the lattice constant a is fixed by the substrate, and the total energy is minimized with respect to β and δ.

TABLE IV. Comparison between calculated and experimental lattice constant a, elastic constants B, $C_{11}-C_{12}$ and C_{44}, Kleinman (Ref. 42) internal-displacement parameter ζ, and the TO-phonon frequency ω in wave numbers cm^{-1}. Also listed are $C_{44}^{(0)}$ which correspond to the value without the internal distortion. The FP-LMTO and PP-PW are the *ab initio* theories, and TB is our tight-binding method discussed in the text. All elastic constants are in units of 10^{11} dyn/cm^2. Experimental values are those listed in Table II.

		Expt.	FP-LMTO[a]	PP-PW[b]	TB
Si	a	5.431	5.41	5.45	5.431
	B	9.923	9.9	9.3	9.923
	$C_{11}-C_{12}$	10.274	10.2	9.8	10.274
	C_{44}	8.036	8.3	8.5	8.013
	$C_{44}^{(0)}$			11.1	11.30
	ζ	0.54[c]	0.51	0.53	0.51
	ω	523	518	521	572
Ge	a	5.65		5.59	5.65
	B	7.653		7.2	7.653
	$C_{11}-C_{12}$	8.189		8.5	8.189
	C_{44}	6.816		6.3	6.84
	$C_{44}^{(0)}$			7.7	9.46
	ζ			0.44	0.49
	ω	303		302	342
GaAs	a	5.642		5.55	5.642
	B	7.69		7.3	7.69
	$C_{11}-C_{12}$	6.63		7.0	6.63
	C_{44}	6.04		6.2	5.791
	$C_{44}^{(0)}$			7.5	7.83
	ζ			0.48	0.50
	ω	273		268	292

[a]Reference 33.
[b]Reference 49.
[c]Reference 50.

B. Chalcopyrite structure (CH)

The ABC_2 semiconductor alloy in the CH structure forms a superlayer structure $ACACBCBCACACBCBC\cdots$ along the $\langle 012\rangle$ direction. The basic lattice vectors can be chosen as $\mathbf{a}_1=(1,1,-2\beta)a/2$, $\mathbf{a}_2=(-1,1,2\beta)a/2$, and $\mathbf{a}_3=(1,-1,2\beta)a/2$, where β again is the c/a ratio, with an ideal value of 1. There are now eight atoms per unit cell: two A atoms at (0,0,0) and $(0,1,\beta)a/2$, two B atoms at $(1,0,\beta)a/2$ and $(1,1,0)a/2$, and four C atoms at $(1+\delta,1,\beta)a/4$, $(1,3+\delta,3\beta)a/4$, $(3,1-\delta,3\beta)a/4$, and $(3-\delta,3,\beta)a/4$, where δ is the internal distortion parameter for the C atoms. Again there are also only two different nearest-neighbor bond lengths in the alloy:

$$d_{AC}=(a/4)[1+(1+\delta)^2+\beta^2]^{1/2}$$

and

$$d_{BC}=(a/4)[1+(1-\delta)^2+\beta^2]^{1/2}.$$

A similar procedure can now be carried out to minimize the total energy with respect to the three parameters a, β, and δ. However, the Hamiltonian $H(\mathbf{k})$ is now 32×32. For the case in which the lattice is matched to the (001) substrate, we again are left with two parameters β and δ to vary for the energy minimization.

C. CuPt structure (CP)

In the CP structure, the alloy forms a $\langle 111\rangle$ superlattice $ACBCACBC\cdots$. Because of a lack of reflection symmetry about any of these planes, the B layer need not be located exactly at the middle position between the two successive A layers. Also the distance between two closest atoms from two different A layers may not need to correlate with that between two A atoms on the same plane. Thus there are a total of five independent parameters required to describe the crystal structure: the lateral lattice constant a for the layers, the spacing D between two successive A layers, and the three spacing parameters for the three layers (one B and two C) inside D.

TABLE V. Comparison between the two different sets of TB parameters described in the text; the resultant coefficients u_0, u_1, and u_2 of u, and the predicted elastic constants, Kleinman internal-displacement parameters ζ, and phonon frequency ω.

		ε_s^A	ε_p^A	ε_s^C	ε_p^C	$V_{ss\sigma}$	$V_{sp\sigma}$	$V_{pp\sigma}$	$V_{pp\pi}$
Si	Chadi	0.0	7.20	0.0	7.20	−2.03	2.55	4.55	−1.09
	Present work	0.0	6.88	0.0	6.88	−2.41	2.59	4.05	−1.15
		u_0	u_1	u_2	C	C_{44}	C_{44}^0	ζ	ω
Si	Chadi	7.29	−9.98	23.90	10.66	7.89	11.38	0.49	620
	Present work	6.93	−9.70	23.42	10.27	7.83	11.39	0.51	592
		ε_s^A	ε_p^A	ε_s^C	ε_p^C	$V_{ss\sigma}$	$V_{sp\sigma}^{AC}$	$V_{sp\sigma}^{CA}$	$V_{pp\sigma}$	$V_{pp\pi}$
GaAs	Chadi	0.0	9.64	5.12	11.56	−1.70	2.40	1.90	3.44	−0.89
	Present work	0.0	10.09	6.79	14.12	−2.34	2.52	2.52	3.94	−1.12
		u_0	u_1	u_2	C	C_{44}	C_{44}^0	ζ	ω	
GaAs	Chadi	5.12	−7.12	18.22	6.36	5.60	8.77	0.54	339	
	Present work	6.53	−8.39	19.90	6.63	5.70	8.53	0.54	322	

There are, however, only four atoms per unit cell, each coming from a different layer, so the $H(\mathbf{k})$ matrix is 16×16. There are also four different bond lengths in the crystal, two for the AC and two for the BC bonds. If the alloy lattice is matched to a (111) substrate, the lateral lattice parameter is fixed, but we still have four parameters to vary for energy minimization. However, if the alloy is matched to a (001) substrate, we shall assume that all the alloying atoms A and B are locked into the fcc lattice points of the substrate, and we are left with only two free parameters which describe the relaxation of the two nonequivalent C layers.

It is useful to comment on the Brillouin-zone (BZ) integration needed for the calculation of the electronic energy. Since our calculation involves relatively small and easily handled matrices, we are able to sample over a large number of $\mathbf{k}$ points. We found that a uniform grid of 1000 $\mathbf{k}$ points inside the BZ always guarantees a convergence of the total energy pair of atom to within 10^{-3} meV and the elastic constants to an accuracy in the third digit. We found that total energies calculated using two special $\mathbf{k}$ points[51] are about 10 meV per atom pair higher than the converged value. However, when extending the two-special-$\mathbf{k}$-point method to the CA and CH structures,[52] most of the errors of the alloy and pure compound cancel, and we found that the final errors in the excess energies are only about 0.5 meV per atom pair.

Before presenting the results, we briefly describe the Keating[36] valence-force-field (VFF) model. The VFF only deals with the strain energy. The energy per unit volume in a strained diamond or zinc-blende crystal is given by

$$U=\frac{3}{8d_0^2}\sum_i \alpha_i[\Delta(\mathbf{d}_i\cdot\mathbf{d}_i)]^2+\frac{3}{8d_0^2}\sum_{i>j}\sum \beta_{ij}[\Delta(\mathbf{d}_i\cdot\mathbf{d}_j)]^2 , \tag{13}$$

where in the first term, the bond-stretching energy, the summation i runs over all the bonds, and in the second term, the bond-angle contribution, the summations include all the pairs of bonds that share common atoms. The d_0 in Eq. (13) is the equilibrium bond length, and $\Delta(\mathbf{d}_i\cdot\mathbf{d}_j)$ is the strain-induced change of the dot product between the two bond vectors which point from the common atom to the nearest-neighbor atoms. For a zinc-blende crystal, there is only one value for the bond-stretching force constants $\alpha_i=\alpha$, and one value for the bond-angle-restoring-force constant $\beta_{ij}=\beta$. Their values are determined[26] by fitting the experimental bulk moduli B and shear elastic constants $C_{11}-C_{12}$, and are tabulated in Table II. We extend Eq. (13) to calculate the strain energy in an ABC_2 alloy by treating each bond and each pair of the same bonds in the same way as in the constituent compounds. However, when dealing with the bond-angle term involving two unlike AC and BC bonds, both d_0 and β_{ij} are taken to be the average values. The results from VFF will be compared with the ETB calculations in Sec. IV.

IV. DISCUSSION OF ALLOY RESULTS

A. Excess energies

Before presenting the results for different kinds of excess energies, the readers should be reminded of the relation between these energies and ordering. The Appendix provides a qualitative discussion of this relation. It also describes how different kinds of excess energies presented below may alter the picture of ordering.

The bulk excess energies ΔE calculated from ETB for the ABC_2 alloys in the three structures of CA, CH, and CP are listed in Table VI along with the results from VFF and LDF. First we observe that ETB and VFF give very similar results. Both models produce very small ΔE values for those alloys composed of compounds with nearly equal lattice constants. However, ΔE for $AlGaP_2$

and $HgCdTe_2$ derived from ETB are slightly negative. The differences between the ETB and VFF results in these systems are a measure of the small size of the chemical-energy contribution to ΔE in the ETB model. For the lattice-mismatched alloys, all the ΔE values from ETB are positive and are slightly larger than those obtained from VFF except for $InAlP_2$. The major reason VFF yields a smaller ΔE is that the VFF used here consistently predicts smaller values of the elastic constant C_{44} than the experimental values, thereby underestimating the strain contribution to ΔE. However, the small chemical energy included in ETB may upset this trend, exemplified by $InAlP_2$.

The quantitative comparison between ETB and LDF is mixed. Starting from the lattice-matched alloys $GaAlAs_2$, the ΔE from LDF calculated by different groups range from 7.5 to 35 meV for the three structures considered, as compared to nearly zero calculated from ETB. Similar differences between LDF and ETB also occur in another lattice-matched system, $HgCdTe_2$. Since the strain energy is nearly zero in these systems, the 10-meV or so difference between the present calculation and the majority of the LDA results represents the discrepancy in the estimate of the chemical energy in ΔE between the two theories. These differences certainly are well within the margins of errors of both ETB and LDF. However, we note the LDA ΔE values for $GaAlAs_2$ in the CA structure are rather consistent except the 35 meV from Ref. 56. For the lattice-mismatched alloys, ETB agrees very well (within 20 meV) with LDF for the III-V alloys in the CP structure and II-VI alloys in the CH structure. The agreement is also reasonable for all alloys in the CA structure. However, the differences between the two calculations are more substantial for the III-V alloys in the CH structure and II-VI alloys in the CP structure. We note that although the trend $\Delta E_{CH} < \Delta E_{CA} < \Delta E_{CP}$ among the three structures holds for both ETB and LDF for these lattice-mismatched alloys, the ΔE_{CH} values from LDF for the III-V systems are considerably lower than those from ETB, particularly the negative values calculated for $InAlP_2$ and $InAlAs_2$. Despite all these quantitative differences, the qualitative trends of our ETB predictions are rather similar to those from LDF by Wei and co-workers[6,27] as shown in Fig. 1. The most important conclusion that can be drawn from Table VI is that these ordered bulk alloys are not the thermal equilibrium states at the experimental growth temperatures shown in Table I. For this to happen, the ΔE value has to be less than -200 meV per four atoms,[25] as mentioned earlier. This ΔE value is far below all the calculated values and is beyond the uncertainties of our ETB model and the LDF calculations listed above.

While the above excess energies do not support a thermally stable ordered bulk alloy at the growth temperature (≈ 600 °C), they may offer different conclusions when applied to the epitaxially grown alloys. The idea was first suggested by Flynn[57] in connection with epitaxial growth of ordered metal alloys. If the substrate material serves as a reservoir for the epitaxial film, then the alloys grown on the substrate have to be in thermal and mechanical equilibrium with the substrate. Since most substrates used in the growth are selected to have their lattice constants close to the equilibrium lattice parameters (see Tables I and IX) of the ordered alloys mentioned above, it takes little strain energy for these alloys to match the substrates. On the other hand, if an alloy on the substrate is to decompose into two substrate lattice-mismatched constituent compounds, it takes energy to constrain the separated systems. Only the case of thin layers in which no dislocations form will be treated. Then the reference constituent energies $E(AB)$ and $E(AC)$ in Eq. (1) which must be used to calculate the excess energy for epitaxial growth, referred to as ΔE_{ep}, are the ones with their lattice parameters matched to the substrate. This coherent strain can make ΔE_{ep} negative for an alloy even when its bulk excess energy ΔE is positive. Of course, the actual values of excess energies are substrate specific. However, an overall assessment can be made when the substrate is also an alloy by choosing the substrate lattice constant to be the average value

$$a_s = \bar{a} = (a^0_{AC} + a^0_{BC})/2 \ .$$

The major difference between ΔE_{ep} and ΔE then comes from the strain energies of the constituent compounds forced to lattice match the substrate. Since the energies of the ordered alloys at their equilibrium lattice constants a only differ by a small amount (1 or 2 meV) from those at $\bar{a}$, we shall use the alloy energies already calculated in Table VI to deduce the epitaxial excess energy.

The strain energy of either constituent compound matched to $a_s = \bar{a}$ on the (001) substrate can be estimated from the following simple formula:

$$E_s(AC) = (C_{11} + 2C_{12} - 4C^2_{12}/C_{11})\delta^2_0 a_s^3/4 \ , \qquad (14)$$

where $\delta_0 = (a^0_{AC} - a^0_{BC})/\bar{a}$ is the percentage lattice-parameter difference between the two constituent compounds. The strain energies used to calculate the values of ΔE_{ep} in Table VII, however, are obtained from ETB, which in fact only produces a small correction (1 to 2

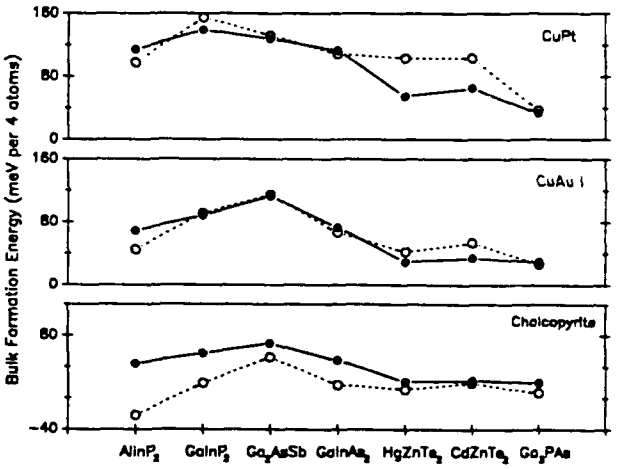


FIG. 1. The excess energies of Eq. (1) for the ordered alloys in three crystal structures calculated from the present TB model (solid circles) and from LDF in Ref. 6 (open circles).

meV at most) to Eq. (14). The relative values of ΔE_{ep} among the three structures on the same substrate, say, (001), should remain the same as those of ΔE, because these two energies only differ in the reference energy. For example, the relation

$$\Delta E_{ep}(\mathrm{CH}) < \Delta E_{ep}(\mathrm{CA}) < \Delta E_{ep}(\mathrm{CP})$$

still holds for alloys composed of lattice-mismatched constituent compounds. On the (001) substrate, the ΔE_{ep} values are essentially all positive in the CP, negative in the CH, and switch between positive and negative values in the CA structure. The CP alloys matched to the (111) substrate have very small ΔE_{ep} with magnitudes smaller than 10 meV. The LDF epitaxial energies calculated by Wei, Ferreira, and Zunger[6] have the same qualitative trends as those from ETB among the three structures. The quantitative differences in ΔE_{ep} between the two calculations are larger and also more scattered than those in ΔE. This is peculiar, because the difference between ΔE and ΔE_{ep} in both calculations comes from the strain energies of the constituent compounds, which can be reliably estimated from Eq. (14) and are well prescribed by our ETB model. Despite these discrepancies, a simple conclusion can also be drawn from Table VII. These numbers still cannot account for the ordering observed experimentally, because all the calculated ΔE_{ep} are still well above the -200-meV value required for the ordering at the growth temperatures.[25]

In the above consideration, the constituent compounds are constrained to match the lateral lattice of the substrate and are allowed to relax fully in the growth direction. For a (001) substrate with a lattice constant $a_s = \bar{a}$, the c/a ratio for either compound is estimated to be

$$c/a = 1 + \delta_0 (C_{12}/C_{11} + 0.5) , \tag{15}$$

where δ_0 again is the percentage bond-length difference between the two constituent compounds. This c/a relaxation results in elongation for one compound and shrinkage for the other along the growth direction. If the lattice constants of two constituent crystals have a substantial difference, these opposite relaxations will produce strained grain boundaries between the AC and BC crystals. The reference state energy $E(AC)+E(BC)$ used in

TABLE VI. The bulk excess energies ΔE (in meV per four atoms) from ETB and comparison with the results from local-density-functional (LDF) theory and the valence-force-field (VFF) model.

	CuAu I			Chalcopyrite			CuPt		
Alloys	ETB	LDF	VFF	ETB	LDF	VFF	ETB	LDF	VFF
$AlGaP_2$	−2.4		1.5	−2.8		1.0	−2.6		0.2
$AlGaAs_2$	0.6	10.8[a] 11.5[c] 15.1[b] 13.5[d] 35[e]	0.3	0.6	11.4[b] 9.8[c]	0.2	0.8	7.5[c]	0.4
$AlInP_2$	69.0	43.0[j]	74.0	44.0	−21.0[j]	47.3	114.2	97.0[f]	111.1
$AlInAs_2$	68.8	35.0[j]	66.6	45.8	−15.0[j]	43.7	107.0		97.6
Ga_2PAs	30.0	26.6[g]	23.1	19.4	6.5[g]	15.0	33.6	37.2[g]	32.7
Ga_2PSb	260.2		207.0	67.4		135.8	290.8		292.6
Ga_2AsSb	113.0	129.2[a] 114.8[h] 115.0[c]	91.8	69.8	52.0[c]	59.9	128.0	132.0[c]	130.1
$GaInP_2$	88.4	115.6[a] 91.0[i] 54.4[k]	83.8	57.2	19.0[i]	54.3	139.4	155.4[g]	124.9
$GaInAs_2$	73.2	60.1[j] 83.6[a] 66.7[f]	67.3	48.2	16.5[g]	43.8	113.0	108.5[g]	99.5
$GaInSb_2$	57.4		51.5	37.4		33.2	85.58		76.2
$ZnCdTe_2$	34.3	54.2[g]	33.7	22.4	19.2[g]	21.1	65.3	103.5[g]	56.0
$ZnHgTe_2$	29.7	42.5[g]	32.1	21.0	11.4[g]	19.9	54.9	103.3[g]	53.5
$CdHgTe_2$	−2.3	12.1[g]	0.61	−2.7	11.3[g]	0.4	−2.7	9.8[g]	1.1

[a]Reference 28.
[b]Reference 31.
[c]Reference 27.
[d]Reference 55.
[e]Reference 56.
[f]Reference 53.
[g]Reference 6.
[h]Reference 30.
[i]Reference 32.
[j]Reference 29.
[k]Reference 54.
[l]Reference 74.

TABLE VII. The excess energy ΔE_{cp} (in meV per four atoms) for epitaxial alloys calculated from the present TB model and comparison with the LDF results.

Structure	CuAu I		Chalcopyrite			CuPt	
substrate	(100)		(100)		(111)	(100)	(100)
(Alloy)	TB	LDF	TB	LDF	TB	TB	LDF
$AlGaP_2$	−2.9		−3.1		0.5	−3.0	
$AlGaAs_2$	0.8		0.7		0.4	0.2	
$AlInP_2$	−11.5	−36.0[a]	−36.3		−0.2	33.8	−18.0[a]
$AlInAs_2$	−4.6		−27.8		−7.4	33.6	
Ga_2PAs	5.3	4.1[a]	−5.3	−21[a]	0.2	8.9	0.5[a]
Ga_2PSb	40.5		−52.2		7.8	71.1	
Ga_2AsSb	15.5	35.0[b]	−27.6	−28[b]	−5.6	30.5	52.0[b]
		17.0[a]		−45.8[a]			11.5[a]
		48.0[c]					
$GaInP_2$	−3.3	−3.0[a]	−34.5	−106[a]	−9.4	47.6	3.4[a]
		12.8[c]					
$GaInAs_2$	0.4	−9.7[a]	−25.4		−9.6	39.5	4.1[a]
		10.5[d]					
		6.7[d]					
		8.4[d]					
		29.6[c]					
$GaInSb_2$	1.4		−18.5		−5.0	29.9	
$ZnCdTe_2$	−5.6	−5.4[a]	−17.5	−57.2[a]	6.6	25.5	0.2[a]
$ZnHgTe_2$	−5.9	−13.5[a]	−14.7	−64.2[a]	3.0	19.3	0.9[a]
$CdHgTe_2$	−2.4		−2.8		−3.0	−2.9	

[a]Reference 6.
[b]Reference 27.
[c]Reference 28.
[d]Reference 29.

the calculation of ΔE_{cp} in Table VII assumed that the decomposed AC and BC phases are macroscopic crystals. The domain-wall energy contributions were neglected because there are few boundaries. However, a realistic path between the completely separated and fully relaxed AC and BC domains is likely to pass through a sequence of intermediate states, including the stage of forming microscopic AC and BC clusters which serve as nucleation centers. As an approximation to this phase space immediately adjacent to the ordered ABC_2 alloy, we have estimated the metastable nucleation energy barrier by assuming that the microscopic AC and BC clusters are lattice matched to the ABC_2 alloy and their energetics can be estimated from the bulk crystals under the same constraint. Because the c/a ratios for the ordered alloy are nearly unity (see Table VIII), this epitaxial energy against nucleation of AC and BC clusters is equivalent to using a reference energy $E(AC)+E(BC)$ which disallows the c/a relaxation. Consequently, this epitaxial excess energy for the alloy, referred to as ΔE_{ep}^{H} (the hard model), is lowered further as shown in Table VIII. Note that the values for a number of alloys have already attained magnitudes that could account for the stability of the observed LRO of lattice-mismatched alloys. However, details of the mechanism that causes the LRO to preferentially form in the first place must still be determined.

B. Bond lengths

In addition to the excess energies, our calculations also produce detailed information about the equilibrium structures of the alloys. Table IX lists the c/a ratio and

TABLE VIII. The excess energy ΔE_{cp}^{H} (in meV per four atoms) for epitaxial alloys without (c/a)-ratio relaxation.

	CuAu I	Chalcopyrite	CuPt
$AlGaP_2$	−3.1	−3.4	−3.3
$AlGaAs_2$	0.7	0.6	0.2
$AlInP_2$	−192.5	−217.4	−147.3
$AlInAs_2$	−149.0	−172.2	−110.8
Ga_2PAs	−37.5	−48.1	−33.9
Ga_2PSb	−354.4	−447.1	−323.8
Ga_2AsSb	−162.2	−205.2	−147.1
$GaInP_2$	−200.4	−231.6	−149.5
$GaInAs_2$	−149.8	−174.8	−109.9
$GaInSb_2$	−113.9	−133.8	−85.3
$ZnCdTe_2$	−131.6	−143.5	−100.5
$ZnHgTe_2$	−150.3	−159.1	−125.1
$CdHgTe_2$	−2.8	−3.2	−3.2

the first-neighbor bond lengths for all the alloys studied. These results are particularly useful if experiments such as EXAFS are carried out to measure the local structures of these LRO alloys. As mentioned earlier, the c/a ratios are nearly unity for all alloys. The notations used in the table are such that the equilibrium bond lengths of the constituent crystals are in the order $d^0_{AC} > d^0_{BC}$. The bond lengths in the CA and CH structures are in general bimodal, with $d_{AC} > d_{BC}$. This result is similar to that found in the disordered bulk alloys. There are four different bond lengths in the CP structure: d^T_{AC}, d^S_{AC}, d^T_{BC}, and d^S_{BC}, where superscripts T and S mean the triple

TABLE IX. The equilibrium bond lengths d_{AC} and d_{BC} and the equilibrium average lattice constants a (all in Å). In CP structure, the first values d_{AC} and d_{BC} are for those bonds along the (111) direction, and the second values are for those in the other three directions.

ABC_2		c/a	a	d_{AC}		d_{BC}	
$AlGaAs_2$	CA	1.000	5.658	2.450		2.449	
	CH	1.000	5.656	2.450		2.449	
	CP	0.998	5.660	2.451	2.449	2.444	2.451
$AlGaP_2$	CA	1.000	5.457	2.365		2.362	
	CH	1.000	5.459	2.366		2.361	
	CP	0.998	5.461	2.371	2.362	2.354	2.366
$InGaSb_2$	CA	1.009	6.272	2.781		2.669	
	CH	0.995	6.298	2.788		2.660	
	CP	1.001	6.295	2.788	2.762	2.650	2.699
$InAlAs_2$	CA	1.013	5.831	2.597		2.479	
	CH	0.992	5.870	2.604		2.469	
	CP	1.001	5.861	2.599	2.578	2.456	2.512
$InGaAs_2$	CA	1.012	5.834	2.597		2.478	
	CH	0.993	5.866	2.605		2.468	
	CP	1.002	5.859	2.602	2.577	2.457	2.510
$InAlP_2$	CA	1.007	5.658	2.520		2.395	
	CH	0.996	5.674	2.527		2.384	
	CP	1.001	5.677	2.516	2.504	2.381	2.425
$InGaP_2$	CA	1.013	5.640	2.518		2.390	
	CH	0.993	5.672	2.526		2.379	
	CP	1.003	5.665	2.520	2.499	2.368	2.425
$AsPGa_2$	CA	1.001	5.549	2.430		2.379	
	CH	0.999	5.552	2.437		2.371	
	CP	0.998	5.556	2.450	2.422	2.354	2.390
$SbAsGa_2$	CA	1.003	5.880	2.601		2.499	
	CH	0.998	5.884	2.613		2.482	
	CP	0.995	5.905	2.637	2.585	2.439	2.531
$SbPGa_2$	CA	1.008	5.776	2.586		2.435	
	CH	0.994	5.787	2.603		2.407	
	CP	0.993	5.822	2.635	2.567	2.344	2.486
$CdHgTe_2$	CA	1.000	6.471	2.805		2.799	
	CH	1.000	6.471	2.806		2.798	
	CP	0.999	6.473	2.817	2.798	2.788	2.805
$HgZnTe_2$	CA	1.010	6.256	2.783		2.656	
	CH	0.995	6.286	2.787		2.650	
	CP	1.003	6.277	2.771	2.770	2.659	2.679
$CdZnTe_2$	CA	1.016	6.252	2.790		2.657	
	CH	0.991	6.302	2.795		2.650	
	CP	1.007	6.275	2.787	2.771	2.647	2.685

TABLE X. The ratios $\gamma \equiv (d_{alloy} - \bar{d})/(d_0 - \bar{d})$ for the ordered alloys calculated from the present TB model and comparison with experimental results for the disordered alloys.

	CA ETB		CH ETB		Disordered alloys Experiments	
ABC_2	γ_{AC}	γ_{BC}	γ_{AC}	γ_{BC}	γ_{AC}	γ_{BC}
$InGaSb_2$	0.71	0.64	0.80	0.76	0.89[a]	0.88[a]
$InGaAs_2$	0.71	0.66	0.80	0.78	0.77[a]	0.80[a]
$InGaP_2$	0.74	0.67	0.83	0.79	0.80[a]	0.76[a]
$AsPGa_2$	0.59	0.57	0.75	0.75	0.76[a]	0.75[a]
$HgZnTe_2$	0.81	0.76	0.86	0.84	0.72[b]	0.73[b]

[a]Reference 58.
[b]Reference 2.

and single bonds, respectively. However, the alloy lattice constants a are quite close to the mean value $\bar{a}$ of the constituent compounds. A more sensitive measure of the bimodal distribution, following Boyce and Mikkelsen,[1] is to look at the ratio $\gamma_{AC} = (d_{AC} - \bar{d})/(d^0_{AC} - \bar{d})$ with $\bar{d}$ being the average bond length, and similarly, γ_{BC}. Table X compares these ratios for the CA and CH structures with the experimental values for the disordered bulk alloys. Note that a value of $\gamma = 1$ corresponds to the totally relaxed case where there is no bond stretching, whereas $\gamma = 0$ corresponds to a rigid virtual crystal where all the atoms are on the zinc-blende crystal sites. In terms of the valence-force model, this ratio is roughly $\gamma = \alpha/(\alpha + 2\beta)$, where α and β [see Eq. (13)] are the average values for the bond-stretching and angular-restoring-force constants of the two constituent compunds. The calculated γ ratios in Table X follow this trend (see the α and β values in Table II). However, the values of γ show that CH structures are more relaxed than the CA structures. This result is consistent with the lower ΔE values in CH than in CA shown in Table VI. The ETB values of γ for $GaInAs_2$, $GaInP_2$, and Ga_2AsP in the CH structure are very close to the experimental values[2,58] for the disordered alloys. This result, when correlated with the lower excess energy, may suggest that the disordered alloys tend to favor a local configuration of the CH structure. However, the calculated γ values for $GaInSb_2$ and $HgZnTe_2$ do not correlate well with the experimental results. As a matter of fact, the trend as a function of the β/α ratio in the experimental results is reversed. It would be interesting to see if these two ordered alloys do have different bond lengths from the disordered states. While the values of γ in this table range from 0.6 to 0.9 for the lattice-mismatched alloys, this trend does not hold for the lattice-matched alloys. For example, Table IX shows that for Hg—Te and Cd—Te bond lengths in $HgCdTe_2$, both CH and CA structures nearly retain their respective constituent crystal values (i.e., $\gamma \simeq 1$), which was also seen in a recent experiment on the bulk alloy.[59] Finally, the bond lengths for the CP structure are characteristically different from those in CA and CH structures. The single bonds along the ordering direction [111] tend to be close to the constituent values while the triple bonds in the other directions have less relaxations (with γ values around 0.5 or less).

TABLE XI. Bulk moduli (in 10^{11} dyn/cm^2) of ordered alloys in three crystal structures calculated from the present TB model and the percentage deviations $\Delta B = (B - B_{av})/B_{av}$ from the average values B_{av} of the constituent compounds.

	B_{CH}	B_{CA}	B_{CP}	B_{av}	ΔB_{CH} (%)	ΔB_{CA} (%)	ΔB_{CP} (%)
$AlGaP_2$	8.858	8.858	8.854	8.8715	−0.150	−0.15	−0.201
$AlGaAs_2$	7.695	7.693	7.689	7.7085	−0.181	−0.20	−0.257
$AlInP_2$	7.876	7.860	7.774	7.9235	−0.605	−0.08	−1.882
$AlInAs_2$	6.705	6.691	6.661	6.7605	−0.828	−1.03	−1.475
Ga_2PAs	8.328	8.291	8.294	8.4165	−1.046	−1.50	−1.457
Ga_2PSb	6.584	6.297	6.188	7.4675	−11.836	−15.68	−17.135
Ga_2AsSb	6.314	6.198	6.157	6.7410	−6.342	−8.05	−8.662
$GaInP_2$	8.007	8.035	8.878	8.1950	−1.437	−1.95	−3.865
$GaInAs_2$	6.610	6.579	6.508	6.7420	−1.961	−2.42	−3.474
$GaInSb_2$	5.226	5.202	5.156	5.3115	−1.607	−2.07	−2.923
$ZnCdTe_2$	4.6105	4.6035	4.3375	4.6500	−0.85	−1.00	−6.72
$ZnHgTe_2$	4.8898	4.8872	4.6323	4.9245	−0.71	−0.76	−5.93
$CdHgTe_2$	4.4697	4.4721	4.4706	4.4845	−0.33	−0.28	−0.31

C. Bulk moduli

Table XI lists the results for the bulk moduli calculated from ETB. Also listed are the percentage deviations from the mean $\Delta B/\bar{B}$. Note that all ΔB values are negative. The magnitudes of ΔB are small except for Ga_2AsSb and Ga_2PSb which have large differences in both B and d between the constituent compounds. Although the magnitudes of ΔB get larger for systems with larger differences in the bond lengths, the dependence does not seem to be a simple function of the bond-length difference. The uniformly negative ΔB values were also obtained from LDF by Ferreira *et al.*[60] for the Ga_2AsSb alloys. One reason for the negative values of ΔB, in a very qualitative argument, is due to the fact that the bulk moduli of semiconductors scales inversely as high powers of the lattice constant,[61] and at the same time, the alloy lattice constant is approximated well by the mean value—the Vegard law. This qualitative behavior also comes out of a simple VFF analysis for disordered alloys.[62]

V. SUMMARY AND CONCLUSION

This work was motivated by our interest in understanding the mechanism of ordering for the semiconductor alloys grown by MBE and MOCVD. We have applied an empirical tight-binding (ETB) model to systematically interpolate the alloy total energies from those of the constituent compounds. Since the strain energy makes a dominant contribution to the excess energy ΔE, particular attention has been given the elastic properties, in addition to the lattice constants and cohesive energies. Our calculated bulk excess energies ΔE are positive for all alloys composed of lattice-mismatched compounds, and nearly zero for the lattice-matched systems. Based on these results, we conclude that all the ordered semiconductor alloys found experimentally are not in their thermodynamic stable states at the experimental growth temperatures. The same conclusion can also be drawn from the most recent LDF calculations[6,20,27–32,53–56] listed in Table VI. We note that several earlier theories[63–66] that concluded a stable ordering for these alloys have all been revised[54,67–70] (see also comments in Ref. 71).

Our calculation also generates detailed information about the structures and bulk moduli of these alloys. These results should be checked experimentally.

To further explore the stability of these alloys, we examined the energetics for spontaneous ordering when the grown materials are constrained to match a substrate lattice. Two kinds of epitaxial energies are calculated: One, denoted ΔE_{ep}, corresponds to the situation in which the constituent compounds are allowed to relax fully along the growth direction; and the other, the hard-model ΔE_{ep}^{H}, does not allow c/a relaxation. Although the sign of ΔE_{ep} can be negative, the magnitudes are too small to account for the observed spontaneous ordering. However, the sign and magnitudes of ΔE_{ep}^{H} for a number of lattice-mismatched alloys are found to be comparable to the energies needed for ordering at the growth temperatures. Whether or not this is a plausible mechanism deserves a more detailed investigation.

Even if ΔE_{ep}^{H} turns out to be a possible mechanism, based on the considerations presented in this paper we are still left without an explanation for the ordering in the lattice-matched alloys such as $GaAlAs_2$. We do not believe that any refined bulk calculation will produce an excess energy with large enough magnitude to account for the observed ordering in $GaAlAs_2$. On the other hand, epitaxial growth is very surface sensitive. Because of changes in the bonding character at surfaces, e.g., dangling bond, charge transfer, and reconstruction, the surface structural energies behave quite differently from those in the bulk, and do not extrapolate from the bulk, energies.[72] New mechanisms for spontaneous ordering may emerge from surface energetic considerations. Some hopeful thoughts[73] along this line have already been suggested.

ACKNOWLEDGMENTS

The work was supported in part by Office of Naval Research Contract No. N00014-88-C0096 and Air Force Office of Scientific Research Contract No. F4920-88-K-0009. The calculation was done by using the Alabama Supercomputer Network. The authors thank Dr. Alex Zunger for providing useful information.

APPENDIX: EXCESS ENERGY AND ORDERING

The excess energies ΔE, ΔE_{ep}, and ΔE_H considered in the text for the ordered alloys have to be considered along with the free energies of other possible phases including the disordered alloy in order to determine if the ordered phase is stable at a given temperature. A schematic mixing free-energy curve ΔF for a disordered pseudobinary alloy $A_{1-x}B_xC$ as a function of alloy composition x at several different temperatures is shown in Fig. 2. Figure 2(a) is for the case $\Delta E > 0$, and Fig. 2(b) for $\Delta E < 0$. (These curves are similar to Fig. 1 in Ref. 57.) If $\Delta E > 0$, then the stable phase is either the segregated phase or a uniform disordered solution depending on whether the temperature T is lower or greater than the

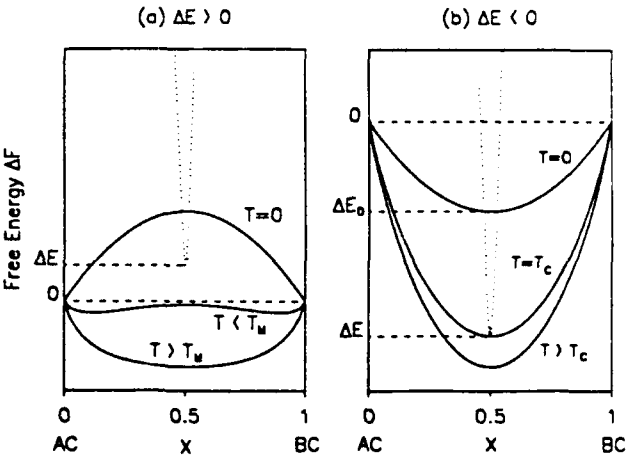


FIG. 2. Schematic plots of the mixing free energy of a disordered alloy as function of alloy composition x (dotted lines) for (a) $\Delta E > 0$ and (b) $\Delta E < 0$. The dashed line represented ΔE for an ordered alloy ABC_2.

critical temperature T_m for phase segregation or forming miscibility gaps. However, if $\Delta E < 0$, then the ordered phase is stable at a temperature T lower than the critical temperature T_c for the order-disorder transition.

A crude estimate of the magnitude of ΔE in Fig. 2(b) required for ordering at a given temperature T can be made by assuming that the disordered phase is a random alloy. Then in the present units of energy (for an ABC_2 molecule), the magnitude of the ordering energy defined as $\Delta E_0 = |\Delta E - \Delta E_D|$ must be greater than $2kT\ln 2$, where ΔE_D is the mixing energy for the disordered phase for the case at $x = \frac{1}{2}$ and $T = 0$, also shown in Fig. 2(b). At a typical epitaxial growth temperature of 600 °C (see Table I), the minimum ordering energy is estimated to be $\Delta E_0 \approx 100$ meV. Taking $\Delta E_D = \Delta E/2$, one finds a critical value of $\Delta E = -200$ meV. However, if a strict pair-potential nearest-neighbor model is used, $\Delta E_D = \frac{1}{4}\Delta E$ for both the CuAu I and chalcopyrite structures and the required ΔE value becomes -400 meV in this estimate.

Finally, consider the implication of different excess energies to ordering. Let us take Ga_2AsSb in the chalcopyrite structure as an example, for which the three excess energies are 70, -28, and -205 meV for ΔE, ΔE_{cp}, and ΔE_{cp}^H, respectively (see Tables VI, VII, and VIII) from the TB calculation. The positive value of ΔE corresponds to the case of Fig. 2(a), which shows that the ordered bulk phase is not stable at any temperature. Although the negative value $\Delta E_{cp} = -28$ meV corresponds to the case in Fig. 2(b), the corresponding critical temperature T_c is very low. This implies a simple constraint that the lateral lattice constants of the grown material matched to the substrate lattice are not enough to produce a metastable ordered alloy at the experimental growth temperature of 600 °C. However, the significant negative value of $\Delta E_{cp}^H = -205$ meV, resulting from a further constraint in which the c axis relaxation is not allowed, may raise the T_c in Fig. 2(b) into the range of the growth temperature and make the metastable ordered phase a possibility.

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

[2] A. Balzarotti, in *Ternary and Multinary Compounds*, Proceedings of the Seventh International Conference, edited by S. K. Deb and A. Zunger (Materials Research Society, Pittsburgh, Pennsylvania, 1987), p. 333.

[3] T. P. Pearsall, R. Carles, and J. C. Portal, Appl. Phys. Lett. **42**, 436 (1983).

[4] K. Kakimoto and T. Katoda, Appl. Phys. Lett. **40**, 826 (1982).

[5] A. Sher, M. van Schilfgaarde, A.-B. Chen, and W. Chen, Phys. Rev. B **36**, 4279 (1987).

[6] S.-W. Wei, L. G. Ferreira, and A. Zunger, Phys. Rev. B **41**, 8240 (1990).

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

[8] H. Nakayama and H. Fujita, in *GaAs and Related Compounds*, Proceedings of the 12th International Symposium, edited by M. Fujimoto, IOP Conf. Proc. No. 79 (Institute of Physics and Physical Society, London, 1986), p. 289.

[9] A. G. Norman, R. E. Mallard, I. J. Murgatroyd, G. R. Booker, A. H. Moore, and M. D. Scott, in *Microscopy of Semiconducting Materials*, edited by A. G. Gullis and P. A. Augustus, IOP Conf. Proc. No. 87 (Institute of Physics and Physical Society, London, 1987), p. 77.

[10] Y. Matsui, H. Hayashi, and K. Yoshida, Appl. Phys. Lett. **48**, 1060 (1986).

[11] T. S. Kuan, W. I. Wang, and E. L. Wilkie, Appl. Phys. Lett. **51**, 51 (1987).

[12] M. A. Shahid, S. Mahajan, D. E. Laughlin, and H. M. Cox, Phys. Rev. Lett. **58**, 2567 (1987). Note that these authors used the vapor-levitation-epitaxy (VLE) method.

[13] S. Yasuami, C. Nozaki, and Y. Ohba, Appl. Phys. Lett. **52**, 2031 (1988).

[14] A. Gomyo, T. Suzuki, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, Appl. Phys. Lett. **50**, 673 (1987); A. Gomoyo, T. Suzuki, and S. Iijima, Phys. Rev. Lett. **60**, 2645 (1988).

[15] P. Bellon, J. P. Chevalier, G. P. Martin, E. Dupont-Nivet, C. Thiebaut, and J. P. André, Appl. Phys. Lett. **52**, 567 (1988).

[16] O. Ueda, M. Takikawa, J. Komeno, and I. Umebu, Jpn. J. Appl. Phys. **26**, L1824 (1987).

[17] M. Kondow, H. Kakibayashi, and S. Minagawa, J. Cryst. Growth **88**, 291 (1988).

[18] H. R. Jen, M. J. Cherng, and G. B. Stringfellow, Appl. Phys. Lett. **48**, 1603 (1986).

[19] I. J. Murgatroyd, A. G. Norman, G. R. Booker, and T. M. Kerr, in *Proceedings of the Eleventh International Congress on Electron Microscopy, Kyoto, Japan, 1986*, edited by T. Imura, S. Maruse, and T. Suzuki (Japan Society of Electron Microscopy, Tokyo, 1986), p. 1497.

[20] S.-H. Wei and A. Zunger, Phys. Rev. B **39**, 6279 (1989).

[21] G. B. Stringfellow, J. Cryst. Growth **27**, 21 (1974); J. Appl. Phys. **54**, 404 (1983).

[22] M. B. Panish and M. Ilegems, Prog. Solid State Chem. **7**, 39 (1972).

[23] V. T. Bublik and V. N. Leikin, Phys. Status Solidi A **46**, 365 (1978).

[24] R. Patrick, A.-B. Chen, and A. Sher, Phys. Rev. B **36**, 6585 (1987).

[25] The magnitude of the excess energy ΔE required to achieve an order-disorder transition at a given transition temperature T_c varies from model to model [see D. de Fontaine, in *Solid State Physics: Advances in Research and Applications*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1979), Vol. 34, p. 73, particularly Table V], with the value ΔE_R for the regular-solution model being the smallest. The value of ΔE_R for the CuAu structure in the present scale is $3kT_c = 225$ MeV per ABC_2 molecule at $T_c = 900$ K. The actual ΔE should be larger than ΔE_R.

[26] A.-B. Chen and A. Sher, Phys. Rev. B **32**, 3695 (1985).

[27] J. E. Bernard, L. G. Ferreira, S.-H. Wei, and A. Zunger, Phys. Rev. B **38**, 6338 (1988).

[28] P. Boguslawski and A. Baldereschi, Phys. Rev. B **39**, 8055 (1989).

[29] T. Ohno, Phys. Rev. B **38**, 13 191 (1988).

[30] A. Qteish, N. Motta, and A. Balzarotti, Phys. Rev. B **39**, 5987 (1989).

[31] D. M. Bylander and L. Kleinman, Phys. Rev. B **34**, 5280 (1986); **36**, 3229 (1987).

[32] G. P. Srivastava, J. L. Martins, and A. Zunger, Phys. Rev. B **38**, 12 694 (1988).

[33]M. Methfessel, C. O. Rodriquez, and O. K. Andersen, Phys. Rev. B **40**, 2009 (1989).
[34]E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).
[35]G. B. Bachelet, D. R. Hamman, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[36]P. N. Keating, Phys. Rev. **145**, 637 (1966).
[37]R. M. Martin, Phys. Rev. B **1**, 4005 (1970).
[38]D. J. Chadi, Phys. Rev. B **19**, 2074 (1979); **29**, 785 (1984).
[39]W. A. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980).
[40]W. A. Harrison, Phys. Rev. B **27**, 3592 (1983).
[41]J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[42]L. Kleinman, Phys. Rev. **128**, 2614 (1962).
[43]A.-B. Chen and A. Sher, Phys. Rev. B **31**, 6490 (1985), Table I.
[44]R. Zallen, in *Handbook on Semiconductors*, edited by T. S. Moss (North-Holland, Amsterdam, 1982), Vol. 1, Chap. 1.
[45]J. C. Phillips, *Bonds and Bands in Semiconductors* (Academic, New York, 1973).
[46]*Numerical Data and Functional Relationship in Science and Technology*, edited by O. Madelung, Landolt-Börnstein, New Series, Group 3, Vol. 17 (Springer-Verlag, Berlin, 1982); *ibid.*, Vol. 22 (1988).
[47]All bond lengths are deduced from the lattice constants quoted by Ref. 44. The values of E_{bond} are taken from Table 7-3 in Ref. 39, except for AlSb, ZnTe, CdTe, and HgTe, which are deduced from Table 8.2 in Ref. 45. The elastic constants and the phonon frequencies are taken from values compiled in Ref. 46.
[48]Note that the parameters u_0, u_1, and u_2 and the calculated elastic constants based on Chadi's TB parameters in this table are not the same as Chadi's published (1979) values, which give bulk moduli about 20% smaller than the experimental values by considerably better phonon frequencies.
[49]O. H. Nielsen and R. M. Martin, Phys. Rev. Lett. **50**, 697 (1983); Phys. Rev. B **32**, 3792 (1985); **32**, 3780 (1985).
[50]C. S. G. Cousins, L. Staun, J. Staun, B. Selsmark, and B. J. Sheldon, J. Phys. C **20**, 29 (1987).
[51]D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).
[52]S. Froyen, Phys. Rev. B **39**, 3168 (1989).
[53]J. E. Bernard, R. G. Dandrea, L. G. Ferreira, S. Froyen, S.-H. Wei, and A. Zunger, Appl. Phys. Lett. **56**, 731 (1990).
[54]J. S. Nelson and I. P. Batra, Phys. Rev. B **39**, 3250 (1989).
[55]B. I Min, S. Massidda, and A. J. Freeman, Phys. Rev. B **38**, 1970 (1988).
[56]S. Ciraci and I. P. Batra, Phys. Rev. Lett. **58**, 14 (1987); Phys. Rev. B **36**, 1225 (1987).
[57]C. P. Flynn, Phys. Rev. Lett. **57**, 599 (1986).
[58]J. B. Boyce and J. C. Mikkelsen, in *Ternary and Multinary Compounds*, Proceedings of the Seventh International Conference, edited by S. K. Deb and A. Zunger (Materials Research Society, Pittsburgh, Pennsylvania, 1987), p. 359.
[59]W.-F. Pong, R. A. Mayanovic, B. A. Bunker, J. K. Furdyna, and U. Debska, Phys. Rev. B **41**, 8440 (1990).
[60]L. G. Ferreira, S.-H. Wei, and A. Zunger, Phys. Rev. B **40**, 3197 (1989).
[61]Based of the VFF analysis, Martin in Ref. 37 shows that B scales as $1/d^5$, while in an empirical formula by M. L. Cohen [Phys. Rev. B **32**, 7988 (1985)] B scales as $1/d^{3.5}$.
[62]A.-B. Chen, A. Sher, and M. A. Berding, Phys. Rev. B **37**, 6285 (1988).
[63]G. P. Srivastava, J. L. Martins, and A. Zunger, Phys. Rev. B **31**, 2561 (1985).
[64]J. S. Cohen and A. G. Schlijper, Phys. Rev. B **36**, 1526 (1987).
[65]W. I. Wang, Appl. Phys. Lett. **58**, 3244 (1985).
[66]M. van Schilfgaarde, A.-B. Chen, and A. Sher, Phys. Rev. Lett. **57**, 1149 (1986).
[67]J. S. Cohen and A. G. Schlijper, Phys. Rev. B **38**, 12 694 (1988).
[68]A.-B. Chen, M. van Schilfgaarde, and A. Sher, Bull. Am. Phys. Soc. **31**, 664 (1986).
[69]D. M. Wood, S.-H. Wei, and A. Zunger, Phys. Rev. B **37**, 1342 (1988).
[70]S.-H. Wei, Phys. Rev. Lett. **59**, 2613 (1987).
[71]The negative ΔE values obtained from LDF in Refs. 63 and 64 were revised in Refs. 67 and 54, respectively. The negative Madelung energy in the ordered alloy proposed in Ref. 65 only considered the contributions from the alloying species. It was pointed out in Refs. 68 and 69 that if all species are considered, the dipolar energy for both ordered and disordered alloys are positive. Finally, the suggestion in Ref. 68 that the total Coulomb energy for a disordered alloy might exceed the ordered system was also revised in Ref. 70.
[72]S. Krishnamurthy, M. A. Berding, A. Sher, and A.-B. Chen, Phys Rev. Lett. **64**, 2531 (1990).
[73]J. E. Bernard, R. G. Dandrea, L. G. Ferreira, S. Froyen, S.-H. Wei, and A. Zunger, Bull. Am. Phys. Soc. **35**, 667 (1990).
[74]R. G. Dandrea, J. E. Bernard, S.-H. Wei, and A. Zunger, Phys. Rev. Lett. **64**, 36 (1990).

$Hg_{1-x}Cd_xTe$: DEFECT STRUCTURE OVERVIEW

M.A. BERDING*, A. SHER*, AND A.-B. CHEN**
*SRI International, Menlo Park, California 94025
**Auburn University, Auburn, Alabama 36849 – 3501

ABSTRACT

Native point defects play an important role in HgCdTe. Here we discuss some of the relevant mass action equations, and use recently calculated defect formation energies to discuss relative defect concentrations. In agreement with experiment, the Hg vacancy is found to be the dominant native defect to accommodate excess tellurium. Preliminary estimates find the Hg antisite and the Hg interstitial to be of comparable densities. Our calculated defect formation energies are also consistent with measured diffusion activation energies, assuming the interstitial and vacancy migration energies are small.

INTRODUCTION

Sophisticated infrared devices are currently made from HgCdTe. While the material has entered a manufacturing phase, there remain many unanswered questions about its nature. Native point defects appear to play a far more important role in the properties of HgCdTe than they do in other technologically important semiconductor materials. The formation energies of these defects and their associated localized states in the band gap are still not completely characterized. Similarly, the characterization of impurities in HgCdTe is far from complete. Beyond point defects it is well known that a number of extended defects are important. The best studied are tellurium inclusions and dislocations which form at heterojunctions and at HgCdTe-substrate interfaces. As grown material is tellurium-rich, the excess Te is accommodated in the form of Hg vacancies and tellurium inclusions. As such, post growth anneals in a Hg vapor overpressure are typically performed to reduce the Hg vacancies prior to other device processing; this converts high purity undoped material from p-type to n-type. The donor causing the n-type doping has not been identified, although it is believed to be a residual impurity, rather than a native defect.

One of the more perplexing properties of HgCdTe stems from the fact that it is an alloy. Until a few years ago it was thought that all tetrahedrally coordinated zincblende semiconductor alloys of the form $A_{1-x}B_xC$ were random, with the C atoms occupying one fcc sublattice and the A and B atoms sharing the other sublattice at random. This has now been demonstrated to be incorrect in many semiconductor alloys. The nature of the correlations can be characterized by counting the populations of those clusters consisting of a centered C atom and its four surrounding A or B atoms. There are five types of clusters of the form $A_{4-n}B_nC$ with n=0, 1, 2, 3, or 4. If the alloy is random the cluster occupation probabilities, p_n, form the binomial distribution, i.e. $p_{n,random} = \binom{4}{n}(1-x)^{4-n}x^n$. In correlated material, deviations from this distribution are found. For example, if at x=0.25 the material tends towards a regular compound, then p_1 is increased while the other cluster occupation probabilities are reduced relative to $p_{n,random}$. If the

Mat. Res. Soc. Symp. Proc. Vol. 216. ©1991 Materials Research Society

material tends towards spinodal decomposition, then p_0 and p_4 are enhanced in separate regions of space, and the other cluster occupation probablilities are reduced. In most semiconductor alloys, the major forces for correlations are strain energies resulting from bond length differences between the AC and BC constituents. Thus theory predicts that HgCdTe will be nearly random because HgTe and CdTe are nearly lattice matched. Calculations that include chemical energies and charge shifts, in addition to the strain energies, have also predicted HgCdTe to be a nearly random alloy. Yet several recent experiments on alloys with x=0.20-0.25 have suggested that HgCdTe is correlated, with $p_n > p_{n,random}$ for n=0 and 1, and $p_n < p_{n,random}$ for n=2, 3, and 4. If this is the case it may have important consequences on diffusion and transport properties. The consequences of such correlations on the defect populations are discussed further below.

NATIVE DEFECTS

The doubly ionized cation vacancy is believed to be the dominant native defect in HgCdTe, that dictates the electrical behavior of the undoped material [1, 2]. We have recently reported the results of calculations of native and impurity defect total energies in HgTe, CdTe and ZnTe [3]. The energies for the formation of various native point defects in HgTe are summarized in Table I; the HgTe solid and the Hg in the vapor are used as references [4]. These calculations used the linearized muffin tin orbital (LMTO) method within the local density approximation (LDA) to the exchange correlation energy. Large supercells containing one defect per cell were repeated periodically, and from the difference in total energies per cell, with and without the defect, the defect formation energy was calculated. To expedite the calculations, the energies were calculated within atomic spheres approximation (ASA) with a small (*spd*) basis set. In the ASA, an approximation to the exact density functional is evaluated; as a result, an error is introduced which is larger than in other LDA methods [5], and relaxation energies cannot be accurately calculated. Thus only those differences in energies here that are $\geq$ 0.5 eV here should be viewed as significant for these calculations.

An appropriate set of mass action constants for the neutral defect reactions is also given in Table I. The notation in that table is as follows. A_B corresponds to an A species occupying a B site, where I corresponds to an interstitial, and V to a vacancy. No subscript on a species indicates it is occupying the correct lattice site, e.g., Hg = Hg_{Hg}. Square brackets, [ ], refer to concentrations. A subscript "g" indicate the species in the gaseous, or vapor, phase and P_{Hg} is the Hg vapor pressure. Most of the reactions in Table I involve the creation or destruction of one or more unit cells. Because the resulting change in volume is accommodated at the surface, the change in the number of unit cells will enter into the determination of the defect equilibrium through the surface entropy. Additionally, surface preparation and orientation will affect the surface free energy. We have assumed for the present that such surface effects are neglible, i.e., that the volume expansions and contractions can occur with negligible changes in the surface properties.

To correctly complete the defect equilibrium determination one must include the equilibration of the electronic charges of the system. To do so, one must have knowledge of the dominant charge states of the defects, and their activation energies with respect to the neutral defect. Such calculations are complicated by the fact that most *ab initio* calculations of the electronic band structure of semiconductors predict an incorrect band gap, E_G, a shortcoming of the local density approximation (LDA.) As such, we shall focus

Table I. Defect reactions and formation energies

Defect reaction	Defect concentration	Energy (eV)
$E_{V_{Hg}} + HgTe \rightleftharpoons V_{Hg}Te + Hg_g$	$[V_{Hg}] = P_{Hg}^{-1} K^0_{V_{Hg}} \exp(\frac{-E_{V_{Hg}}}{kT})$	2.24*
$E_{Te_{Hg}} + 2HgTe \rightleftharpoons Te_{Hg}Te + 2Hg_g$	$[Te_{Hg}] = P_{Hg}^{-2} K^0_{Te_{Hg}} \exp(\frac{-E_{Te_{Hg}}}{kT})$	4.53
$E_{Te_i} + HgTe \rightleftharpoons Te_i + Hg_g$	$[Te_i] = P_{Hg}^{-1} K^0_{Te_i} \exp(\frac{-E_{Te_i}}{kT})$	4.96
$E_{V_{Te}} + Hg_g \rightleftharpoons HgV_{Te}$	$[V_{Te}] = P_{Hg} K^0_{V_{Te}} \exp(\frac{-E_{V_{Te}}}{kT})$	3.12**
$E_{Hg_{Te}} + 2Hg_g \rightleftharpoons HgHg_{Te}$	$[Hg_{Te}] = P_{Hg}^2 K^0_{Hg_{Te}} \exp(\frac{-E_{Hg_{Te}}}{kT})$	-0.42
$E_{Hg_I} + Hg_g \rightleftharpoons Hg_I$	$[Hg_I] = P_{Hg} K^0_{Hg_I} \exp(\frac{-E_{Hg_I}}{kT})$	0.84, 0.98

* Experimental number from Vydyanath [1, 2].
** Calculated using tight binding Hamiltonian[6].

on the neutral native defects here, and the established or expected charge states of these defects. In wide band gap materials the defects equilibration can be substantially affected by the Fermi level; for example the formation energy of a donor will decrease when the Fermi energy is near the valence band edge, since the donor electron can drop into a vacant state near the valence band, thereby lowering the energy by $\sim E_G$. Because we are discussing HgCdTe with a narrow band gap, we expect the Fermi effects to be small, but not insignificant at high temperatures. Because of the small conduction band effective mass, though, in n-type material the filling of the conduction band states by electrons can shift the Fermi energy significantly; combined with the increase in the band gap for the high temperatures at which most defect studies are done, the effective band gap can be significantly larger than the usual 77 K bandgap associated with a given concentration of HgCdTe.

First we consider the defects which accommodate excess tellurium, the first three defects in Table I. The mass action constants are given by a product of the form

$$K_i = K_i^0 e^{\frac{-\Delta E_i}{kT}} \tag{1}$$

where ΔE_i is the change in the enthalpy for the i^{th} reaction. For the first three equations, the K_i^0 are given by

$$K^0_{V_{Hg}} = C_o(kT)^{\frac{5}{2}}(2\pi m_{Hg})^{\frac{3}{2}}h^{-3}\exp(\frac{\Delta S_{V_{Hg}}}{k}) \quad , \tag{2}$$

$$K^0_{Te_{Hg}} = C_o(kT)^5(2\pi m_{Hg})^3h^{-6}\exp(\frac{\Delta S_{Te_{Hg}}}{k}) \quad , \tag{3}$$

and

$$K^0_{Te_I} = C_o(kT)^{\frac{5}{2}}(2\pi m_{Hg})^{\frac{3}{2}}h^{-3}\exp(\frac{\Delta S_{Te_I}}{k}) \quad . \tag{4}$$

Here T is the temperature in Kelvin, k is Boltzmann's constant, m_{Hg} is the mass of the mercury vapor atoms, h is Planck's constant, ΔS_i is the change in vibrational entropy upon formation of the defect, and C_o converts from site fraction to volume concentrations. Because two unit cells of HgTe are destroyed when a tellurium antisite is created, compared to one unit cell when a mercury vacancy is created, we do not expect that $\exp((\Delta S_{Te_{Hg}} - \Delta S_{V_{Hg}})/kT) \sim 1$. While we have not completed the evaluation of these entropy terms, our preliminary estimates indicate this ratio is $\sim 10^{-4}$. For the tellurium interstitial and the mercury vacancy we expect that $\exp((\Delta S_{Te_i} - \Delta S_{V_{Hg}})/kT) \sim 1$ will be correct within a factor of 10. Evaluating the numerical constants, we find

$$\frac{[Te_{Hg}]}{[V_{Hg}]} \simeq 10^{-10} \tag{5}$$

and

$$\frac{[Te_i]}{[V_{Hg}]} \simeq 10^{-18} \tag{6}$$

for T=500 C and $P_{Hg} = 1$ atm. The conclusion from Equations 5 and 6 that the mercury vacancy is the dominant native defect is consistent with experimental observation. This conclusion is unchanged if we include the possibility that the species may be ionized at the equilibration temperature where the material is expected to be intrinsic. Although the tellurium antisite density decreases more rapidly with decreasing Hg pressure than does the mercury vacancy density, the point at which the concentrations are comparable is at less than $P_{Hg} \simeq 10^{-10}$ atm, and certainly the HgTe phase boundary is reached before such low Hg pressures can be achieved. This is also consistent with the fact that no p-to-n conversion is observed for low mercury pressures, as would be expected if tellurium antisites became the majority native defect. Because the pressure dependences of the tellurium interstitial and the mercury vacancy concentrations are the same, the above conclusion will hold independent of the mercury pressure.

We have checked the sensitivity of the calculated concentration ratios to the magnitude of the reaction enthalpy. Because the enthalpies enter in the exponents, small changes in the enthalpies will result in large changes in the predicted defect concentrations. For example, let us assume that our calculated antisite formation enthalpy is in error by 0.5 eV; in this case the ratio of antisite to vacancy concentrations (at 500 C) will be reduced to $\sim 1 \cdot 10^{-11}$. For an antisite formation enthalpy in error by 1.0 eV, this ratio is reduced to $\sim 1 \cdot 10^{-7}$. We do not expect the ASA errors to exceed 0.5 eV [5].

If the HgCdTe is not completely annealed, and tellurium precipitates are still present, the defect equilibrium will not be that predicted by the mass action equations given in Table I. For example, near the inclusions we can assume that the defects will be nearly in equilibrium with the tellurium solid; thus

$$E'_{V_{Hg}} + Te_s \rightleftharpoons V_{Hg}Te \tag{7}$$

and

$$E'_{Te_{Hg}} + 2Te_s \rightleftharpoons Te_{Hg}Te \tag{8}$$

will be the appropriate reactions. The formation energies for a tellurium antisite and a Hg vacancy from the tellurium solid are 1.63 eV and 1.15 eV, respectively. Although the

difference in the formation energies isless than when both defects are referenced to the mercury vapor (~ 0.5 eV compared to ~ 2 eV), the gas phase entropy factor does not enter into the ratio of the defect concentrations. Using the same estimate of the entropy ratio, the defect concentration ratio using tellurium solid as the reference state is

$$\frac{[Te_{Hg}]}{[V_{Hg}]} \simeq 10^{-8} \quad . \tag{9}$$

Thus, near the inclusions we expect higher relative concentration of tellurium antisites, as compared to the rest of the material equilibrated with the Hg vapor. Additionally, the absolute defect concentrations may differ substantially in the two regions of the material. A better calculation of the vibrational entropy is needed before we can predict these absolute defect concentrations. Differences in the defect concentrations arising from different equilibration condition are a possible source of spatial variability of the HgCdTe material. If the material is not fully annealed to equilibrium, for example because of an abundance of tellurium precipitates, this history may affect subsequent processing.

In the above we have discussed the defect energies for HgTe and applied them directly to the small x, $Hg_{1-x}Cd_xTe$ system. Because we are dealing with the native defects of an alloy we expect a number of complexities to affect the above analysis. First the variation of the defect formation energies for vacancies is sensitive to the alloy environment, in particular for the vacancies of the nonsubstituted species, such as tellurium in HgCdTe [6]. Even for vacancies of the substituted species, we have found the formation energies may vary by several tenths of an electron volt. Because of this variation in the formation energy, the fraction of defective sites will vary by as much as a factor of 100 from one class of sites to the next. Consider various classes of Hg sites in ideal HgCdTe, which can be distinguished to first order by specifying the number of Hg and Cd atoms in the second neighbor shell (the four first neighbors are always tellurium), $Hg_{12-j}Cd_j$, with a concentration given by [j]. The total vacancy concentration is given by

$$[V_{Hg}] = \sum_{j=1}^{12} [j] P_{Hg}^{-1} K_j^0 \exp\left(\frac{-E_j}{kT}\right) \tag{10}$$

where E_j is the vacancy formation energy for the j^{th} cluster. The populations of vacancies in each class of cluster, j, can be expected to differ because of differences in the cluster populations, [j], and the formation energies, E_j. Additionally, the defect energy levels may differ in the various classes of sites, possibly leading to different ionization states for vacancies in different classes of sites. If the cations in the alloy are randomly arranged, such differences may be difficult to infer experimentally. If, on the other hand, the cations are correlated, exhibiting short range order, more complex behavior may be present. Such short range order has been demonstrated using Raman scattering [7, 8], infrared reflectivity [9], and nuclear magnetic resonance [10]. In these cases, the contribution to the vacancy densities from the dominant class of clusters will be increased. Because the studies finding short range order focus on the tellurium centered 5-atom clusters of the form $Hg_{4-n}Cd_n$ rather than on cation centered clusters of the form $Hg_{12-n}Cd_n$, higher level 5-atom cluster-cluster correlations must be known to predict the effects on the vacancy populations.

Next we examine the defects which accommodate excess Hg in the solid. The existence region for HgCdTe is always tellurium rich and thus the native defects which accommodate excess tellurium are expected to dominate. For these equations in Table I, K_i^0 is given by

$$K^0_{V_{Te}} = C_o^{-1}(kT)^{-\frac{3}{2}}(2\pi m_{Hg})^{-\frac{3}{2}}h^3\exp(\frac{\Delta S_{V_{Te}}}{k}) \quad , \tag{11}$$

$$K^0_{Hg_{Te}} = C_o^{-1}(kT)^{-5}(2\pi m_{Hg})^{-3}h^6\exp(\frac{\Delta S_{Hg_{Te}}}{k}) \quad , \tag{12}$$

and

$$K^0_{Hg_i} = C_o(kT)^{-\frac{3}{2}}(2\pi m_{Hg})^{-\frac{3}{2}}h^3\exp(\frac{\Delta S_{Hg_i}}{k}) \quad . \tag{13}$$

If we assume the change in entropy is comparable for all three defects, we find

$$\frac{[Hg_{Te}]}{[V_{Te}]} \simeq 10^{+14} \tag{14}$$

and

$$\frac{[Hg_i]}{[V_{Te}]} \simeq 10^{+14} \quad , \tag{15}$$

for T=500 C and $P_{Hg} \simeq 1$ atm. From Equations (14) and (15) we see that the tellurium vacancy, $[V_{Te}]$, is a minority defect species. For the pressure and temperature considered, the density of Hg antisites is predicted to be comparable to the density of Hg interstitials. Because the ratio of $[Hg_{Te}]$ to $[Hg_i]$ is nearly unity, any errors in the calculation of the activation energy could push the balance toward one side or the other. Thus we must depend on the next generation of calculations, with the ASA removed and full relaxation included plus a quantitative comparison of the entropy differences between the mercury antisite and the tellurium interstitial, to determine the dominant defect in this class.

As mentioned above, we have shown that the tellurium vacancy formation energy varies significantly with the alloy environment. Because the tellurium vacancy is not expected to be a dominant defect in HgCdTe, and the tellurium diffuses by an interstitial mechanism, we do not expect any measurable manifestation of this variation. On the other hand, the Hg antisite may be the dominant Hg-excess defect, and its formation energy may vary significantly with the alloy environment. We are currently calculating the magnitude of this variation.

DIFFUSION

HgCdTe exhibits a complex tracer diffusion profile, with both a fast and a slow branch. The fast branch is attributed to a vacancy and interstitial diffusion in parallel mechanism where the dominant diffuser is determined by the pressure and temperature, while the slow component fits a vacancy and interstitial in series mechanism [11]. The activation energies for the fast branch are 2.10 eV and 0.61 eV for the vacancy and interstitial mechanisms, respectively. Our calculated formation energy for the mercury interstitials are 0.89 and 0.98 eV for the anion- and cation-surrounded tetrahedral interstitial sites, respectively, and the experimental formation energy for the mercury vacancy is 2.24 eV. Comparing these energies to the experimental activation energies we find close agreement, indicating that the migration energy contribution to the diffusion activation energies are small for both interstitials and vacancies.

In a recent experiment on mercury diffusion in ion-implantation damaged HgCdTe, an activation energy of several tenths of an electron volt was measured [12]. The disparate result can be interpreted as a measure of only the defect migration contribution to the diffusion activation energy, since defects in excess of the equilibrium concentration were likely formed during implantation. It is not evident that the measured activation energy corresponds to the vacancy or the interstitial mechanism. The conclusion that the diffusion activation energies are largely defect formation energies, with the migration energies being much smaller, is in agreement with the above interpretation of the Richter and Kalish [12] experiment.

CONCLUSIONS

We have incorporated our calculated defect energies into the mass action equations for the neutral defects in HgCdTe. In agreement with experiment, we find the mercury vacancy to be the dominant native defect in tellurium rich material. We also find the mercury antisite and interstitial defect densitities to be comparable, although a better calculation of the vibrational entropy is needed to confirm this result. Comparing the defect formation energies to the diffusion measurements by Tang and Stevenson [11], we find agreement with their diffusion activation energies for both the vacancy and the interstitial mechanism, if we assume that the migration energy is small. Further work is in progress to incorporate the defect charge states into the calculation, and also to calculate the fully relaxed defect energies with the full potential LMTO.

The work was supported by NASA contract NAS1-18226, by ONR contract N00014-88-C0096, and by AFOSR contract F49620-88-K-0009.

REFERENCES

[1] H. R. Vydyanath, J. Electrochem. Soc. 128, 2609 (1981).

[2] H. R. Vydyanath, J.C. Donovan, and D.A. Nelson, J. Electrochem. Soc. 128, 2625 (1981).

[3] M.A. Berding, M. van Schilfgaarde, A.T. Paxton, and A. Sher, J. Vac. Sci. Technol. A 8, 1103 (1990).

[4] A different reference is used here than was used in Table III in Reference 3. An error appears in that table owing to the incorrect use of an energy of $2E_b$ per unit cell rather than $4E_b$. The defect energies in Table I in Ref. 3 from which the energies in Table III were derived, are correct.

[5] A. Sher, M. van Schilfgaarde, and M.A. Berding, presented at the 1990 HgCdTe Workshop in San Francisco, CA, and accepted for publication in J. Vac. Sci. Technol.

[6] M.A. Berding, A. Sher, and A.-B. Chen, J. Appl. Phys. 68, 5064 (1990).

[7] P.M. Amirtharaj and F.H. Pollak, Appl. Phys. Lett. 45, 789 (1984).

[8] A. Compaan, R.C. Bowman, and D.E. Cooper, Appl. Phys. Lett. 56, 1055 (1990).

[9] L.K. Vodopyanov, S.P. Kozyrev, Y. A. Aleshchenko, R. Triboulet, and Y. Marfaing, Appl. Phys. Lett. 56, 1057 (1990).

[10] D. Zamir, K. Beshah, P. Becla, P.A. Wolff, R.G. Griffin, D. Zax, S. Vega, and N. Yellin, J. Vac. Sci. Technol. 6, 2612 (1988).
[11] M.-F.S. Tang and D.A. Stevenson, J. Vac. Sci. Technol. 7, 544 (1990), and references therein.
[12] V. Richter and R. Kalish, J. Appl. Phys. 67, 6578 (1990).

Surface energies for molecular beam epitaxy growth of HgTe and CdTe

M. A. Berding, Srinivasan Krishnamurthy, and A. Sher
SRI International, Menlo Park, California 94025

(Received 5 November 1990; accepted 18 December 1990)

We present results for the surface binding energies for HgTe and CdTe that will serve as input for molecular beam epitaxy growth models. We have found that the surface binding energies are surface orientation dependent and are not simply proportional to the number of first-neighbor bonds being made to the underlying layer. Moreover, because of the possibility of charge transfer between cation and anion surface states, one may have large differences between the binding energy for the first and the last atom in a given layer, and these differences will be different for the narrow-gap, less ionic materials than for the wide-gap, ionic materials. We also find that the surface states associated with an isolated surface atom or vacancy are extended in materials with small gaps and small effective masses, and thus call into question the modeling of surface binding by simple pair interactions.

While considerable advances have been made in recent years in molecular beam epitaxy (MBE) growth of HgCdTe, modeling of the growth process is still primitive because of the complexity of the process. One important input common to all growth models is the binding energy of atoms to the particular growth surfaces. These energies are usually approximated as being equal to the number of first-neighbor bonds made to the surface times the energy per bond in the bulk. We have shown[1] that this is in fact a very poor approximation, and that the surface binding energies often differ considerably from those estimates, and are sensitive to the local surface stoichiometry. In this article, several implications of these results on the nature of the growth surface of HgTe and CdTe and their alloys are discussed.

We have calculated[1] the surface binding energies in two limits. The first is the *concentrated* limit in which the atom is added to complete the surface layer, resulting in an ideal surface. The second is the *dilute* limit in which an isolated atom is added to an underlying ideal surface; this limit corresponds to the initiation of a new layer. The surface binding energies were calculated using a tight-binding Green's function technique that has been presented elsewhere.[2,3] Surface binding energies of the constituents in these two limits for HgTe and CdTe were calculated for the (111)B and (100) surfaces. Results are summarized in Table I. Also shown for reference are the experimental bulk energies per bond.

The results in Table I demonstrate that surface binding energies are not proportional to the number of bonds made to the surface. Because no in-plane bonds are made on the (111) and (100) growth surfaces, this is evident by the fact that the binding energies in the dilute (E_d) and concentrated (E_c) limits are not in general equal, as they would be if the simple proportionality relationship held. Even when the dilute and concentrated energies are equal, as in the case of tellurium on the HgTe (111)B surface, these energies are not equal to three times the bulk energy per bond. Although the surface binding energies are not linearly proportional to the number of bonds being made to the surface, we do find that averaging over the dilute and concentrated cation and anion binding energies for a particular surface yields the cohesive energy per bulk layer, as it should.

Differences in the E_a and E_c can be attributed, in part, to charge transfer effects on the surface.[1,2] Briefly, because the anion surface states lie at lower energies than the cation surface states, surface cations will always transfer charge to surface anions, until all the surface anion states are fully occupied. Thus, for example, for cadmium on the (100) surface, differences in the binding energies result because in the dilute limit the cation added to the surface can transfer charge to the surface anions, thereby lowering the Cd surface binding energy, while in the concentrated limit, no surface anion states are available into which to transfer charge. In contrast, in the concentrated limit electrons must be promoted out of the anion surface states (which had been transferred there from other cations already present on the surface) when the cation completing the layer is added. Both of these result in a stronger binding of the cation in the dilute limit, as seen in Table I. This charge transfer effect will be largest in ionic semiconductors with large band gap because of the large difference in anion and cation surface state energies.

From Table I we note that on both the (111)B and (100) surfaces mercury is more weakly bound in the dilute limit than in the concentrated limit. As a consequence, the initiation of a new layer of mercury atoms will be more difficult than the completion of a mercury layer, and thus initiation of islands in layer-by-layer growth will be disfavored with respect to the completion of islands. The Te is strongly bound to both the (111)B and (100) surfaces, indicating that the tellurium-stabilized growth will be preferred.

For CdTe, we see that the nucleation of a new layer will not be as troublesome as in HgTe, as indicated by the large values for E_d in Table I, while the completion of the surface will be more difficult. This has an important consequence on the nature of the growing surface. Because $E_d < E_c$ for Cd in CdTe, the cations on the surface prefer to nucleate new layers, as opposed to completing layers. This can be interpreted as an effective repulsive surface interaction among the Cd atoms on the surface. This may have important consequences on the growth of this and other wide-gap ionic materials such as ZnTe. Because the surface atoms prefer not to sit adjacent to one another, there is the possibility of the atoms on the surface forming a superlattice with the surface

TABLE I. Surface binding energies for HgTe and CdTe (eV).

		(111)B		(100)		
		Dilute	Concentrated	Dilute	Concentrated	Bulk
HgTe	Hg	~0.0	−0.3	~0.0	−1.0	−0.82
	Te	−2.8	−2.8	−2.4	−3.1	
CdTe	Cd	−2.2	−1.3	−2.1	−0.1	−1.10
	Te	−5.2	−2.1	−5.0	−2.6	

vacancies. This phenomenon has been predicted and observed on GaAs, where the same mechanism is responsible.[4,5] The nature of this surface superlattice will depend on the particular surface, the magnitude of the energy differences between E_d and E_c, and the temperature of the growing surface.

Because the surface binding energies are dependent on the surface stoichiometry in so far as it controls the states available for charge transfer, we expect the contributions to surface binding energies to change relatively abruptly at some specific surface coverage. As an example we consider the addition of Cd to the (100) surface of CdTe. As discussed above, when an isolated Cd atom is added to the surface, the binding energy is lowered because of the charge transfer to adjacent Te atoms, while when Cd is added a to nearly complete Cd surface, the binding energy does not benefit from such a charge transfer because there are no empty Te surface states. At 50% surface coverage, the number of Cd and Te surface states will be equal, with the lower-energy Te states being completely filled and the higher-lying Cd states being empty. Now we consider an arbitrary intermediate surface coverage to determine if the Cd surface binding energy will benefit from a charge transfer to adjacent Te. If the surface coverage is below 50%, empty Te states will still be available for charge transfer, while above 50% coverage, all Te states will be already full. Thus, near 50% we expect the value of the surface binding energies to change rather abruptly, with binding energies below 50% coverage being closer to E_d and binding energies for coverages above 50% being closer to E_c. The situation is similar for other polar surfaces. For example, for addition of Te to the (111)B surface the changeover between dilute-like and concentrated-like binding energies occurs at 75% coverage. This is because the density of surface states per cation on the surface is three times the density of surface states per anion. Thus, when the surface is 75% anions, the density of cation and anion surface states will be equal, and full charge transfer will occur. This is similar to the results found by Chadi[4] and Tong *et al.*[5] for GaAs.

Because the charge transfer depends on the band gap of the material, we expect the behavior in HgCdTe and HgZnTe with band gaps in the infrared to have effective attractive surface interactions. This is consistent with preliminary calculations using the supercells/slab method.[6] Thus, at sufficiently low temperatures, although higher than temperatures for which surface diffusion is too slow for surface equilibration, the surfaces will grow via the formation of smooth islands of like atoms. As the surface temperature is raised, entropy will dominate and the surface will be rough, with smooth islands being replaced by disordered, randomly arranged atoms.

The lateral extent of the surface states produced by an isolated surface atom or vacancy is found to differ for HgTe and CdTe. This has been demonstrated using the same tight-binding Hamiltonian in a slab calculation[6] and in a Green's function calculation. In the slab calculation, surfaces are modeled by constructing multiple layers of the semiconductor and vacuum, where the number of vacuum layers is large enough to completely decouple the two surfaces and the number of semiconductor layers is chosen large enough so that the center layer looks bulk like. In the Green's function method, a truly semiinfinite surface can be created. Although the slab method does not mimic a real surface as well as the Green's function calculation, it permits the calculation of the formation energy for a periodic array of defects. To examine the in-plane coupling of the vacancies, a regular array of Hg atoms was removed from the (111)A HgTe surface, where each surface vacancy created was completely surrounded by atoms. This removal energy per atom is expected to be equal to $-E_c$, if the wave functions are localized at the removal site. However, we find that the energy required (per atom from a unit cell of four atoms) to remove a regular array of Hg atoms from the (111) HgTe A surface differs substantially from $-E_c$. When the size of the supercell is increased to include nine atoms, the calculated energy (per atom) is larger than the previous one but still smaller than E_c. This is because the created surface vacancies are far from each other and consequently couple less when compared to the previous case. However, when these calculations were repeated for removal of cations from CdTe and ZnTe surfaces, the removal energies calculated with nine atoms per unit supercell agreed exactly with E_c. It suggests that the surface wave function is well extended on HgTe surfaces, but terminates near the third neighbor on CdTe and ZnTe surfaces. Whenever surface wave functions have a large spatial extent, coupling substantially to one another, the removal energies per atom will be affected.

The difference in spatial extent on HgTe and CdTe surfaces can be understood from the complex band structure of these materials. The complex band structure is the relationship between real energy and complex states[7] and connects the real bands through the forbidden gap in the complex wave vector plane. In the region of fundamental gap, the complex bands start from valence and conduction band edges and meet at a branch point in the complex plane. The extent of the wave function decreases exponentially with the magnitude of the imaginary part k_i of this branch point. When the hole mass is much larger than electron mass, it can be found that k_i is directly proportional to conduction band effective mass and band gap. Consequently, for HgTe in which both band gap and electron mass are very small, the surface wave functions are well extended. In the case of CdTe and ZnTe, with large gap and large effective mass, the calculated k_i is large and wave function is less extended.

We conclude by noting several features of the surface binding energies that must be incorporated into any MBE growth model. First, surface binding energies are surface

orientation dependent and are not simply proportional to the number of first-neighbor bonds being made to the underlying layer. Moreover because of the possibility of charge transfer between cation and anion surface states, one may have $E_d < E_c$ or $E_d > E_c$, where the former will occur mostly for wide-gap ionic compounds and the latter for narrow- (or zero) gap, less ionic compounds. Finally, because the surface states associated with an isolated surface atom or vacancy are extended in materials with small gaps and small effective masses, the approximation of surface binding using simple pair interactions is highly suspect. We expect that the narrow-gap semiconductors HgCdTe and HgZnTe will behave more like HgTe than CdTe

Acknowledgments: The financial support provided for this work through ONR Contract No. N00014-88-C0096 and NASA Contract No. NAS1-18226 is gratefully acknowledged.

[1] S. Krishnamurthy, M. A. Berding, A. Sher, and A.-B. Chen, Phys. Rev. Lett. **64**, 2531, (1990).
[2] S. Krishnamurthy, M. A. Berding, and A. Sher, J. Appl. Phys. (in press).
[3] A.-B. Chen, Y.-M. Lai-Hsu, and W. Chen, Phys. Rev. B **39**, 923 (1989).
[4] D. J. Chadi, Phys. Rev. Lett. **52**, 1911 (1984).
[5] S. Y. Tong, G. Xu, and W. N. Mei, Phys. Rev. Lett. **52**, 1693 (1984).
[6] M. A. Berding, S. Krishnamurthy, and M. van Schilfgaarde (unpublished).
[7] C. Herring, Phys. Rev. **52**, 365 (1937); S. Krishnamurthy and J. A. Moriarty, Phys. Rev. B **32**, 1027 (1985).

Computer Aided Innovation of New Materials
M. Doyama, T. Suzuki, J. Kihara and R. Yamamoto (Editors)
© Elsevier Science Publishers B.V. (North-Holland), 1991

SURFACE ENERGIES AND ORDER-STATE: EFFECTS ON SEMICONDUCTOR GROWTH

Srinivasan KRISHNAMURTHY, M. A. BERDING, A. SHER
SRI International, Menlo Park, California 94025, USA

A.-B. CHEN
Physics Department, Auburn University, Auburn, AL 36849, USA

The energies to remove atoms from various surfaces of semiconductors have been calculated using a new Green's function formalism. By comparing the energy to remove atoms from a nearly completed (concentrated) surface, E_c, and from a nearly empty (dilute) one, E_d, we can deduce the surface interaction energy. We find that E_c and E_d differ significantly even when there are no first-neighbor bonds. Because an important contribution comes from charge transfer, $E_c - E_d$ is positive (indicating attractive interactions) for pure covalent and narrow gap materials, and may become negative (repulsive surface interactions) for wide-gap ionic materials. For attractive interactions,below a critical temperature, T_c, the equilibrium surface atoms collect in islands and two-dimensional layer-by-layer growth can occur. Above T_c, rough three-dimensional growth is the rule. When $E_c - E_d$ is negative, there is also a critical temperature, T_c^*, below which the atoms and vacancies arrange into regular arrays. Calculated energies and temperatures are given for Si, GaAs, CdTe and HgTe on the (111), ($\bar{1}\bar{1}\bar{1}$), (110), and (100) surfaces.

1. INTRODUCTION

Properties of semiconductor surfaces during epitaxial growth are sensitively dependent on the nature of the interactions between surface atoms. When the interaction between the atoms is attractive, the surface grows with formation of islands at temperatures below a critical temperature, T_c, known as roughness transition temperature. However, when the interaction between the atoms is repulsive, the surface undergoes an order-disorder transition. At growth temperatures below the order-disorder transition temperature, T_c^*, surface atoms and vacancies arrange themselves in a superlattice pattern that minimizes the number of atom-atom pairs. Regardless of the nature of the interaction, the sites are occupied uniformly at growth temperatures well above the corresponding critical temperature.

Growth by low-temperature epitaxial methods such as atomic-layer epitaxy (ALE) and molecular-beam epitaxy (MBE) is believed to take place under nonequilibrium growth conditions.[1,2] However, when the surface relaxation rate is much larger than the arrival rate, each surface atom will have enough time to minimize its free energy before the arrival of the next atom, so that a thermodynamic model can be expected to qualitatively describe the morphology of the growth surface. Statistical models with various degrees of approximations are frequently used to study the nature of the surface.[3-5] In this paper, we extend the studies to consider the surface arrangement of vacancies in a zinc-blende lattice. We obtain the strength of a pair interaction on various ideal semiconductor surfaces using the tight-binding Green's function method.[6-7] The pair interaction energies are then used in a thermodynamic model to obtain the minimized free energy. The qualitative nature of various surfaces is discussed.

2. PAIR INTERACTION ENERGIES

The surface sublimation energies for various surface orientations are obtained using the tight-binding Green's function (GF) method. Because the details of the calculational procedure have been published elsewhere,[6-8] this discussion will be brief.

We define surface sublimation energy (SSE) as the difference between the total energy of a surface plus a free atom and that of the surface with the

atom bonded to it. In this convention, a positive SSE means that atoms will require energy to go from surface to vapor. The calculation of SSE proceeds in two steps. First, the GF for an ideal atomically smooth surface is calculated in terms of bulk GF, using Dyson's expansion. The modified GF with an atom added to or removed from the surface is then calculated in terms of surface GF. The change in density of states introduced by perturbations is used to obtain the electronic energy contribution to SSE.

The perturbation to the surface that includes the removal of an atom, leaving a surface vacancy, is referred to as surface sublimation from a *concentrated* surface and the corresponding SSE is denoted as E_c. This removal entails breaking bonds with the layer below and in-plane and, sometimes, breaking second-neighbor surface bonds. The perturbation to the surface that involves the removal of an atom and leaves an ideal flat surface is referred to as sublimation from a *dilute* surface and the corresponding SSE is denoted as E_d. Here, the removal necessitates breaking bonds with the layer below only.

In this model, it is clear that the effective atom-atom pair interaction energy, ε_0, between atoms on the surface is simply $(E_c - E_d)/\eta_0$, where η_0 is the number of near-neighbor surface bonds. A positive (negative) ε_0 implies an attractive (repulsive) interaction between surface atoms. Calculated values of E_c and E_d are given in Table I for (111), $(\bar{1}\bar{1}\bar{1})$ and (100) surfaces of silicon, GaAs, CdTe and HgTe. Notice that the (111) surface can terminate either with triply bonded cations or with singly bonded anions. The opposite bonding arrangement is found in the $[\bar{1}\bar{1}\bar{1}]$ direction.

Intuitively, one would expect $E_c > E_d$. The removal of an atom from the concentrated limit, in addition to breaking interlayer bonds, requires the breaking of in-plane surface bonds, which are first-neighbor bonds for the (110) surface and second-neighbor bonds for the other surfaces. Even in the absence of first-neighbor bonds on the surface, the surface dangling bonds in the concentrated limit interact to form a partially filled band, which lowers energy. When a surface atom is removed, the removal of an electron from these broadened surface states will in general require more energy than if the dangling states were isolated, as in the dilute case. As seen from Table 1, this conventional interpretation explains the trends observed in silicon, HgTe, and in GaAs in most cases. However, for other cases, we find that E_c is smaller than E_d.

In order to understand the mechanisms that drive some semiconductors to have $E_c < E_d$, we examine the nature of the anion and cation surface states. As an illustration, we choose the $(\bar{1}\bar{1}\bar{1})$ surface of CdTe where E_c is less than E_d. The surface density of states for two ideal surfaces, the Cd- and Te-terminated $(\bar{1}\bar{1}\bar{1})$, are shown in Figure 1. For an unreconstructed surface, some of the

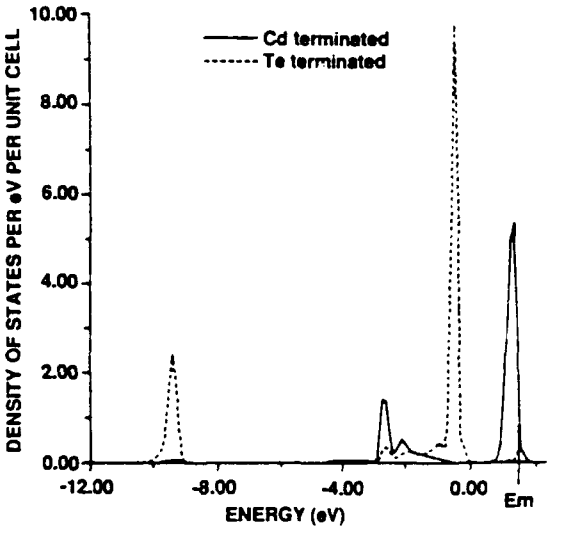


FIGURE 1
Density of dangling hybrid states of Te (dashed) and Cd (solid) for $(\bar{1}\bar{1}\bar{1})$ oriented CdTe. E_m is the highest occupied level for Cd terminated $(\bar{1}\bar{1}\bar{1})$ surface.

surface states normally lie in the fundamental gap. Compounds have both cation- and anion-derived hybrid surface states, but the cation-derived states most often lie higher in energy. The energy

exception. Owing to an increased ionicity in II-VI compounds, the energy separation between the dangling states is large. As illustrated above, the effect of the charge transfer is substantial in CdTe and in most cases causes E_c to be less than E_d. Although HgTe is a II-VI compound, the charge transfer does not occur because it is a semimetal with no forbidden gap.[6,7]

From Table 1, we conclude that Si, HgTe, and most GaAs surfaces will exhibit smooth to rough transitions while most of the CdTe surfaces and the singly bonded, gallium-terminated (111) surface will undergo an order-disorder transition. During growth, the actual nature of a surface will be determined by the growth temperature with respect to the critical temperature.

3. RESULTS AND DISCUSSION

In ALE, grown from a free atom vapor phase, only one type of atom species (anion or cation) is present at any given time. The atom-atom pair energies calculated in Section 2 can be used in a single-layer growth model to obtain minimized free energy as a function of layer coverage, x. At temperatures below T_c, the growth will take place with the formation of islands, and atomic concentration in those islands will be given by the location of multiple minima in the curve of x versus free energy. At temperatures well below T_c, the minima lie near $x = 0$ and $x = 1$. In these cases, islands on the surface are nearly fully occupied and the remaining area is nearly empty. As the surface coverage is increased, the islands grow in size with a fixed vacancy concentration until the layer is fully grown, i.e., two-dimensional, layer-by-layer growth.

Because the (100) surface is a square lattice, the exact value of $T_c = 0.57\ \varepsilon_0/k_B$ given by Onsagar[3] is used, where k_B is Boltzmann's constant. For the hexagonal lattice, observed in the (111) orientation, the exact value[11] of T_c is $0.91\ \varepsilon_0 / k_B$. The calculated T_c are given in Table 2 for Si, HgTe, and most GaAs surfaces. T_c is much larger than the typical growth temperature for these elements and compounds [except for the $(\bar{1}\bar{1}\bar{1})$ Te-terminated HgTe surface]. Thus, ALE growth of these materials is dominated by the formation of nearly perfect islands.

When the interaction between surface atoms is repulsive, as in a Ga-terminated $(\bar{1}\bar{1}\bar{1})$ GaAs surface and all CdTe surfaces, the free energy of a homogeneous surface always remains negative with one minimum. However, these surface energies may reduce the free energy with ordered arrangement of atoms and vacancies. The nature of the superlattice arrangements below T_c^* and the area of different superlattices depend on the growth temperature and surface coverage. For example, at temperatures well below T_c^* for a square lattice of 50% coverage, the superlattice pattern will look like a checkerboard, where every near-neighbor site of an atom (vacancy) is occupied by vacancies (atoms). This type of arrangement can be envisioned for a monolayer growth on (100) orientations. T_c^* calculated using Onsagar's result for (100) surfaces and that using an exact value[12] of $0.35\ \varepsilon_0/k_B$ for (111) orientations are given in Table 2. We see that the calculated temperatures are well above the typical ALE or MBE growth temperatures for these compounds [except for the $(\bar{1}\bar{1}\bar{1})$ Ga-terminated GaAs surface]. Hence, the

TABLE 2. Transition temperatures in Kelvins. Parentheses show number of bonds broken with layer below.

Orientation	Removed atoms	Si T_c K	GaAs T_c K	CdTe T_c K	HgTe T_c K
(111)	a(1)	3700	2600	500	3500
	c(3)	3900	200	*1400	2400
$(\bar{1}\bar{1}\bar{1})$	a(3)	3900	700	*2100	0
	c(1)	3700	410	*600	900
(100)	a(2)	3200	2200	*4000	1200
	c(2)	3200	1700	*3300	2200

*Order-disorder transition temperature

partly filled growth surface is expected to exhibit superlattice arrangements with atoms and vacancies.

In MBE where both species are present, simultaneous growth of an anion-cation double layer can be expected. We extend our model to study the double-layer growth of semiconductors with zinc-blende lattice. The intralayer pair interaction energies deduced for the single-layer growth model are still valid. The interlayer anion-cation pair interaction energies are simply E_d/η_1, where η_1 is the number of interlayer near-neighbor sites. We define an anion-stabilized growth to be the case where relative anion and cation fluxes are such that any exposed surface atoms are anions. When the interaction between the atoms in the upper layer and that in the lower layer is strong (compared with kT_g), the surfaces always grow in the smooth limit. (100) surfaces of HgTe and GaAs and all surfaces of Si are predicted to fall into this category. However, when the interaction between adjacent layers is not always strongly attractive, as in the case of $(\bar{1}\bar{1}\bar{1})$ HgTe and (111) GaAs surfaces, the nature of growth depends on whether the surface is anion or cation stabilized.

The free energy calculated in quasi-chemical approximation (QCA)[13] at a growth temperature of 185°C for Te- and Hg-stabilized $(\bar{1}\bar{1}\bar{1})$ surfaces is plotted respectively in Figure 2(a) and (b) as a function of Te and Hg surface concentration. Each corner of this plot represents the surface fully covered by the element noted there. In ALE, growth takes place along the sides. Whenever the minimum occurs between the two corners, that surface is predicted to grow in the rough limit where atoms and vacancies randomly occupy the surface sites. Whenever the minima occur near the corners, the surface is predicted to grow in the smooth limit with formation of nearly perfect islands. In MBE where both species are present in double-layer stochastic growth of compounds, the above arguments are applied to the free-energy curve along the diagonal line connecting the origin and upper-right corner. Accordingly, we see from Figure 2 that, while ALE growth of the Hg and Te

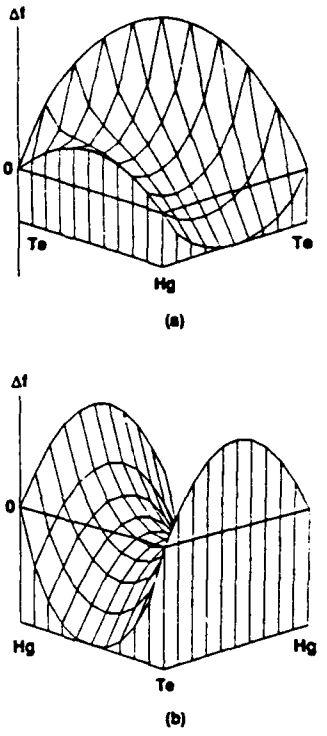


FIGURE 2

(a) Free energy for Te-stabilized surface $(\bar{1}\bar{1}\bar{1})$ surface of HgTe at 185°C temperature as a function of Hg and Te surface concentrations.

(b) Free energy for Hg-stabilized surface $(\bar{1}\bar{1}\bar{1})$ surface of HgTe at 185°C temperature as a function of Hg and Te surface concentrations.

layer takes place in the smooth and rough limits, respectively, MBE growth of the Te-stabilized surface takes place in the smooth limit. This is because the pair energy for Te (upper layer)-Hg (lower layer) is strong, 0.9 eV. However the Hg-stabilized surface grows in the rough limit because the Hg in the upper layer is weakly attracted to Te below (nearly 0 eV) and atoms do not congregate into islands to maximize the cation-anion pairs.

4. CONCLUSIONS

We have calculated intralayer and interlayer atom-atom pair interaction energies using a tight-binding Green's function method. The calculated values imply that sublimation energies do not increase linearly with number of broken bonds. Also, the sublimation energy depends on the coverage. The possibility of repulsive surface interactions due to a charge transfer mechanism is predicted. These energies are then used in a thermodynamic model to study the nature of growing surfaces. As the growth temperature is varied, attractive interaction between surface atoms leads to smooth-to-rough transition, and Si, HgTe, and most GaAs surfaces are calculated to fall in this category. The repulsive interactions lead to superlattice ordered-to-disordered transition, and $(\bar{1}\bar{1}\bar{1})$ GaAs and most CdTe surfaces are expected to be in this category. Smooth or rough surface growth is found to be decided by in-plane interactions in an ALE growth and is also influenced by interplane interactions in double-layer MBE growth. In our calculations of pair energies, we considered only ideal surfaces. The critical temperatures and other conclusions in this paper, can, at best, be expected to represent the trends. Nonidealities such as reconstruction, dimerization, relaxation of adatoms and of surface can lower surface energy substantially.[14,15] For quantitatively more accurate values, the effects of these mechanisms on the values of $(E_c - E_d)$ and statistical approximations that are better than QCA with pairs will have to be included.

ACKNOWLEDGMENT

The work was supported by NASA Contract NAS1-18226, by ONR Contract N00014-88-C0096, and by AFOSR Contract F49620-88-K-0009.

REFERENCES

1. A. Madhukar and S. Ghaisas, CRC Crit. Rev. Sol. State and Mat. Sci. **14** (1988) 1.

2. J. Greene, CRC Crit. Rev. Sol. State and Mat. Sci. **11**, (1983) 47.

3. L. Onsagar, Phys. Rev. **65** (1944) 117.

4. K.A. Jackson, *Liquid Metals and solidification* (American Society for Metals, Cleveland, Ohio, 1958) 174 .

5. R. Kikuchi, J. Chem. Phys.**57** (1972) 4633.

6. S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, Phys. Rev. Lett. **64** (1990) 2531.

7. S. Krishnamurthy, M.A. Berding, A. Sher, and A.-B. Chen, J. Appl. Phys. (in print).

8. A.-B. Chen, Y.-M. Lai-Hsu, and W. Chen, Phys. Rev. **B39** (1989) 923.

9. D.J. Chadi, Phys. Rev. Lett. **52** (1984) 1911.

10. S.Y. Yang, G. Xu, and W.N. Mei, Phys. Rev. Lett. **52** (1984) 1693.

11. R.J. Baxter, *Exactly solved models in statistical mechanics* (Academic Press, New York, 1982) 308.

12. M. Schick, J.S. Walker, and M. Wortis, Phys. Rev. **B16** (1977) 2205.

13. A. Sher, M. van Schilfgaarde, A.-B. Chen and W. Chen, Phys. Rev. **B36** (1987) 4279.

14. K. C. Pandey, Phys. Rev. Lett. **49** (1982) 223.

15. J.E. Northrup, Phys. Rev. Lett. **57** (1986) 154.

InTlSb AS AN LWIR MATERIAL

July 1992

A. Sher, M. van Schilfgaarde, M.A. Berding
SRI International, Menlo Park, CA 94025

J.M. MacCallum
DDDRE/R&AT/ET The Pentagon, Washington, DC 20301

A.-B. Chen
Auburn University, Auburn, NC 36849

ABSTRACT

We have examined the prospects of $In_{1-x}Tl_xSb$ as a long-wave infrared (LWIR) material for focal-plane array applications. We find there is a near bond-length match between InSb and TlSb in the zinc blende phase, the cohesive energy is typical of InSb, and the concentration x needed to narrow the gap into the LWIR range is about $10 \pm 5\%$. The calculated phase diagram shows a wide temperature range where this alloy is stable. While this diagram shows that the alloy cannot be grown from the liquid phase, growth should be possible in a vapor-phase process.

We have investigated the prospects of $In_{1-x}Tl_xSb$ as an IR detector material with a gap that can be adjusted to zero just as that of $Hg_{1-x}Cd_xTe$. In the search for a III-V compound-based alternative to $Hg_{1-x}Cd_xTe$ for long-wave infrared (LWIR) focal-plane arrays,[1,2] a number of materials are currently under investigation, including $InAs_{1-x}Sb_x$ based layer superlattices[1] and $Al_{1-x}Ga_xAs$ based quantum well structures.[2] InSb is often used in mid-wave infrared (MWIR) devices, but its bandgap is too wide to respond in the LWIR. The alloy $InSb_{1-x}Bi_x$ has been examined[3,4] as one with an adjustable narrower gap that might serve the purpose. However, BiSb does not form in the zinc blende (ZB) structure[5] and the

---

1 G.C. Osbourn, L.R. Dawson, R.M. Biefeld, T.E. Zipperian, I.J. Fritz, and B.L. Doyle, *J. Vac. Sci. Technol.* **A5**, 3150 (1987).

2 B.J. Levine, C.G. Bethea, K.G. Glogovsky, J.W. Stayt, and R.E. Leibenguth, *Semicond. Sci. Technol.* **6**, C11 (1991).

3 J. Zelko and J.E. Greene, *Appl. Phys. Lett.* 33, 254 (1978); *J. Appl. Phys.* **51**, 1549 (1980).

4 M.A. Berding, A. Sher, A.-B. Chen, and W.E. Miller, *J. Appl. Phys.* **63**, 107 (1988).

5 O. Madelung, *Physics of III-V Compounds*, John Wiley and Sons, New York, Ch. 1 (1964).

alloy is not miscible in equilibrium. It has been prepared by Ag ion-assisted epitaxy, but no practical growth system has evolved from this work.

We have theoretically examined $In_{1-x}Tl_xSb$ based on a combination of full-potential linearized-muffin-tin orbital approximation (FP-LMTO) whose accuracy is limited only by the local density approximation (LDA)[6] and a parametrized calculation of the $In_{1-x}Tl_xSb$ liquid-state free energy. Details of these calculations will be presented elsewhere. Here we summarize only the results relevant to the IR community.

Table 1 contains the atomic volume Ω, cohesive energy per cation E, bulk modulus B, one shear coefficient C_{11}-C_{12}, and the bandgap Eg for the materials HgTe, CdTe, InSb, and TlSb. For InSb and TlSb the calculations are presented for both the ZB and CsCl crystal structures. Other structures were also examined, but their cohesive energies are small. The HgTe and CdTe ZB numbers were done for comparison.

Table 1

Calculated and experimental optical and structural properties for CdTe, HgTe, InSb, and TlSb in a AB lattice at 0 K. Atomic volume Ω is in Å, cohesive energy per cation E and bandgap E_g are in eV; C_{11}-C_{12} and bulk modulus B are in 10^{12} erg/cm^3. Calculated values for InSb in the CsCl and TlSb in both ZB and CsCl lattices are shown. Column *Adj* adjusts the calculated values of TlSb in the ZB structure for errors in the LDA.

	CdTe		HgTe		InSb			TlSb		
ZB	Calc	Expt	Calc	Expt	Calc	Expt	CsCl	ZB	Adj	CsCl
Ω	33.40	34.2	34.2	33.4	34.1	33.8	25.6	36.4		27.1
E	5.28	4.40	4.24	3.24	6.45	5.60	6.08	5.70		5.79
$C_{11}-C_{12}$	0.20	0.17	0.19	0.18	0.31	0.31		0.17		
B	0.44	0.42	0.44	0.47	0.45	0.48	0.72	0.38		0.49
E_g	0.54	1.60	-0.95	-0.30	-0.46	0.26		-2.20	-1.5	

A number of important points can be deduced from Table 1.

- For HgTe, CdTe, and InSb where experimental numbers are available,[7] the calculated structural properties Ω, B, and $C_{11}-C_{12}$ agree well with experiment. E values are generally slightly high, because the free-atomic energies are not done well in LDA. However, differences between solid energies are reliable. The bandgaps are always too small in LDA, but the reasons are well understood and, once again, reliable adjustments can be made.
- In agreement with experiment,[5] we find the TlSb ground state is the CsCl structure, but the energy difference with the ZB structure is small—88 meV. The energy difference for InSb favors the ZB structure by 380 meV.

---

6 M. Methfessel and M. van Schilfgaarde, unpublished.

7 K.-H. Hellwege, editor-in-chief *Landholt-Bornstein: Numerical Data and Functional Relationships in Science and Technology*, Vols. 17 & 22, Springer-Verlag, Berlin, Heidelberg, New York (1982).

- For the ZB structure, the atomic volumes of InSb and TlSb are small, corresponding to a bond length difference of 2.1%. Thus, they are matched nearly as well as $Hg_{1-x}Cd_xTe$.
- The adjusted zero-temperature bandgaps of the InSb (0.26 eV) and TlSb (–1.4 eV) are such that the concentration needed for a 0.1 eV gap is x = 0.09. At 77 K, we estimate $x \leq 0.15$ will bring the gap into the LWIR range. These numbers are the most uncertain of those presented.
- The predicted cohesive energy per cation of the $In_{0.9}Tl_{0.1}Sb$ is 6.38 eV, compared to 5.05 eV for $Hg_{0.78}Cd_{0.22}Te$: it is 26% more strongly bound.
- Even assuming no bandgap bowing, dEg/dx is comparable for InTlSb and HgCdTe alloys.

Thus, the $In_{1-x}Tl_xSb$ alloy looks like a viable LWIR material if it can be prepared in the ZB structure with high enough concentration. We address that issue next.

Our calculated phase diagram of $In_{1-x}Tl_xSb$ is shown in Figure 1. The lower part of the diagram, that below the solidus curves, is based on our FP-LMTO results and is likely to be accurate. The solidus and liquidus curves are based on some liquidus properties that have been projected from the measured properties of InSb, and a theoretical estimate of the melting point of TlSb in the CsCl structure. The melting point of InSb is 808 K,[7] the calculated eutectic temperature and composition are 687 K and 0.68, the point at which the ZB miscibility curve intersects its solidus is at 700 K and 0.16, the point at which the CsCl strikes its solidus is 700 K and 0.99, and the melting point of the TlSb in the CsCl structure is 747 K. It is evident from this phase diagram that growing the proper alloy from the melt is impossible, which is in agreement with reported experiments.[8,9]

There is, however, a substantial temperature range where the proper alloy composition is stable and it should be possible to grow the material either by molecular beam epitaxy (MBE) or metal organic chemical vapor deposition (MOCVD) techniques. One attempt to grow the material by MBE failed,[10] but that was done without the benefit of the information in Figure 1.

Finally, we conclude that the $In_{1-x}Tl_xSb$ alloy should be stable in the concentration range needed for LWIR focal-plane arrays, and because the proper composition is mostly InSb, it should process in much the same way. However, this remains to be tested once the material is prepared.

Acknowledgments: We wish to thank A.T. Paxton, G. Cinader, and A. Sanjurjo for helpful suggestions and information. We also are indebted to ONR through Contract N00014-88-C-0096 for partial support of this effort.

---

8 S.B. Evgen'ev, O.V. Sorokina, and V.G. Zinov'ev, *Inorganic Materials* **21**, 1747 (1985); O.V. Sorohina, S.B. Sorokina, and I.V. Maschenko, *Russ. J. of Inorganic Chem.* **28**, 916 (1983).

9 V.I. Fedorov and V.I. Machnev, *Teplofizika Vysokikh, Temperature* **8**, 447 (1970).

10 C.E.C. Wood, A. Noreika, and M. Francombe, *J. Appl. Phys.* **59**, 3610 (1986).

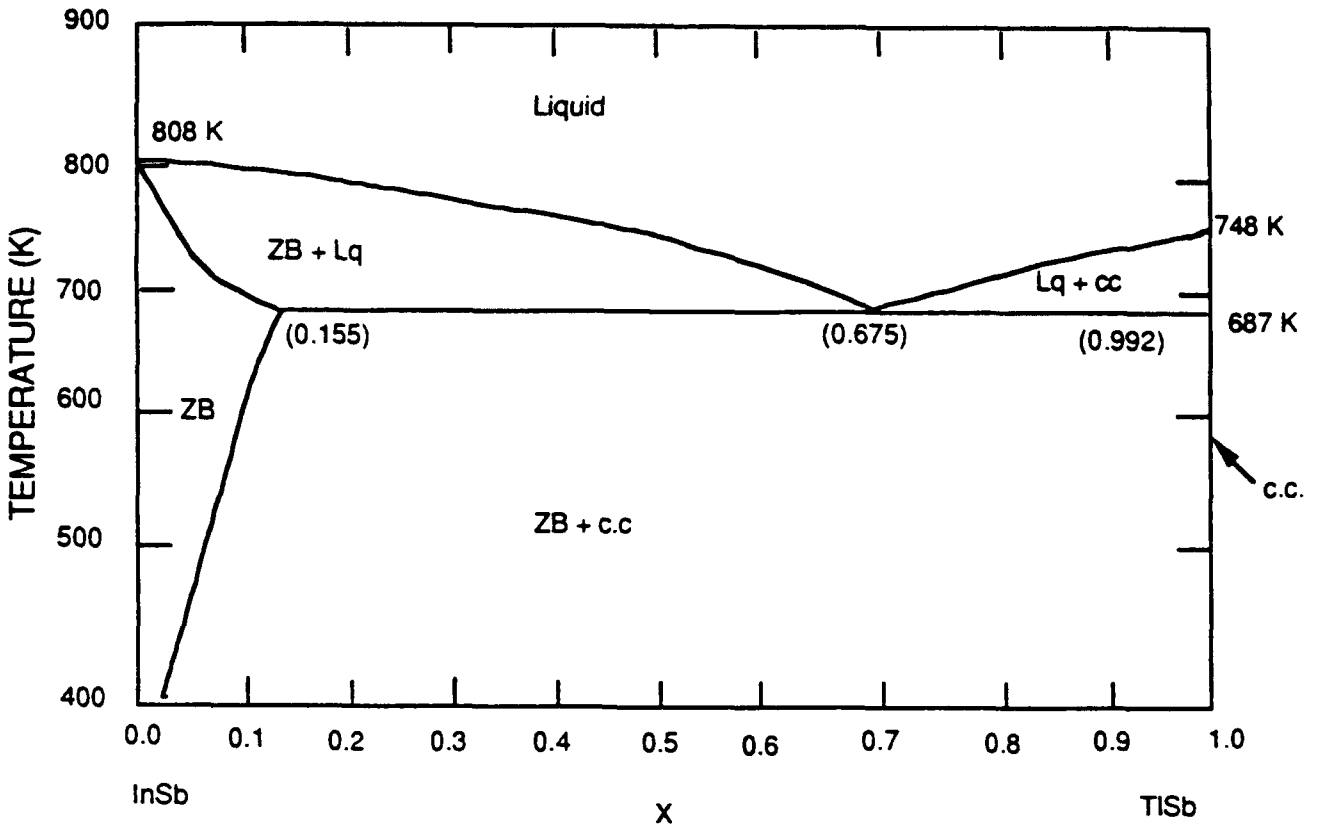


Figure 1. Phase diagram of $In_{1-x}Tl_xSb$. For $x << 1$ the zinc blende structure is stable, and for $x = 1$ the CsCl (cc) structure is the more stable one. The eutectic point is $x = 0.675$ and $T = 687$ K.

CHAPTER 1

Elastic Constants and Related Properties of Semiconductor Compounds and Their Alloys

A.-B. Chen

PHYSICS DEPARTMENT
AUBURN UNIVERSITY
AUBURN, ALABAMA

and

Arden Sher

PHYSICAL ELECTRONICS LABORATORY
SRI INTERNATIONAL
MENLO PARK, CALIFORNIA

and

W.T. Yost

NASA LANGLEY RESEARCH CENTER
HAMPTON, VIRGINIA

ISBN 0-12-752137-2

I. Introduction

1. Goals

The goals of this paper are to review the current state of knowledge of the elastic constants of elemental, compound, and pseudobinary alloy semiconductors. To accomplish this objective, we will review

- experimental methods currently used to measure elastic constants,
- experimental results,
- binding and elastic constant theory, and
- related mechanical properties.

Heavy emphasis is placed on comparisons between theory and experiment, and the accuracy of approximations currently in vogue. The theories discussed range from first-principles methods, requiring heavy comutations, to parametrized physical models. The intent is to identify a logical path between these extremes, and thereby provide insight into the connection between atom potentials and semiconductor mechanical properties. In the course of this presentation, there are a number of instances where improvements in the theory or additional experimental results, are needed. We have tried to highlight these situations and suggest possible remedies.

2. Definition and Calculation of Elastic Constants

The theory of elasticity of solids has been well formulated in many treatises (Love, 1944; Landau and Lifshitz, 1959), so we need not review the different formalisms and conventions. However, as we do wish to present a coherent account of the essence of the theory, it is necessary to define terms and describe general calculations to be used later.

In linear elasticity theory, deformation is assumed to be infinitesimal. The relative displacement vector $\vec{x}'$ between two points in a deformed solid is related to the corresponding vector $\vec{x}$ in the undeformed solid by the following equation, in component form, through a nine-component strain sensor ε:

$$x'_\alpha = x_\alpha + \Sigma_\beta \varepsilon_{\alpha\beta} x_\beta . \tag{1}$$

The change in the internal energy associated with ε is also small and will be denoted UV, where V is the equilibrium volume of a solid. Under the

condition that the entropy and the electrostatic displacement field are constant, U is only a function of ε, and is quadratic in ε:

$$U = \tfrac{1}{2}\Sigma_{\alpha\beta\mu\nu}\varepsilon_{\alpha\beta}C_{\alpha\beta\mu\nu}\varepsilon_{\mu\nu}, \tag{2}$$

where $C_{\alpha\beta\mu\nu}$ are the elastic stiffness coefficients, which are characteristic properties of the solid. From this definition, $C_{\mu\nu\alpha\beta} = C_{\alpha\beta\mu\nu}$ is required for U to be an analytical function of ε. Moreover, $C_{\alpha\beta\mu\nu} = C_{\beta\alpha\mu\nu} = C_{\alpha\beta\nu\mu}$ is also required to ensure that U is zero under any infinitesimal rigid rotation. In light of these properties, the energy density can be expressed in terms of a symmetrical strain tensor η as

$$U = \tfrac{1}{2}\Sigma_{\alpha\beta\mu\nu}\eta_{\alpha\beta}C_{\alpha\beta\mu\nu}\eta_{\mu\nu}, \tag{3}$$

where $\eta_{\alpha\beta}$ is defined as

$$\eta_{\alpha\beta} = \tfrac{1}{2}(\varepsilon_{\alpha\beta} + \varepsilon_{\beta\alpha}). \tag{4}$$

The strain tensor η is a thermodynamic parameter with stress tensor σ as its conjugate variable (Brugger, 1964). The components of σ are given by

$$\sigma_{\alpha\beta} = \frac{\partial U}{\partial \eta_{\alpha\beta}} = \Sigma_{\mu\nu}C_{\alpha\beta\mu\nu}\eta_{\mu\nu}. \tag{5}$$

It is clear that σ is also a symmetrical tensor.

The most frequently used notation is the engineering convention, in which the strain tensor e is related to η of Eq. (4) by $e_{\alpha\alpha} = \eta_{\alpha\alpha}$ for the diagonal components, but $e_{\alpha\beta} = 2\eta_{\alpha\beta}$ for $\alpha \neq \beta$. Furthermore, since e has at most six independent components, it is treated as a six-component vector, with the vector components 1 to 6 corresponding respectively to the tensor components xx, yy, zz, yz, xz and xy. In terms of e the strain energy density is written as

$$U = \tfrac{1}{2}\Sigma_{ij}\, C_{ij}e_ie_j, \tag{6}$$

where C_{ij} can be identified as $C_{\alpha\beta\mu\nu}$ with $i = \alpha\beta$ and $j = \mu\nu$. Because C_{ij} is symmetrical, it has at most 21 independent components for any crystal. Crystal symmetries reduce this number further (Love, 1944; Ashcroft and Mermin, 1976). For a cubic lattice, to which class the zincblende and diamond semiconductors belong, there are only three independent components, namely C_{11}, C_{12}, and C_{44}.

The three independent elastic constants of the diamond and zincblende (zb) semiconductors can be calculated by considering the following three strains.

(i) Under a uniform expansion, which changes a displacement $\vec{x}$ into $\vec{x}' = (1 + e)\vec{x}$, then $e_1 = e_2 = e_3 = e$ and the other strain components are zero; the elastic energy density is then given by

$$U = 3(C_{11} + 2C_{12})e^2/2. \tag{7}$$

U can also be expressed in terms of the adiabatic bulk modulus B defined by $\delta P = -B(\delta V/V)$, where V is the crystal volume, δV is its change, and δP is the corresponding pressure change. The result is $U = B(\delta V/V)^2/2 = 9Be^2/2$, because the dilatation is $\delta V/V = 3e$ in the present case. This establishes the relationship

$$B = (C_{11} + 2C_{12})/3. \tag{8}$$

(ii) The next case to consider is a tetragonal shear strain e which changes a displacement according to

$$(x, y, z) \rightarrow (x + ex, y - ey, z). \tag{9}$$

The only nonzero strain components are $e_1 = -e_2 = e$. Then U becomes

$$U = (C_{11} - C_{12})e^2. \tag{10}$$

(iii) To calculate C_{44} we consider a shear strain e that changes a displacement according to

$$(x, y, z) \rightarrow (x + ey/2, y + ex/2, z). \tag{11}$$

This strain contains e_6 as the only nonzero component. Because C_{44} equals C_{66}, the energy density is simply

$$U = C_{44}\, e^2/2. \tag{12}$$

Although the macroscopic crystal distortion of the Bravais lattice caused by this strain is described by Eq. (11), microscopically there is a relative displacement $\tilde{u} = (0, 0, u)$, the so-called Kleiman (1962) internal displacement, between two successive atomic planes perpendicular to the z axis. In other words, the relative displacements between the atoms on the same fcc sublattice are governed by Eq. (11), but there is an additional induced relative

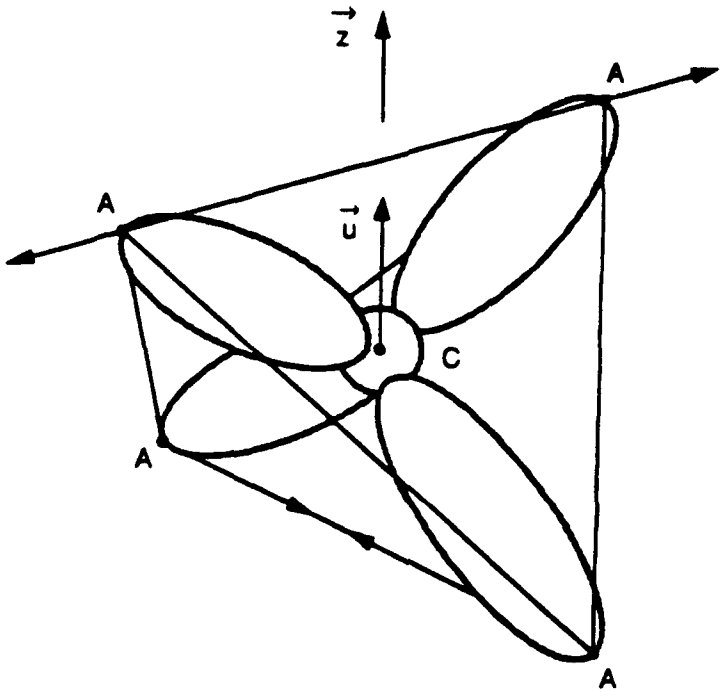


FIG. 1. Distortion of a tetrahedron corresponding to the C_{44} elastic constant; u is the internal dispacement between the anion and cation sublattices.

displacement $\tilde{u}$ between the two sublattices. The directions of the displacements of atoms in a tetrahedral cell are shown in Fig. 1. In calculations one can use an arbitrary infinitesimal pair of e and u to obtain the coefficients in the following quadratic expansion of the strain energy density:

$$U = \Phi u^2/2 + Deu + C_{44}^{(0)} e^2/2, \tag{13}$$

where the force function Φ is related to the transverse optical (TO) phonon frequency ω at Γ (the center of Brillouin zone) by $\Phi = \mu\omega^2$, with μ being the reduced mass. In Eq. (13), D is a constant, and $C_{44}^{(0)}$ would be the shear stiffness coefficient if the internal displacement were not allowed. Kleiman (1962) defined an internal displacement parameter that is related to the equilibrium value of u by $u = \zeta ae/4$ for a fixed e, where a is the lattice constant. Taking the first derivative of U with respect to u in Eq. (13) and setting it equal to zero, one finds that the ζ value is given by

$$\zeta = -4D/(a\Phi). \tag{14}$$

Finally, the sought-after C_{44} is given by

$$C_{44} = C_{44}^{(0)} - \zeta^2 a^2 \Phi/16. \tag{15}$$

Thus, the internal displacement is an essential part of C_{44}. These procedures will be used in the theoretical calculations to be presented in subsequent sections.

3. Elastic Constants and Sonic Wave Propagation

The purpose of this section is to give the reader sufficient information to measure ultrasonic wave velocities and to deduce from them the elastic coefficients of materials of interest to the electronics industry. We start from the definition of strain and develop equations that relate the elastic coefficients with wave propagation velocities in various crystallographic directions. The cases taken as examples, cubic and isotopic, were chosen because of their prevalent role in electronic materials.

A number of techniques used to measure wave propagation velocities in materials are discussed. The researcher's decision on technique depends upon a number of factors, such as accuracy needed, size and crystallographic orientation of available samples, equipment, etc. Sample preparation is significant, since a carelessly prepared sample yields useless velocity data. Because the equipment needs vary widely among the techniques that can be used, references containing information on each technique are given. Key papers chosen for citation were picked for their readability and their direct application to the measurement technique. Primary emphasis has been given to techniques that are generally accepted and that give results that are understood. Throughout this section a general theory of elasticity due to Murnaghan (1951) is used.

a. Elastic-to-Wave Relationship

Consider, as before, a point in a lossless elastic medium at rest, whose displacement vector, $\vec{x} = (x_\alpha)$, $\alpha = 1, 2, 3$, are the cartesian coordinates of a point in the medium. In the presence of a stress wave disturbance, the point moves to a new location, $\vec{x}' = (x'_\alpha)$ at time t. The Lagrangian strains, η, defined consistently with Eqs. (1) and (4) but with higher-order terms retained (Landau and Lifshitz, 1986; Murnaghan, 1951), are as follows:

$$\eta_{\eta\nu} = \frac{1}{2}\left(\frac{\partial x'_\alpha}{\partial x_\mu}\frac{\partial x'_\alpha}{\partial x_\nu} - \delta_{\mu\nu}\right) \tag{16}$$

where $\delta_{\mu\nu}$ is the Kronecker delta. Einstein summation convention over repeated indices is used in this section.

By considering the internal energy per unit mass of the medium $E(x_\alpha, \eta_{\alpha\beta}, S)$, where S is the entropy, we obtain a relationship between wave propagation and the internal energy of the medium. If we restrict this discussion to a medium that is initially unstressed, then the internal energy per unit volume of the medium is $\rho_0 E(r_\alpha, \eta_{\alpha\beta}, S)$, where ρ_0 is the mass density of the medium in the unstrained state. By expanding this in a power series of the Lagrangian strains, we obtain

$$\rho_0 E(x_\alpha, \eta_{\alpha\beta}, S) = \rho_0 E(x_\alpha, 0, S) + \frac{1}{2!} C_{\mu\nu\alpha\beta}\eta_{\mu\nu}\eta_{\alpha\beta} + \frac{1}{3!} C_{\mu\nu\alpha\beta\gamma\delta}\eta_{\mu\nu}\eta_{\alpha\beta}\eta_{\gamma\delta} + \cdots . \tag{17}$$

The coefficients $C_{\mu\nu\alpha\beta}$ and $C_{\mu\nu\alpha\beta\gamma\delta}$ are the second-order and third-order elastic coefficients (adiabatic) as defined by Brugger (1964):

$$C_{\mu\nu\alpha\beta} = \rho_0 \left(\frac{\partial^2 E}{\partial\eta_{\mu\nu}\partial\eta_{\alpha\beta}} \right), \tag{18}$$

$$C_{\mu\nu\alpha\beta\gamma\delta} = \rho_0 \left(\frac{\partial^3 E}{\partial\eta_{\mu\nu}\partial\eta_{\alpha\beta}\partial\eta_{\gamma\delta}} \right). \tag{19}$$

The wave equation can be derived (Goldstein, 1965, p. 347) from Lagrange's equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_\mu}\right) + \frac{d}{dx_\alpha}\left(\frac{\partial L}{\partial\left(\frac{\partial x'_\mu}{\partial x_\alpha}\right)}\right) - \frac{\partial L}{\partial x'_\mu} = 0, \tag{20}$$

where $\dot{x}_\mu = \partial x_\mu/\partial t$ and L is given by

$$L = \tfrac{1}{2}\rho_0 \dot{u}_\mu \dot{u}_\mu - \rho_0 E(x_\alpha, \eta_{\alpha\beta}, S). \tag{21}$$

Defining the particle displacement vector, $\mathbf{u}$, as

$$\mathbf{u}_\mu(\vec{r}, t) = x'_\mu - x_\mu, \tag{22}$$

and differentiating it gives

$$\frac{\partial u_\mu}{\partial x_\nu} = \frac{\partial x'_\mu}{\partial x_\nu} - \delta_{\mu\nu}. \tag{23}$$

Combining Eqs. (22), (23), and (16) gives the strain tensor in terms of the particle displacements,

$$\eta_{\mu\nu} = \frac{1}{2}\left(\frac{\partial u_\mu}{\partial x_\nu} + \frac{\partial u_\nu}{\partial x_\mu} + \frac{\partial u_\beta}{\partial x_\mu}\frac{\partial u_\beta}{\partial x_\nu}\right). \tag{24}$$

Using Eqs. (21) to (24) in Eq. (20), and retaining first-order terms in $\partial u_\mu/\partial x_\nu$, gives

$$\rho_0 \ddot{u}_\mu = \frac{d}{dx_\nu}\left(C_{\mu\nu\alpha\beta}\frac{\partial u_\alpha}{\partial x_\beta}\right). \tag{25}$$

We assume a plane wave solution to Eq. (25) of the form

$$u_\mu = \hat{e}_\mu \cdot \tilde{\mathbf{u}} \cos(k_\nu x_\nu - \omega t), \tag{26}$$

where

$\hat{e}_\mu$ = unit displacement vector in the μth direction,
$\tilde{\mathbf{u}} = (u_\mu)$ = particle displacement vector,
$\tilde{\mathbf{k}} = (k_\nu)$ = wave propagation vector, and
ω = angular frequency.

Substituting Eq. (26) into Eq. (25), we obtain the eigenvalue-eigenvector equation

$$|C_{\mu\nu\alpha\beta}\kappa_\nu\kappa_\beta - \rho_0 v^2 \delta_{\mu\alpha}| = 0, \tag{27}$$

where

(κ_μ) = the set of direction cosines of $\tilde{\kappa}$,
$v = \omega/|\tilde{k}|$ is the phase speed, and
$\tilde{\kappa} = \tilde{k}/|k|$ indicates the direction of wave propagation in the medium.

For any given wave propagation direction, there are three eigenvalue solutions to Eq. (27); these correspond to one quasilongitudinal and two quasitransverse polarization modes. Relationships have also been derived for wave propagation in the presence of residual stresses and for retention of the higher-order elastic coefficients (Cantrell, 1982; Breazeale and Ford, 1965).

Because both the stress and strain tensors are symmetric, one can reduce the number of independent second-order elastic coefficients to 21 (Landau and Lifshitz, 1986, p. 32). It is also convenient to use the Voight (1928)

contraction of the indices as in Eq. (6), which is used in the remainder of this section. The number of independent elastic coefficients varies with crystal structure[1] and is listed in the accompanying table.

Crystal Class	Number of Independent Elastic Coefficients
Triclinic	21
Monoclinic	13
Orthorhombic	9
Tetragonal (C_4, S_4, C_{4h})	7
Tetragonal (C_{4v}, D_{2d}, D_4, D_{4h})	6
Rhombohedral (C_3, S_6)	7
Rhombohedral (C_{3v}, D_3, D_{3d})	6
Hexagonal	5
Cubic	3
(Isotropic)	2

Green (1973) and others have solved Eq. (27) for isotropic materials and for cubic crystals in the [100], [110], and [111] directions. Using the symmetry properties of the crystals, one obtains the results presented in Section I.3.b for cubic crystals, and in Section I.3.c for isotropic solids.

b. Cubic Crystals

For plane waves propagating along the [100] direction:

$$(\kappa_1 = 1, \kappa_2 = 0, \kappa_3 = 0);$$

longitudinal (compressional) waves (pure mode with particle displacements u_1 in direction of propagation, u_2 and $u_3 = 0$) have a phase speed

$$v_1 = \sqrt{\frac{C_{11}}{\rho_0}}; \tag{28}$$

and transverse (shear) waves (pure mode with particle displacements $u_1 = 0$, u_2 and u_3 perpendicular to the direction of propagation) have values given by

$$v_2 = v_3 = \sqrt{\frac{C_{44}}{\rho_0}}. \tag{29}$$

[1]We assume that the angles defining the orientation of axes in the crystal are not specified. For a further discussion, see Landau and Lifshitz (1986).

For plane waves propagating along the [110] direction:

$$(\kappa_1 = 1/\sqrt{2}, \kappa_2 = 1/\sqrt{2}, \kappa_3 = 0);$$

for longitudinal (compressional) waves (pure mode with particle displacements $u_1 = u_2$, $u_3 = 0$), one obtains

$$v_1 = \sqrt{\frac{C_{11} + C_{12} + 2C_{44}}{2\rho_0}}; \tag{30}$$

and for transverse (shear) waves (pure mode with $u_1 = -u_2$, $u_3 = 0$), one obtains

$$v_2 = \sqrt{\frac{C_{11} - C_{12}}{2\rho_0}}; \tag{31}$$

or for transverse (shear) waves (pure mode with $u_1 = u_2 = 0$, $u_3 \neq 0$), one obtains

$$v_3 = \sqrt{\frac{C_{44}}{2\rho_0}}. \tag{32}$$

For plane waves propagating along the [111] direction:

$$(\kappa_1 = 1/\sqrt{3}, \kappa_2 = 1/\sqrt{3}, \kappa_3 = 1/\sqrt{3}); \tag{33}$$

the speed for longitudinal (compressional) waves (pure mode with $u_1 = u_2 = u_3$) is given by

$$v_1 = \sqrt{\frac{C_{11} + 2C_{12} + 4C_{44}}{3\rho_0}}; \tag{34}$$

and that for transverse (shear) waves (pure mode with $u_1\kappa_1 + u_2\kappa_2 + u_3\kappa_3 = 0$ or particle displacement perpendicular to wave propagation) is given by

$$v_2 = v_3 = \sqrt{\frac{C_{11} - C_{12} + C_{44}}{3\rho_0}} \tag{35}$$

c. Isotropic Solids

For the case of isotropy, all directions are equivalent, we are left with two independent constants (Lamé constants). The elastic coefficients can be

expressed in terms of the Lame' constants (Green, 1973) as follows:

$$\begin{aligned} C_{11} &= C_{22} = C_{33} = \lambda + 2\mu, \\ C_{12} &= C_{13} = C_{23} = \lambda, \\ C_{21} &= C_{31} = C_{32} = \lambda, \\ C_{44} &= C_{55} = C_{66} = \mu. \end{aligned} \tag{36}$$

For plane waves propagating in any direction (e.g., the x-direction):

$$(\kappa_1 = 1, \kappa_2 = 0, \kappa_3 = 0).$$

the speed for longitudinal (compressional) wave (pure mode with particle displacements u_1 in direction of propagation, u_2 and $u_3 = 0$) is given by

$$v_1 = \sqrt{\frac{\lambda + 2\mu}{\rho_0}}, \tag{37}$$

and that for transverse (shear) waves (pure mode with particle displacements $u_1 = 0$, and u_2 and u_3 perpendicular to the direction of propagation) is given by

$$v_2 = v_3 = \sqrt{\frac{\mu}{\rho_0}}. \tag{38}$$

II. Measurement Methods

4. Velocity Measurements

To measure the second-order elastic coefficients (SOEC), one can determine the sound velocity and the density of the sample, and calculate the combination of SOEC that the direction of propagation requires. For cubic systems, sound velocities measured in the pure mode directions equivalent to [100], [110], and [111] make it possible to determine the three independent elastic constants, C_{11}, C_{12}, and C_{44}. Using care in sample preparation, and appropriate corrections for bond thickness and diffraction, the sound velocities can be determined to parts in 10^5. The accuracy of a velocity determination is usually limited by the accuracy of the path length measurement in pulsed and continuous wave techniques. With optical techniques, deflection angles, wavelength measurements, and frequencies are determinant factors.

In general, there is little problem in obtaining a sound velocity measurement on the order of 1% uncertainty. But to improve on this, one must exercise additional care in the preparation of surfaces, the control of

temperature, the determination of acoustic path length, and the determination of travel time of the acoustic wave.[2] For single transducer configurations, where a transducer is used to send and receive the acoustic wave, the base equation for determining the velocity using pulsed techniques or pulse echo without corrections is as follows:

$$v = \frac{2L}{T}. \tag{39}$$

where v is the wave velocity, L is the sample length, and T is the round-trip time. Corrections[3] are made in the evaluation of T and vary with the technique used in the measurement. Corrections for bond thickness (McSkimin, 1961), and for diffraction effects (Papadakis, 1967) can be measured and/or calculated. Accuracies of some of the more frequently used pulsed techniques are given in Table I. A comparison of the accuracies of various techniques are discussed by Papadakis (1972, 1976). Using continuous wave techniques, the propagating plane wave model (Bolef and Miller, 1971), the basic equation without corrections is

$$v = 2L\Delta f, \tag{40}$$

where Δf is the frequency difference between two adjacent mechanical resonance modes. Typically, velocity measurement accuracy is good to about 1% to 10%. A correction factor (Chern *et al.*, 1981) can be applied to the right side of Eq. (39) that corrects for the effects of the transducer and bond on the mechanical resonances of the sample. When applied, the inaccuracy can be as low as five parts per 100,000, neglecting inaccuracies in path length measurements.

If the sample is transparent, optical techniques offer a convenient method to measure sound velocity. For some of these techniques, the accuracy is on order of parts per thousand. Still, some offer accuracies of parts per ten thousand and can be used with small samples, and very high-frequency ultrasonic waves (Breazeale *et al.*, 1981). In general, the calculation for

[2]The accuracy depends upon the correct choice for the resonant frequency of the transducer, which can be obtained from the fact that it is half a wavelength in thickness. Using a micrometer, and the wave velocity of the transducer material, one can calculate the resonant frequency of the transducer. Transducer off-resonant conditions can have a relatively large influence on the measurement of round-trip time for pulsed studies.

[3]Other influences on the measurement of round-trip time are bond thickness and diffraction. These are discussed in various papers, including McSkimin (1961) and Papadakis (1967, 1972).

TABLE I

ACCURACIES OF MEASUREMENT TECHNIQUES

Method	Features	Accuracy	Corrections
Pulsed			
Pulse echo	· cm or longer samples · one transducer · commercial units available	a few parts in 1,000	none
Gated double pulse superposition[a]	· cm or longer samples · matched transducers	1 part in 10,000	bond
Pulse superposition[b,c]	· cm or longer samples · one transducer	1 part in 5,000[c]	bond, diffraction estimate (b)
Pulse echo overlap[d,e]	· cm or longer samples · one transducer · commercial units available	(delay time accuracy several parts per 1,000,000)	bond, diffraction
Long pulse buffer rod[f]	· thin lossy samples (as thin as 2 mm)	1 or 2 parts in 10,000	bond
Continuous Wave	· bonds must be thin · cm or longer length · sample preparation more critical		requires corrections for all results. Otherwise, errors can be very large.[e] Difficult to correct for diffraction. All these techniques require compound resonator corrections.
Continuous wave transmission[g,h,i]	· can be automated · electromagnetic crosstalk can give problems	parts in 10,000,000 under ideal conditions; considerably less with thick bonds and/or thin samples[k]	
Sampled continuous wave method[j,k]	· single transducer · eliminates electromagnetic crosstalk · can be automated		
Spectrometer methods[l,m]	· can track small changes · can require specialized transducers		
Optical Methods (for Use with Optically Transparent Media)			
Measurement of wavelength	· uses traveling microscope to measure wavelength	0.1 to 0.01%	

TABLE I (*Continued*)

ACCURACIES OF MEASUREMENT TECHNIQUES

Method	Features	Accuracy	Corrections
Raman-Nath diffraction	· measurement of angles of incidence and diffraction to measure wavelength of sound	1%	
Bragg diffraction	· useful at frequencies between ≈ 100 MHz and several GHz · useful for measurement of local variations in velocity	0.01 to 0.1%	
Brillouin scattering	· measurement of optical frequency shift and angle · *useful near phase transitions*	1 to 3% (0.1% precision)	
Stimulated Brillouin scattering[a,e]	· frequencies in the GHz range · needs high power electromagnetic source · measurement can result in fracture of sample · frequencies to 60 GHz reported	3%	

[a]Williams and Lamb (1985).
[b]McSkimin (1961).
[c]Papadakis (1972).
[d]Papadakis (1967).
[e]Papadakis (1972).
[f]McSkimin (1950).
[g]Bolef *et al.* (1962).
[h]Bolef and Miller (1971).
[i]Chern *et al.* (1981).
[j]Miller (1973).
[k]Chern *et al.* (1981).
[l]Conradi *et al.* (1974).
[m]Heyman (1976).
[n]Chiao *et al.* (1964).
[o]Brewer and Rieckhoff (1964).
[p]Chern *et al.* (1981).

velocity involved the determination of acoustic wavelength in the medium, by using the equation

$$v = f\lambda, \tag{41}$$

where f is the frequency of the sound wave, determined by the drive frequency of the transducer, and λ is the wavelength. Brillouin scattering also makes it possible to measure the sound velocity (Beyer and Letcher, 1969). In addition to measurement of an angle, the accuracy also depends on determination of the wavelength of light in the medium; this requires a determination of the index of refraction of the medium.

5. Ultrasonic Measurement Techniques

Surveys of techniques and details for measuring ultrasonic velocities are available from various sources (Breazeale *et al.*, 1981; Papadakis, 1976; Truell *et al.*, 1969). Consideration in this work will be given to three classes of techniques for absolute velocity measurements. A compilation of some of the features of each technique is given in Table I.

a. Sample Preparation

The tolerance selected for preparing the sample depends largely upon the accuracy needed in the determination. If a 1% measurement is in order, surface preparation and parallelism are less critical than if an accuracy of 0.01% is desired. Also important is the consideration of correction factors needed to compensate for bond thickness and diffraction effects. At present, diffraction corrections exist only for compressional waves. Therefore, under equivalent experimental conditions, the most accurate determination of combinations of elastic coefficients would be those that are calculated from compressional wave velocities.

It is assumed that a surface on the sample has been ground to optical tolerances, and that the crystallographic direction of the axis of this surface has been determined. Typically, the crystallographic directions are measured to within minutes of arc (McSkimin and Andreatch, 1964). The remaining critical issues are the parallelism of the reflecting surface to the reference surface, and the flatness of these surfaces. Parallelism of the surfaces can be measured, for example, with an autocolliminator or a He-Ne gas laser and some mirrors. By placing the sample's reference surface on a stationary flat, dust-free surface, rotating the sample, and measuring the diameter of the circle traced by the beam reflected from the sample's top surface, and

measuring the total path length of the beam from the sample to the image screen, one can determine the parallelism of the surfaces. The surfaces should be parallel to an angle better than 0.01 $\lambda_{acoustic}$/transducer diameter. For a typical half-inch diameter, 10 MHz transducer and a typical solid, the surfaces must be parallel to better than 4×10^{-4} radians (1 degree, 23 minutes). Higher frequencies, multiple reflections, and continuous-wave techniques require proportionally smaller tolerances.

Typically, one also tries for a flatness of better than 1/100 of an acoustic wavelength for accurate determination of transit times. If the resonant frequency of the transducer is 10 MHz, then the wavelength in a solid is in the neighborhood of 5×10^{-4} meters, which puts the flatness requirement near 5 micrometers. Using an optical flat and Newton's rings analysis, one can determine the flatness of the sample. For higher-frequency transducers and continuous-wave techniques, the number of fringes are appropriately decreased. At 100 MHz, pulsed mode, for example, the tolerance is down to less than several fringes. By using flat lapping surfaces, this is easily achieved. When multiple reflections are involved in the measurement, one must consider that for each reflection, changes occur in the wavefront direction; the phase change across the surface of the transducer at each reflection is due to lack of parallelism between the faces. For a case where a 20 MHz transducer was used, and multiple reflections employed for the measurement, Papadakis (1967) quotes a sample surface parallelism and flatness of better than 10^{-4} inches per inch.

b. Piezoelectric Transducers

Selection of the piezoelectric transducer depends largely upon the wave mode, the electronic equipment to be used for the measurement, and the personal preference of the researcher. In general, piezoelectric crystals[4] such as quartz (high electric impedance) and lithium niobate, or poled ceramics such as PZT (lead zirconate titanate), are chosen. Their physical and electrical properties are covered elsewhere (O'Donnell *et al.*, 1981). Other piezoelectric transducers, including polymeric materials,[5] are also available. The transducer diameter should be smaller than the sample to assure that propagation modes are not affected by the location of lateral boundaries (Tu

[4]Information on piezoelectric transducers is available from Valpey-Fisher Corp., 75 South Street, Hopkinton, Massachusetts 01748, and Crystal Technology Inc., 1060 E. Meadow Circle, Palo Alto, California 94303.

[5]A film sold under the tradename KYNAR is an example of this. Information on this material can be obtained from Pennwalt Corp., Box C, King of Prussia, Pennsylvania 19406-0018. Some of the properties are covered in Bloomfield *et al.* (1978). Lead attachment to KYNAR films is covered in Scott and Bloomfield (1981).

et al., 1955). This is especially important where measurements are taken on small samples that may have their lateral sides close to each other. One also must consider the problem as it relates to wave propagation, since wave modes other than those considered here can be excited in materials of small dimensions. From studies on cylindrical specimens, the minimum sample diameter can be no less than approximately 2.5λ, where λ is the wavelength of the ultrasound in the medium. The transducer diameter is chosen after the frequency is selected. Generally, the transducer diameter should be no longer than half the diameter of the sample, to prevent any interference in the measurement from reflections off the lateral boundaries caused by diffraction effects.

Selection of a bonding material depends upon the type of wave (compressional or shear), and the temperature range through which the measurements are to be taken. At room temperature, a good choice for both types of waves in phenyl salicylate or phenyl benzoate (Papadakis, 1964), while Dow-Corning DC-200 silicone is good for longitudinal waves. At other temperatures (McSkimin, 1957), other bonding materials, such as Nonaq stopcock grease, are used for compressional and shear waves. Various resins can be used with shear wave transducers; commercially prepared bonding materials are also available.[6]

Application of the bond requires care to keep surfaces clean and free of dust. The bond should be as thin as practical, taking care not to break the transducer. Often, applying some heat helps with viscious bonding materials. When using phenyl salicylate or phenyl benzoate for bonds, certain procedures can be followed to assure uniform bonds (Papadakis, 1964a).

c. Noncontacting Transducers

There are three types of non-contacting excitation transducers: capacitive transducers (compressional wave excitation only) (Cantrell and Breazeale, 1977), electromagnetic transducers (EMAT) (Vasile and Thompson, 1977; Johnson and Mase, 1984), and optically stimulated acoustic transducers (Prosser and Green, 1985). With the exception of the electrostatic transducer (compressional with circular piston geometry), diffraction correction data have not been developed. However, bond corrections are not necessary as these methods generate the wave directly on the sample surface. Generally, the signal levels are quite small, and require high gain and specialized circuits or devices to bring the signals to usable levels.

[6]For example, suppliers of damped ultrasonic transducers can supply material suitable for high temperature and shear measurements. Two companies are Panametrics, Inc., 221 Crescent St., Waltham, Massachusetts 02254, and Harisonics, Inc., 7 Hyde St., Stamford, Connecticut 06907.

6. Corrections to Pulse Measurements: An Example

With some pulsed systems, and with care in preparing the sample and in taking the measurements, one can expect round-trip time determinations to have standard deviations in the neighborhood of 100 picoseconds. As corrections for bond thickness and diffraction are often larger than this, it is necessary to correct for these sources of systemic error where possible. As an example, consider, for a properly prepared sample, that the pulse-echo overlap technique is to be used to measure the round-trip time for a tone burst. The time between the first-received and the second-received echos will be determined. For simplicity, we will treat the case of only one round trip in the sample. The measured time T is composed of several terms: the true travel time δ, the bond thickness contribution Δ^B, and the diffraction contribution Δ^D.

a. Bond Thickness

The model for the sample-bonded transducer system has been developed by Williams and Lamb (1958) and has been used extensively by McSkimin (1961) and by Papadakis (1967). They showed that if the transducer bond sample system is treated as an acoustic transmission line, it is possible to calculate the effects of bond thickness on the measurement of transit time. McSkimin (1961) showed that by measuring the change in the transit time when the system is detuned by a known amount, it is possible to determine the bond thickness in terms of measured quantities. This permits one to correct for the transit time through the bond material. Papadakis (1967) discussed the correction in some detail, and applied it to his measurements on fused quartz and silicon.

Consider an undamped transducer bonded to one end of a sample with some bonding material of known velocity and density; a relatively long wave is reflected from the interface between the sample and the transducer bond system. The reflected wave from this interface experiences a phase shift from the impedance mismatch, which can be calculated from the model. The calculation uses the real and imaginary parts of the effective impedance of the components to obtain the contribution to the phase shift. By detuning the frequency of the applied tone burst, we can measure a corresponding change in transit time. Then, a comparison with the measured values permits the calculation of the bond thickness, so that its effect on the measurement can be determined. The approach is outlined below.

The reflection coefficient for the system described above is given by

$$\frac{E_b}{E_i} = \frac{Z_d - Z_s}{Z_d + Z_s}, \tag{42}$$

where E_b is the reflected pressure wave, E_i is the incident pressure wave, Z_s is the specific acoustic impedance (Elmore and Heald, 1969) of the sample ($Z_s = \rho c$, where ρ is the mass density, and c is the wave propagation speed), and Z_d is the effective specific acoustic impedance of the transducer bond system, given by McSkimin (1961),

$$Z_d = jZ_1 \left[\frac{\left(\frac{Z_1}{Z_2}\right)\tan k_1\Delta_1 + \tan k_2\Delta_2}{\left(\frac{Z_1}{Z_2}\right) - \tan k_2\Delta_1 \tan k_1\Delta_1} \right], \tag{43}$$

where Z_1 and Z_2 are the specific acoustic impedances of the bond material and transducer material, respectively. Using similar conventions, k_1 and k_2 are propagation constants in the respective materials, as Δ_1 and Δ_2 are the respective bond and transducer thicknesses, and j is $\sqrt{-1}$.

The phase angle (McSkimin, 1961) is calculated by writing Eq. (42) in complex polar form and gives

$$\gamma = \arctan\left(\frac{2|Z_d|Z_s}{|Z_d|^2 - Z_s^2}\right), \tag{44}$$

where $|Z_d|$ is the modulus of Z_d, and $(\pi + \gamma)$ is the phase shift between the reflected and incident pressure waves impinging on the interface of the transducer bond system with the sample.

When the system is driven at the transducer resonant frequency, f_r can be calculated by using the fact that the transducer at fundamental resonance has a thickness of one-half of a wavelength; then $|Z_d|$ becomes $Z_1 \tan k_1\Delta_1$, which is generally quite small for a thin bond. Frequencies of off-transducer resonance, generally chosen as $0.9f_r$, give a larger phase angle, since the dependence of Z_d upon off-resonance frequency excitation is large. As outlined in McSkimin (1961) and Papadakis (1967), we can calculate the time difference between phase-matched echo trains, ΔT, caused by the change of drive frequency, and compare it with experimental results to determine the correct condition for overlap.

Assume an alignment of, for example, the first and second echoes in the sample. The change in the measured time between echos is given by

$$\Delta T = \frac{1}{f_L}\left(n - \frac{p\gamma_L}{2\pi}\right) - \frac{1}{f_R}\left(n - \frac{p\gamma_R}{2\pi}\right), \tag{45}$$

where p is the number of round trips (for our case, $p = 1$) of the ultrasonic tone burst in the sample, n is the number of cycles of mismatch in the tone burst, f_L is the off-transducer resonance drive frequency, f_R is the drive frequency at transducer resonance, γ_L is the phase angle at the off-resonance frequency (radians), and γ_R is the phase angle (radians) at the transducer resonance. One can calculate ΔT for the case of no cycle mismatch ($n = 0$) and no bond thickness ($\gamma_R = 0$) from Eqs. (43) through (45). The measured value closest to the calculated value determines the correct experimental cycle for cycle match. Following the experimental identification of the correct match, the measured ΔT can be used to determine the correction for bond thickness by adjusting Δ_1 in Eq. (43) and solving Eq. (44) to bring Eq. (45) into agreement with the measured ΔT values. The phase angle at resonance can be determined, and the travel time correction for bond thickness, Δ^B, is given by

$$\Delta^B = \frac{\gamma_R}{2\pi f_R}, \tag{46}$$

where γ_R and f_R are the phase angle, and drive frequency respectively, both taken at transducer resonance.

b. Diffraction Corrections

Diffraction effects for compressional waves in various crystalline symmetries have been treated by several investigators (Seki *et al.*, 1956; Papadakis, 1963, 1964b, 1966; Williams, 1970), while others have treated isotropic media (Benson and Kiyohara, 1974; Khimunin, 1972; Rogers and Van Buren, 1974). Papadakis (1972) has shown that without diffraction corrections, one can expect errors in travel times as large as $0.25/f_R$. He also discussed (1972) diffraction corrections for the technique of pulse-echo overlap technique. Using the dimensionless quantity

$$S = \frac{z\lambda}{a^2}, \tag{47}$$

where S is the Seki parameter, a is the transducer radius, λ is the wavelength in the sample, and z is the distance of propagation, he wrote the phase correction due to diffraction, Δ^D, as

$$\Delta^D = \frac{[\phi(S_n) - \phi(S_m)]}{2\pi f}, \tag{48}$$

where ϕ is the phase shift in radians of the received wave front that is due to diffraction effects. For the example, the sample is of length L, which gives $S_n = 2L\lambda/a^2$, and $S_m = 4L\lambda/a^2$.

The true travel time, δ, can be written in terms of the measured time, T, and the corrections:

$$\delta = T + \Delta^B + \Delta^D. \tag{49}$$

The thickness of the sample, L, is measured by conventional means, such as a high-precision micrometer, and the velocity of sound is given by

$$v = \frac{2L}{\delta}. \tag{50}$$

7. Optical Techniques

Optical techniques have been used with success in the measurement of sound velocities in transparent media. The methods include diffraction and scattering of light by sound waves. A direct method of optically measuring an acoustic wavelength in the material can also be used.

The diffraction techniques considered here permit the determination of the wavelength of sound. The wavelength, λ^*, is combined with the ultrasonic frequency, f, of the sound beam to calculate its velocity, according to the expression

$$f\lambda^* = v, \tag{51}$$

where v is the speed of sound in the sample. In any optical technique, however, one must specify the type of diffraction experienced by the light.

a. Fraunhofer Diffraction of Light by Sound

Consider an ultrasonic transducer, bonded to a transparent solid, generating ultrasonic compressional waves that propagate into the solid. Further

suppose that the sound wave encounters light traversing the same medium so that the light is diffracted. There are two different physical regimes that can produce the diffraction effects. One involves the formation of a corrugation in the phase fronts of the light that is due to the spatial variation in the index of refraction of the solid, as caused by the sound waves. This is called Raman-Nath diffraction (Raman and Nath, 1935a, 1935b, 1936a, 1936b, 1936c; Born and Wolf, 1970). The second regime involves the reflection of light from the evenly spaced crests of the sound waves. These reflections occur under some conditions that are similar to x-ray diffraction by a crystal lattice; this is called Bragg diffraction (Bhatia and Noble, 1953). Both types of diffraction can be used to determine the sound velocity of the material.

In order to determine which type of diffraction effect predominates (Nomoto, 1942), a dimensionless parameter (Klein *et al.*, 1965), Q, is defined as

$$Q = \frac{K^{*2}L}{\mu_0 K}, \tag{52}$$

where K^* is the ultrasonic propagation constant, L is the width of the ultrasonic beam, μ_0 is the index of refraction, and K is the propagation constant of light in vacuum. If $Q > 9$, one has Bragg diffraction. If $Q < 1$, Raman-Nath diffraction occurs. For $1 < Q < 9$, the diffraction is mixed. For illustrative purposes, consider a typical transparent solid ($c_{sound} \approx 5 \times 10^3$ m/s, index of refraction ≈ 1.5, and ultrasonic beam width $\approx 1.27 \times 10^{-2}$ m), illuminated with light at wavelength 632.8 nm. If the ultrasonic frequency is approximately 27 MHz or less, the interaction satisfied Raman-Nath diffraction conditions. If the frequency is greater than approximately 81 MHz, the interaction is governed by Bragg diffraction conditions.

b. Raman–Nath Diffraction

For the case of Raman-Nath diffraction, we consider the light beam impinging on the medium at an angle ϕ, with the light normal to the sound beam. The location of the diffraction orders are given by the expression (Breazeale *et al.*, 1981)

$$\sin(\theta_n + \phi) - \sin(\phi) = \frac{n\lambda}{\lambda^*}, \tag{53}$$

where ϕ is the angle of incidence, θ is the angle of diffraction, λ is the wavelength of light in the medium, λ^* is the ultrasonic wavelength, and n is an

integer. One can experimentally set $\phi = 0$, which reduces Eq. (53) to

$$\sin(\theta_n) = \frac{n\lambda}{\lambda^*}. \tag{54}$$

The ultrasonic wavelength can be determined from Eq. (53). With the measurement of the ultrasonic frequency, one can determine the sound velocity of the compressional wave (Barnes and Hiedemann, 1957). In determining the wavelength measurement and its uncertainty, it is necessary to make an analysis of the optical setup, including the effects of Snell's Law at interfaces.

c. Bragg Diffraction

For the case of Bragg diffraction, the angle of incidence, ϕ, is set to the angle of diffraction, and

$$n\lambda = 2\lambda^* \sin \phi_B, \tag{55}$$

where ϕ_B is the Bragg angle. Bragg diffraction is used to measure wave velocities in the frequency range from approximately 100 MHz to the low end of the gigahertz scale. For example (Krischer, 1968), the technique has been used to measure wave velocity to an estimated accuracy of better than 0.1%. It is also useful in measuring the local velocities within a sample (Simondet *et al.*, 1976; Michard and Perrin, 1978). Measurements with precision of better than 0.01% in homogenous samples have been reported (Simondet *et al.*, 1976).

d. Direct Measurements (Fresnel Diffraction)

Consider a standing ultrasonic wave in a sample through which collimated light is passed so that the collimated light beam is perpendicular to the sound beam. A measuring microscope or similar optical device is focused so that the image of the wave can be viewed at the instrument focal plane. Because the images of the wave fronts are $\lambda/2$ apart, it is possible to determine the ultrasonic velocity from measurements of ultrasonic frequency and measurements of the wavelength in the medium. The technique is sensitive enough to detect local variations in velocity greater than 0.01% (Hiedemann and Hoesch, 1934, 1937; Mayer and Hiedemann, 1958, 1959).

e. Brillouin Scattering

Consider the scattering of a photon by a high-frequency phonon traveling in a specific direction within a crystal. Application of conservation of energy and momentum to the scattering, coupled with the approximation that any

photon frequency shift is small (Benedek and Fritsch, 1966; Beyer and Letcher, 1969, pp. 47-50), gives

$$v^* = \pm 2v\left(\frac{v}{C}\right)\sin\left(\frac{\theta}{2}\right), \tag{56}$$

where v is the photon frequency, v^* is the phonon frequency as well as the difference in frequency of the scattered photons, C is the speed of the photon in the medium, v is the speed of the phonon, and θ is the scattering angle. If desirable, the Bragg condition can be used for constructive reinforcement by adjusting θ (Pollard, 1977). The technique gives three lines in the scattered photon spectra, which can be separated and measured with appropriate optical devices, such as Fabry-Perot interferometers, to obtain v^*. Equation (56) can be solved for v in terms of the other quantities. Brillouin scattering is useful in the investigation of the sound velocity of a material near a phase transition (Fleury, 1970, pp. 37-42). The technique has been used for both longitudinal and mixed modes. However, in one study on cubic crystals, no Brillouin scattering was observed from an acoustic branch that consisted of pure transverse waves (Benedek and Fritsch, 1966). Uncertainties in the index of refraction, the scattering angle, and the width of the Stokes and anti-Stokes lines influence the accuracy of the determination. The precision is considerably better, however, with values of 0.1% as mentioned.

In stimulated Brillouin scattering, the photon scattering process is dependent upon the intensity of the radiation striking the surface. With high enough light intensity, nonlinear effects occur, which result in scattering by frequencies and harmonics created by harmonic generation (Brewer, 1965). A large buildup of acoustic intensity, both compressional and shear waves, accompanies a threshold in optical intensity (Chiao *et al.*, 1964). Other effects include the possibility of sample destruction from the intense radiation, and the line pulling effects of the laser cavity on the scattered light, which affects accuracy (Fleury, 1970, pp. 57-58). Amplified acoustic frequencies have been reported as high as 60 GHz.

III. Theoretical and Experimental Results

8. *Ab Initio* Theory

It is clear from Section 2 that the calculation of elastic constants requires an accurate computation of the variation in the total energy of a solid, from equilibrium to distorted configurations. Thanks to the availability of powerful computers, great advances in *ab initio* total-energy calculation have been

made in recent years (Anderson *et al.*, 1985). In this section we briefly discuss the approximations used in *ab initio* theory and summarize the calculated results.

The Hamiltonian of a solid consists of five parts: the two kinetic energies of the ions and electrons, K_I and K_e respectively, and the potential energies U_{II} among ions, U_{ee} among electrons, and U_{eI} between electrons and ions. The Born-Oppenheimer (1927) adiabatic approximation is a simple way to separate the electronic from the ionic variables. In this approximation, because the ion speed is at least two orders smaller than the electron speed, one freezes the ionic motion in a configuration specified by a set of ion position vectors $\vec{R}_n$, then solves the Schrodinger equation for that part of the Hamiltonian that involves the set of electronic coordinates $\{\vec{r}_i\}$:

$$H\psi_\gamma(\{\vec{r}_i\}, \{\vec{R}_n\}) = E_\gamma(\{\vec{R}_n\})\psi_\gamma(\{\vec{r}_i\}, \{\vec{R}_n\}), \tag{57}$$

where $H = K_e + U_{ee} + U_{eI}$. Thus the energy E_γ is a function of the ionic configuration. The lowest energy curve of the sum of $E_g = E_\gamma + U_{II}$ as a function of $\{R_n\}$ then serves as the potential energy for the ionic motion. A Taylor expansion of E_g about its minimum value E_0, takes the form

$$E_g\{\vec{R}_n\} = E_0 + \tfrac{1}{2}\Sigma_{nm\alpha\beta}\Phi^{\alpha\beta}_{nm}u_{n\alpha}u_{m\beta} + \cdots, \tag{58}$$

where $u_{n\alpha}$ is the α-component of the small displacement $\vec{R}_n - \vec{R}_n(0)$ with $\{\vec{R}_n(0)\}$ the equilibrium ionic positions at the minimum energy E_0. The force functions $\Phi^{\alpha\beta}_{nm}$ are the second derivatives of E_g with respect to the displacements evaluated at the ion equilibrium positions. The force functions are directly related to the elastic constants. For example, for a Bravais lattice they are given by Aschcroft and Mermin (1976) as

$$c_{\alpha\beta\mu\nu} = -\tfrac{1}{2}\Sigma_{\vec{L}}L_\alpha\Phi^{\beta\mu}(\vec{L})L_\nu, \tag{59}$$

where the sum is over all the lattice vectors $\vec{L}$.

Equation (57) is computationally the most difficult part of the problem, because it deals with about 10^{23} electrons that are interacting with each other and with the ions; the wave functions have to be the properly antisymmetrized many-body functions. Self-consistent density-functional theory (SCDFT) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965; Callaway and March, 1984), which casts the Hamiltonian into a density functional, reduces the many-body problem to an effective one-electron problem. This theory has been tested in many crystalline solids, and the resulting elastic constants have been excellent, especially for semiconductors.

In SCDFT, the ground-state energy of a solid is completely specified by single-particle wave functions of the occupied states $\{\phi_v\}$. First, the electron density is given by

$$\rho(\vec{r}) = \Sigma_v |\phi_v(\vec{r})|^2 . \tag{60}$$

Then, the ground state energy is computed as follows:

$$E_g = K_e + U_{ee} + U_{el} + U_{II} + U_{xc}, \tag{61}$$

where the different terms are given by the expressions

$$K_e = \Sigma_v \int \phi_v^*(\vec{r}) \frac{p^2}{2m} \phi_v(\vec{r})\, d^3r, \tag{62}$$

$$U_{ee} = \frac{e^2}{2} \int\int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r\, d^3r' . \tag{63}$$

$$U_{el} = -e^2 \Sigma_n Z_n \int \frac{\rho(\vec{r})}{|\vec{r} - \vec{R}_n|} d^3r, \tag{64}$$

$$U_{II} = \frac{e^2}{2} \Sigma'_{mn} Z_m Z_n / |\vec{R}_m - \vec{R}_n|, \tag{65}$$

$$U_{xc} = \int \rho(\vec{r}) \varepsilon_{xc}[\rho(\vec{r})] d^3r . \tag{66}$$

Note that $n = m$ is excluded in the sum in Eq. (65). The meaning and the notations of the above equations are mostly self-evident, except the $\varepsilon_{xc}[\rho]$ in Eq. (66); this is the correction term arising from the many-body exchange and correlation effects. The square bracket means that ε_{xc} is a *functional* of ρ. Several different expressions for ε_{xc} as a function of ρ, available in the literature (for example, Wigner, 1934; Hedin and Lunquist, 1971; Ceperley and Alder, 1980; Perdew and Zunger, 1981), have yielded similar results for the structural properties of semiconductors. A minimization of E_g with respect to ϕ_v^*, with the constraint that the total number of electrons is a constant, leads to the familiar Schroedinger single-particle equation:

$$\left[\frac{p^2}{2m} + V(\vec{r})\right] \phi_v(\vec{r}) = \varepsilon_v \phi_v(\vec{r}), \tag{67}$$

where ε_v is a Lagrange multiplier and V is an effective one-electron potential containing three parts:

$$V = V_{ee} + V_{el} + V_{xc}, \tag{68}$$

$$V_{ee} = e^2 \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r'. \tag{69}$$

V_{el} is the Coulomb potential due to ionic charges Z_n:

$$V_{el} = -e^2 \Sigma_n Z_n / |\vec{r} - \vec{R}_n|. \tag{70}$$

Finally, the exchange-correlation potential V_{xc} is given by

$$V_{xc} = \varepsilon_{xc} + \rho \frac{\partial \varepsilon_{xc}}{\partial \rho}. \tag{71}$$

Thus, Eqs. (60) and (67)-(71) form a repeated loop, $\rho \rightarrow V \rightarrow \phi \rightarrow \rho$, and the calculation must be iterated until self-consistency is achieved. Following this recipe for calculating the total energy, one calculates E_s as a function of ionic positions, finds the equilibrium configuration, then imposes a strain and calculates the strain energy to deduce the elastic constants, following the prescription of Section 2. The problem then becomes strictly computational. The most challenging task is an accurate solution for the single-particle eigen states of Eq. (67). For a crystalline solid, lattice translational symmetry simplifies the problem, and band-structure techniques can be applied to obtain the solution. Because the strain energy is many orders of magnitude smaller than the total energy, very precise computation is required if one hopes to obtain meaningful elastic constants. So far, at least two band-structure methods have been demonstrated as reliable for all three elastic constants: the plane-wave method using pseudopotentials (PP-PW) (Nielsen and Martin, 1983, 1985a), and the full-potential linearized-muffin-tin-orbital method (FP-LMTO) (Methfessel *et al.*, 1989). Although the full-potential argumented plane-wave method (FP-APW) (Krahauer *et al.*, 1979; Wimmer *et al.*, 1981; Wei and Krahauer, 1985; Ferreira *et al.*1989) has produced excellent lattice constants, structural energies, and bulk moduli, the complete semiconductor elastic constants based on this method are not yet available.

Even if the total energies at different distortions can be calculated accurately, there is still the problem of searching for the equilibrium atomic positions in a distorted crystal, and the numerical determination of elastic constants from energy differences. If the strain energy can be calculated directly without taking the difference between two large energies, or if the

derivatives can be calculated directly, not only can the computation time be shortened, but the numerical errors will also be reduced. The quantum mechanical theory of forces and stresses of Nielsen and Martin (1985b) and the closely related direct calculation of elastic constants from linear response theory of Baroni *et al.* (1987) represent the status of efforts in this direction. The former has been carried out for all three elastic constants for Si, Ge and GaAs (Nielsen and Martin, 1985a), while the latter has been done only on the bulk modulus for Si; both are based on the PW-PP method. Table II shows a

TABLE II

COMPARISON BETWEEN CALCULATED AND EXPERIMENTAL CONSTANTS[a]

		Expt[b]	FP-LMTO[c]	PP-PW[d]	TB
SI					
	a	5.431	5.41	5.45	5.431
	B	9.923	9.9	9.3	9.923
	$C_{11}-C_{12}$	10.274	10.2	9.8	10.274
	C_{44}	8.036	8.3	8.5	8.013
	$C_{44}^{(0)}$			11.1	11.30
	ζ	0.54[e]	0.51	0.53	0.51
	ω	523	518	521	572
Ge					
	a	5.65		5.59	5.65
	B	7.653		7.2	7.653
	$C_{11}-C_{12}$	8.189		8.5	8.189
	C_{44}	6.816		6.3	6.84
	$C_4^{(0)}$			7.7	9.46
	ζ			0.44	0.49
	ω	303		302	342
GaAs					
	a	5.642		5.55	5.642
	B	7.69		7.3	7.69
	$C_{11}-C_{12}$	6.63		7.0	6.63
	C_{44}	6.04		6.2	5.79
	$C_{44}^{(0)}$			7.5	7.83
	ζ			0.48	0.50
	ω	273		268	292

[a]Comparison between calculated and experimental lattice constant a, elastic constants B, $C_{11}-C_{12}$ and C_{44}, Kleiman (1962) internal distortion parameter ζ, and the TO optical phonon ω in wave numbers 1/cm. Also listed are $C_{44}^{(0)}$ defined in Eq. (13). The FP-LMTO and PP-PW are the *ab initio* theories described in Part I, and TB is the tight-bonding theory in Part III, Section 10. All elastic constants are in units of 10^{11} dynes/cm^2.

[b]All the experimental lattice constants are those tabulated by Zallen (1982). The experimental elastic constants are taken from Table III, and the phonon frequencies are from Table V.

[c]Methfessel *et al.* (1989).

[d]Nielsen and Martin (1985a).

[e]Cousins *et al.* (1987).

comparison between theoretical calculations and experimental results. From this comparison, it is fair to say that we have a very reliable *ab initio* theory for the elastic constants for crystalline semiconductors based on the self-consistent local density-functional theory. Note that Table II also lists the results from an empirical tight-binding (TB) theory to be discussed in Section 10.

9. Valence Force Field Model

The preceding section showed that *ab initio* theory for the elastic constants requires complicated computations. Accurate *ab initio* calculations for semiconductors have been obtained only recently and only for several systems. On the other hand, phenomenological microscopic models of elastic constants for all semiconductors have been available for some time. Of these, the valence force-field model (VFF) is perhaps the simplest and the most useful. This topic has been reviewed and well analyzed in a paper by Martin (1970), and its conclusions constitute the main body of this section.

a. Diamond Structure

The original VFF model by Musgrave and Pople (1962) was for the diamond structure, in which the elastic energy is a quadratic form, in terms of the changes in each bond length Δr_i, in bond angles $\Delta\theta_{ij}$, and in the products $\Delta r_i \Delta r_j$ and $\Delta\theta_{ij}$, between nearest-neighbor bonds. For elastic constants, Keating (1966) showed that the VFF can be simplified by the following approximation to the elastic energy of the crystal:

$$\Delta E = \frac{3\alpha}{8d^2}\sum_i [\Delta(\vec{r}_i \cdot \vec{r}_i)]^2 + \frac{3\beta}{8d^2}\sum_{i>j} [\Delta(\vec{r}_i \cdot \vec{r}_j)]^2, \tag{72}$$

where the bond index i runs over all the bonds, but the i and j sum only over those pairs of bonds that are connected to a common atom. In Eq. (72), d is the equilibrium bond length and $\Delta(\vec{r}_i \cdot \vec{r}_j)$ is the change in the dot product of the two bond vectors that start at the common atom, point along the bond directions, and end at the first neighbor atoms. Following the calculational procedure described in Section 2 for uniform expansion, the U in the VFF under a uniform expansion can be shown to be $U = 2(3\alpha + \beta)e^2d^2/\Omega$, where $\Omega = a^3/4$ is the equilibrium volume per unit cell. Thus,

$$B = (C_{11} + 2C_{12})/3 = (\alpha + \beta/3)/a. \tag{73}$$

For the shear strain described by Eq. (9), Eq. (72) yields $U = 4\beta e^2/a$. Thus, according to Eq. (10),

$$C_{11} - C_{12} = 4\beta/a. \tag{74}$$

For the shear strain described by Eq. (11) and with an internal displacement $\vec{u} = (0, 0, u)$, Eq. (72) yields the following expression:

$$U = [\alpha(e - \eta)^2 + \beta(e + \eta)^2]/(8a), \tag{75}$$

where $u = \eta a/4$. A comparison between Eqs. (13) and (75) shows that $\phi = 16(\alpha + \beta)/a^3$, $D = -4(\alpha - \beta)/a^2$, and $C_{44}^{(0)} = (\alpha + \beta)/a$. Using these results in Eq. (14), we find that the Kleiman internal displacement parameter in the present model is given by

$$\zeta = (\alpha - \beta)/(\alpha + \beta) = 2C_{12}/(C_{11} + C_{12}). \tag{76}$$

Equation (15) then produces

$$C_{44} = 2\alpha\beta/[(\alpha + \beta)a]. \tag{77}$$

The three elastic constants given above are not independent and can be shown to related to each other by the Keating identity (1966), or

$$I_K = 2C_{44}(C_{11} + C_{12})/[(C_{11} - C_{12})(C_{11} + 3C_{12})] = 1. \tag{78}$$

b. Zincblende Structure and Coulomb Force

The Keating identity (1966) holds very well for systems with the diamond structure but not so well for the zincblende compounds (see Table III). One obvious difference between the two structures is the presence of Coulomb interactions arising from charge shifts between the cation and anion sublattices in zb semiconductors. Martin (1970) incorporated Blackman's (1959) treatment of the Coulomb forces in the Keating VFF in the following manner. First, the Coulomb energy was treated as a screened Madelung energy E_M. For example, in a uniformly expanded crystal with a bond length r, the Coulomb energy was taken to be $E_M = -N\alpha_M Z^{*2}e^2/(\varepsilon r)$, where N is the total number of unit cells, $\alpha_M = 1.6381$ is the Madelung constant, and Z^{*2}/ε is the effective charge defined by the optic-mode splitting:

$$S = Z^{*2}/\varepsilon = \mu(\omega_l^2 - \omega_t^2)/(4\pi e^2). \tag{79}$$

TABLE III

EXPERIMENTAL VALUES FOR CUBIC SEMICONDUCTORS[a]

	C_{11}	C_{12}	C_{44}	s	α	β	I_K	I_M	I_{BOM}
C[b]	107.640	12.520	57.740	0.0	129.100	84.573	1.00	1.00	1.02
Si[b]	16.772	6.498	8.036	0.0	49.247	13.951	1.00	1.00	1.13
Ge[b]	13.112	4.923	6.816	0.0	39.438	11.583	1.08	1.08	1.05
AlSb[b]	8.769	4.341	4.076	1.684	33.768	6.653	1.11	1.05	1.08
GaP[b]	14.390	6.520	7.143	3.815	46.965	10.448	1.12	1.05	1.08
GaAs[b]	12.110	5.480	6.040	2.827	40.895	9.159	1.12	1.06	1.05
GaSb[b]	9.089	4.143	4.440	1.569	33.123	7.412	1.10	1.06	1.06
InP[b]	10.220	5.760	4.600	3.766	41.095	6.250	1.20	1.07	1.03
InAs[b]	8.329	4.526	3.959	2.820	33.744	5.531	1.22	1.11	1.00
InSb[b]	6.918	3.788	3.132	1.372	29.909	4.951	1.17	1.11	1.05
ZnS[c]	9.420	5.680	4.360	6.788	37.026	4.571	1.33	1.07	0.95
ZnS[c]	10.790	7.220	4.120	6.788	45.126	4.341	1.28	1.02	1.01
ZnS[c]	9.810	6.270	4.483	6.788	39.947	4.300	1.42	1.13	0.90
ZnS[c]	10.460	6.530	4.630	6.788	41.880	4.828	1.33	1.08	0.95
ZnSe[b]	8.95	5.39	3.984	4.368	34.432	4.716	1.28	1.09	0.98
ZnSe[c]	8.59	5.06	4.06	4.368	34.519	4.673	1.32	1.13	0.95
ZnSe[c]	8.720	5.240	3.920	4.368	35.469	4.603	1.29	1.10	0.98
ZnTe[c]	7.130	4.070	3.120	2.566	29.976	4.452	1.18	1.06	1.05
ZnTe[c]	7.220	4.090	3.080	2.566	30.204	4.558	1.14	1.03	1.08
CdTe[b]	5.33	3.65	2.04	3.105	27.058	2.455	1.34	1.07	0.98
CdTe[c]	6.150	4.300	1.960	3.105	31.546	2.731	1.16	0.94	1.13
HgTe[b]	5.971	4.154	2.259	2.381	30.300	2.542	1.37	1.16	0.96
HgTe[c]	5.63	3.66	2.11	2.381	26.919	2.542	1.37	1.15	0.95

[a]Experimental elastic constants for some cubic semiconductors and the parameters of Eq. (81) taken from Martin (1970), with the force constants α and β obtained from Eqs. (82) and (83) and the identity relations I_K, I_M, and I_{BOM} given by Eqs. (78), (88), and (119), respectively. The elastic constants are in units of 10^{11} dynes/cm^2, and the force constants are in 10^3 dynes/cm.

[b]Data quoted from "Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology," New Series, Vols. 17 and 22.

[c]Listed in the review by Mitra and Massa (1982).

In Eq. (79), ω_1 and ω_t are, respectively, the longitudinal and transverse phonon frequencies in the long-wave-length limit. Then, to counterbalance the Coulomb forces, a repulsive force term was added and assumed to contribute to the bond-stretching energy in the form

$$\Delta E_R = -\sum_i \alpha_M Z^{*2} e^2 \, \Delta r_i/(4\varepsilon d^2). \tag{80}$$

With the above two contributions added, the total strain energy is $\Delta E_T = \Delta E + \Delta E_M + \Delta E_R$, where ΔE is the VFF contribution in Eq. (72) and ΔE_M is

the change in the Madelung energy. The energies ΔE_T are expanded in a power series, and only terms up to the second power in the strain are kept. The ΔE_M contributions arising from fixed values of the charge shift $S = Z^{*2}/\varepsilon$ on the atomic sites under different strains were worked out by Blackman (1959). Using these results and defining

$$s = Se^2/d^4 = e^2Z^{*2}/(d^4\varepsilon), \tag{81}$$

Martin (1970) obtained the following modified expressions for the elastic constants:

$$C_{11} + 2C_{12} = (3\alpha + \beta)/a - 0.355s, \tag{82}$$

$$C_{11} - 2C_{12} = 4\beta/a + 0.053s, \tag{83}$$

$$\zeta = [(\alpha - \beta)/a - 0.294s]/C_M, \tag{84}$$

$$C_{44} = (\alpha + \beta)/a - 0.136s - C_M\zeta^2, \tag{85}$$

where C_M is defined as

$$C_M = (\alpha + \beta)/a - 0.266s. \tag{86}$$

The above equations can be combined to yield

$$\zeta = (2C_{12} - C')/(C_{11} + C_{12} - C'), \tag{87}$$

where $C' = 0.314s$. Since the extra parameter s is fixed by the optic modes and the bond length, the above results combine into a new identity, the Martin identity (1970):

$$I_M = \frac{2C_{44}(C_{11} + C_{12} - C')}{(C_{11} - C_{12})(C_{11} + 3C_{12} - 2C') + 0.831C'(C_{11} + C_{12} - C')}. \tag{88}$$

Table III lists a set of experimental values of the elastic constants and the s values for a number of diamond and zincblende semiconductors. These values are used to compute the force constants α and β and the identity expressions I_k and I_M given in Eqs. (78) and (88), respectively. Several sets of data are quoted for some of the systems to show the uncertainties in the experiments for these systems. The table results clearly show that the inclusion of the Coulomb energies improves the identity relation; the deviations of I_M from unity are 15% or less. Also listed are the values for another identity relation, I_{BOM} from Eq. (119), based on a tight-binding

model to be discussed in Section 10. Martin further studied trends as functions of the bond lengths d and the ionicity scale f_i of Phillips (1973) and Van Vechten (1969). He found that α scales roughly as $1/d^3$, i.e.,

$$\alpha d^3/e^2 = \text{constant}, \tag{89}$$

where e is the electron charge. He also found the ratio between the bond-angle and bond-stretching forces tend to decrease as f_i increases and scales roughly as

$$\beta/\alpha \propto 1 - f_i. \tag{90}$$

He further observed that if S of Eq. (79) is set equal to f_i and if the α and β values are extrapolated, using Eqs. (89) and (90), from those fitted to the average values of the B and $C_{11} - C_{12}$ for Si and Ge, then all the elastic constants can be predicted from Eqs. (82) through (86) to an accuracy of 10%.

It is interesting to compare Eq. (82) using the results of Eqs. (89) and (90), with Cohen's (1985) empirical formula for the bulk modulus

$$B = (1971 - 22\lambda)/d^{3.5}, \tag{91}$$

where B is in GPa, d in Å, and $\lambda = 0$, 1, and 2, respectively, for the group IV, II-V, and II-VI semiconductors. Both Martin and Cohen's formulas give B values to better than 10% for all materials tabulated in Table III. The B in Eq. (91) scales as $1/d^{3.5}$, while in VFF it scales as $1/d^4$.

10. Tight-Binding Theory

In the semi-empirical tight-binding (TB) approach, the total energy of a semiconductor crystal is assumed to be the sum of the electron energies $\varepsilon_v(\vec{k})$ in the valence bands plus repulsive pair energies u_{ij} between the nearest-neighbor atoms (Chadi, 1978):

$$E_T = E_{bs} + U_r = \sum_v \sum_k \varepsilon_v(\vec{k}) + \sum_{i>} \sum_j u_{ij}. \tag{92}$$

Furthermore, the band energies are interpolated by using a TB Hamiltonian that contains term values of the atoms, and a handful of interaction parameters between orbitals of the neighboring atoms. Despite the simplicity of Eq. (92), recent first-principles theories have given some support to this approximation.

One virtue of the TB approach over valence force-field models is that it is a quantum theory without much complication. As compared to first-principles theory, the TB approach is easier to execute, particularly when applied to complicated systems such as alloys and superlatices. In actual applications, the TB calculation either is carried out using the full band-structure calculation (BS), or is approximated by simpler local theory such as Harrison's (1980, 1983a, and 1983b) bond orbital model (BOM).

a. Band Structure Calculations

The simplest TB Hamiltonian contains the s and p atomic term values ε_s and ε_p for both cations and anions, and the nearest-neighbor two-center interactions $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$, and $V_{pp\pi}$. To be more explicit, the 8×8 k-dependent Hamiltonian contains the term values as the diagonal matrix elements, while the off-diagonal matrix elements between the cation and anion orbitals are given by

$$H_{\gamma\gamma'}(\vec{k}) = \sum_{\vec{d}} e^{i\vec{k}\cdot\vec{d}} h_{\gamma\gamma'}(\vec{d}), \tag{93}$$

where the sum runs over the four first neighbor atoms specified by the bond displacements $\vec{d}$. The γs are the orbital indices for s, p_x, p_y, and p_z. The $h_{\gamma\gamma'}$ values are related to the two-center V's by the Slater-Koster (1954) relations:

$$h_{ss} = V_{ss\sigma}, \tag{94}$$

$$h_{sx} = \alpha_1 V_{sp\sigma}, \tag{95}$$

$$h_{xx} = \alpha_1^2 V_{pp\sigma} + (1 - \alpha_1^2) V_{pp\pi}, \tag{96}$$

$$h_{xy} = \alpha_1 \alpha_2 (V_{pp\sigma} - V_{pp\pi}), \tag{97}$$

where $\alpha_i = x_i/d$ are the direction cosines of $\vec{d}$ and the V's depend only on the length d.

Once the values of these TB parameters and their dependences on the bond length are known, the Hamiltonian at each k inside the Brillouin zone (BZ) can be evaluated, and the summation of k carried out to obtain the band-structure energy, which, when added to the repulsive energy, gives the total energy of any specified geometry. All the elastic constants, associated internal displacements, and transverse optical phonon frequencies, are easily calculable following the procedure of Section 2. The only point to note is the k-

sum, which without strain can be calculated accurately by using the 10 special k points (Chadi and Cohen, 1973) in the irreducible wedge of the BZ. Under strain, the crystal symmetry changes; one needs to extend these special k points to other nonequivalent wedges. However, since the sum of the valence-band energies as a function of k is a rather smooth function, a uniform sampling over the whole BZ converges very quickly. A $5 \times 5 \times 5$ grid is sufficiently accurate for the required calculations. To avoid the numerical inaccuracy inherent in direct energy subtractions, one can also calculate the second derivatives directly by using perturbation theory.

Perturbation theory starts with the expansion of the k-dependent Hamiltonian H in powers of the infinitesimal strain parameter e, keeping terms up to second-order:

$$H(k) = H_0 + H_1 e + \tfrac{1}{2} H_2 e^2, \tag{98}$$

where H_0 is the strain-free Hamiltonian, and H_1 and H_2 are, respectively, the first and second derivatives with respect to e evaluated at $e = 0$. The band energy contribution to the strain coefficient then comes from the second derivative of E_{bs} with respect to e, denoted by

$$\frac{\partial^2 E_{bs}}{\partial e^2} = \sum_v \sum_k \langle v\vec{k}|H_2|v\vec{k}\rangle + 2 \sum_v \sum_c \sum_k \frac{|\langle v\vec{k}|H_1|c\vec{k}\rangle|^2}{\varepsilon_v(\vec{k}) - \varepsilon_c(\vec{k})}, \tag{99}$$

where $\varepsilon_c(\vec{k})$ and $|c\vec{k}\rangle$ are, respectively, the eigen energies and eigenvectors of H_0 for the conduction bands, and $v\vec{k}$ stands for the valence bands. Note that the inter-valence-band contributions in the second-order perturbation sum cancel exactly so they are not needed in Eq. (99). To evaluate these matrix elements one needs to have the first and the second strain derivatives of the two-center interactions and the direction cosines α_i. For the strain parameters e defined in Section 2, and for the two center interactions V that scale as $1/d^n$, the following results are useful.

(i) For the bulk modulus, the direction cosines do not change, and we have

$$\partial V/\partial e = -nV \qquad \text{and} \qquad \partial^2 V/\partial e^2 = n(n+1)V.$$

(ii) For $C_{11} - C_{12}$ with e specified in Eq. (9), we get

$$\partial \alpha_i/\partial e = \alpha_i(\delta_{i1} - \delta_{i2}),\ \partial^2 \alpha_i/\partial e^2 = \alpha_i[3(\delta_{i1} + \delta_{i2}) - 4]/3, \qquad \partial V/\partial e = 0,$$

and

$$\hat{c}^2 V \partial e^2 = 4nV/3.$$

(iii) For C_{44} with the strain e given in Eq. (11) and an internal displacement u, we find

$$\begin{aligned}
\partial V/\partial e &= -n\alpha_1\alpha_2 V, \qquad \partial V/\partial u = n\alpha_3 v/d, \\
\partial\alpha_i/\partial e &= (\delta_{i1}\alpha_2 + \delta_{i2}\alpha_1)/2 - \alpha_i\alpha_1\alpha_2, \\
\partial\alpha_i/\partial u &= -\delta_{i3}/d + \alpha_i\alpha_3/d, \qquad \partial^2 V/\partial e^2 = n(n-1)V/9, \\
\partial^2 V/\partial u^2 &= n(n-1)V/(3d^2), \\
\partial^2\alpha_i/\partial e^2 &= -\alpha_i(\delta_{i1} + \delta_{i2})/12, \qquad \partial^2\alpha_i/\partial u^2 = -2\delta_{i3}\alpha_3/d^2, \\
\partial^2 V/(\partial e\partial u) &= -n(n+2)\alpha_1\alpha_2\alpha_3/d,
\end{aligned}$$

and finally

$$\partial^2\alpha_i/(\partial e\partial u) = (\delta_{i1}\alpha_2\alpha_3 + \delta_{i2}\alpha_1\alpha_3 + 2\delta_{i3}\alpha_1\alpha_2 - 6\alpha_i\alpha_1\alpha_2\alpha_3)/(2d).$$

Also note that three second derivatives of the band-structure energy are needed, namely $\partial^2 E_{bs}/\partial e^2$, $\partial^2 E_{bs}/(\partial e\partial u)$, and $\partial^2 E_{bs}/\partial u^2$ for the evaluation of $C_{44}^{(0)}$, D, and ϕ of Eq. (13), respectively.

b. Bond Orbital Model

Harrison's (1980, 1983a, and 1983b) bond orbital model (BOM) emphasizes calculations of the TB total energy in terms of local energies. One special feature of the BOM is its universality. Another feature is that its simple and often analytical forms provide direct insight into the essential physics. Although BOM aims at predicting trends, it is reasonably accurate in many cases.

The band-structure energy, or the center of gravity of the valence band, in BOM is computed in the following steps. The terms involved are indicated in Fig. 2.

(1) Construct the sp^3 hybrid orbitals $|h\rangle$ for each atom; these hybrid orbitals are directed toward the neighboring atoms. For example, the hybrid in the [111] direction is given by

$$|h\rangle = (|s\rangle + |p_x\rangle + |p_y\rangle + |p_z\rangle)/2. \tag{100}$$

The hybrid energy is then given by $\varepsilon_h = \langle h|H|h\rangle = (\varepsilon_s + 3\varepsilon_p)/4$. In the zincblende structure, the cation hybrid energy ε_h^C is in general different from the anion hybrid energy ε_h^A. Note that the two hybrid orbitals of the same atom but in two different directions are now coupled by the so-called metallic energy $V_1 \equiv \langle h|H|h'\rangle = (\varepsilon_s - \varepsilon_p)/4$.

(2) Construct the bonding and antibonding molecular orbitals, $|b\rangle$ and $|a\rangle$, from the two hybrid orbitals $|h^C\rangle$ and $|h^A\rangle$ directed toward each other

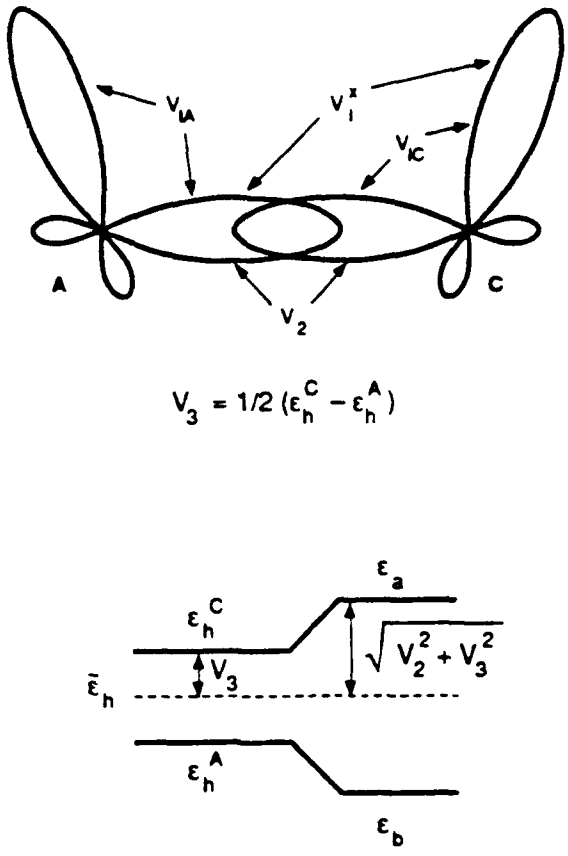


FIG. 2. Labels of the interactions between hybrids associated with an adjacent anion-cation pair. $V_{1A}(V_{1C})$ is the interaction between two hybrids on the same anion (cation); the "covalent energy" V_2 is the interaction between anion and cation hybrids that point toward one another along the bond direction; V_1^x is the interaction between an anion hybrid in one direction and an adjacent cation hybrid pointing in a different direction. The "ionic energy" V_3 is half the difference betwen cation and anion term values. The lower segment of the figure depicts the splitting of the hybrid energy levels by the V_2 and V_3 interactions.

along the same bond by diagonalizing a 2×2 matrix with ε_h^C and ε_h^A on the diagonal, and $V_2 = \langle h^c|H|h^A\rangle$ as the off-diagonal matrix elements. The resulting energies for $|b\rangle$ and $|a\rangle$ are $\varepsilon_b = \bar{\varepsilon}_h - \sqrt{V_2^2 + V_3^2}$ and $\varepsilon_a = \bar{\varepsilon}_h + \sqrt{V_2^2 + V_3^2}$, respectively, where $\bar{\varepsilon}_h = (\varepsilon_h^c + \varepsilon_h^A)/2$ is the mean hybrid energy, V_2 is the covalent energy, and $V_3 = (\varepsilon_h^c - \varepsilon_h^A)/2$ is the polar energy. The eigen states can also be written explicitly in terms of these energies:

$$|b\rangle = \sqrt{(1 + \alpha_p)/2}|h^A\rangle + \sqrt{(1 - \alpha_p)/2}|h^c\rangle, \tag{101}$$

$$|a\rangle = -\sqrt{(1 - \alpha_p)/2}|h^A\rangle + \sqrt{(1 + \alpha_p)/2}|h^c\rangle, \tag{102}$$

where α_p is called the polarity and is defined as

$$\alpha_p = V_3/(V_2^2 + V_3^2)^{1/2}. \tag{103}$$

(3) The quantity ε_b would be the center of gravity of the valence bands, if interactions between states on different bonds were neglected. Harrison (1983b) incorporated these interactions in a perturbation theory in which the change of the bonding energy is given by

$$\Delta\varepsilon_b = \sum_{a'} |\langle b|H|a'\rangle|^2/(\varepsilon_b - \varepsilon_{a'}), \tag{104}$$

where the sum runs over the antibonding states of the surrounding bonds. Note that the interactions among the bonding states lead to the formation of the valence bands, but do not shift their center of gravity; therefore, these interactions need not to be considered in the total energy calculation. Including the energy correction $\Delta\varepsilon_b$, the final band-structure energy per bond (which contains two electrons) is given by

$$E_b = 2\varepsilon_b + 2\Delta\varepsilon_b. \tag{105}$$

Harrison (1980) referred to the second term as the metallization energy. Besides providing a simple means for evaluating the center of gravity of the valence electron, Harrison's BOM also provides a set of universal TB parameters. Based on comparison with the free-electron band width (Froyen and Harrison, 1979) and with empirical TB parameters, Harrison (1983b) deduced the following set of universal two-center interactions:

$$V_{aa'} = \eta_{aa'}\hbar^2/(md^2), \tag{106}$$

with $\eta_{ss\sigma} = -1.32$, $\eta_{sp\sigma} = 1.42$, $\eta_{pp\sigma} = 2.22$, and $\eta_{pp\pi} = -0.63$, where m is the free-electron mass and d the bond length. In units where d is in Å and V in eV, $V_{aa'} = 7.62\,\eta_{aa'}/d^2$. The pair-repulsive energy u in the BOM is taken as resulting from the overlap of wavefunctions of the orbitals on the two centers and was shown to have the form

$$u = u_0(d_0/d)^4, \tag{107}$$

where d_0 is the equilibrium bond length. The value of u_0 is determined by requiring that d_0 is the experimental value. Note that the d dependences of both $V_{aa'}$ and u are taken to be the proper scaling not only among different

systems, but also within the same system, as the bond length varies under distortions.

Under these assumptions, the bond energy E_{bond}, which is defined as the difference between the energy per bond in a semiconductor and the average energy per two electrons in the free atoms, i.e., $E_{\text{bond}} = E_b + u_0 - 2\bar{\varepsilon}$, takes the following simple form for a nonpolar semiconductor (Harrison 1983a):

$$E_{\text{bond}} = V_2(1 - \alpha_m + 9\alpha_m^2/16), \tag{108}$$

where α_m is called the metallicity and is defined as $2V_1/V_2$. For a polar semiconductor, E_{bond} becomes slightly more complicated:

$$E_{\text{bond}} = 2\bar{\varepsilon}_h - 2\bar{\varepsilon} - 2(V_2^2 + V_3^2)^{1/2} \times \left[1 - \frac{1}{2}\alpha_c^2 + \frac{9}{16}\alpha_c^4(V_{1C}^2 + V_{1A}^2)/(V_2^2 + V_3^2)\right], \tag{109}$$

where $\alpha_c = \sqrt{1 - \alpha_p^2}$ is called the covalency, and V_{1C} and V_{1A} are the metallic energies for the cation and the anion, respectively. The bulk modulus also takes a very simple form; for a group-IV semiconductor, it reads

$$B = -2V_2(1 - 9\alpha_m^2/16)/(\sqrt{3}\, d^3), \tag{110}$$

and for a polar semiconductor, it becomes

$$B = -2V_2\left[\alpha_c^2 - \frac{9}{8}\alpha_c^3(5\alpha_c^2 - 4)(V_{1C}^2 + V_{1A}^2)/(V_2^2 + V_3^2)\right]\Big/(\sqrt{3}\, d^3). \tag{111}$$

These expressions show that the bulk modulus varies as $1/d^5$ in the pure covalent case, and as $1/d^9$ in the extreme ionic limit $V_3 \gg V_2$. Note that this result is different from the $1/d^{3.5}$ dependence in Cohen's (1985) formula and the $1/d^4$ scale in VFF.

Shear strains cause a semiconductor to shift away from perfect tetrahedral symmetry. To deal with the shear elastic coefficient, the BOM has to be modified. A simple approximation is the rigid hybrid model (Harrison, 1983b; van Schilfgaarde and Sher, 1987a and b), in which the hybrid orbitals of each atom are assumed to remain in their original tetrahedral directions despite the lattice distortion. Then the hybrids of two nearest-neighbor atoms making up the bonding and antibonding states no longer are directed toward

each other, as shown in Fig. 3. There is a misalignment angle θ between each hybrid and the line connecting the two atoms, and the convalent energy V_2 is given by

$$V_2(\theta) = \frac{1}{4}\left[V_{ss\sigma} - 2\sqrt{3}\cos\theta\, V_{sp\sigma} - \cos^2\theta\, V_{pp\sigma} + 3(1-\cos^2\theta)V_{pp\pi}\right]. \quad (112)$$

The lowest-order change δV_2, caused by an infinitesimal angular misalignment $\delta\theta$, is then given by

$$\delta V_2 = \frac{1}{4}(\sqrt{3}V_{sp\sigma} + 3V_{pp\sigma} - 3V_{pp\pi})(\delta\theta)^2. \quad (113)$$

Under the strain e described in Eq. (9) for $C_{11} - C_{12}$, there is no bond length change, and $(\delta\theta)^2 = 2e^2/3$. If one assumes that the metallization coupling is only through the metallic energies V_{1C} and V_{1A}, as has been assumed so far, then the change of the crystal energy is the change of the band-structure energy due to δV_2. Then, according to Eqs. (10) and (105),

$$\begin{aligned} C_{11} - C_{12} &= \frac{\sqrt{3}}{2d^3}\frac{\partial E_b}{\partial V_2}\delta V_2/e^2 \\ &= \frac{\sqrt{3}}{4d^3}\alpha_c(\sqrt{3}V_{sp\sigma} + 3V_{pp\sigma} - 3V_{pp\pi}) \\ &\quad \times\left[1 + \left(\frac{3}{4} - \frac{9}{8}\alpha_c^2\right)(V_{1C}^2 + V_{1A}^2)/(V_2^2 + V_3^2)\right]. \end{aligned} \quad (114)$$

Harrison (1983b) has pointed out, however, that in addition to V_1, other interactions such as V_1^x shown in Fig. 2 produce important contributions to the shear elastic constant. By arguing that these other contributions must cancel those associated with the change δV_2 arising from the metallization energy in a rigid rotation, Harrison deduced the following expression:

$$C_{11} - C_{12} = \frac{\sqrt{3}}{4d^3}\alpha_c^3(\sqrt{3}V_{sp\sigma} + 3V_{pp\sigma} - 3V_{pp\pi}). \quad (115)$$

Under the strain e for the C_{44} given in Eq. (11) and with an internal displacement given by $u = \eta d/\sqrt{3}$ as described in Section 2, the bond misalignment angles for the four bonds have the same magnitude with $(\delta\theta)^2 = 2(\eta + e/2)^2/9$. The four bond lengths also change, with the change for

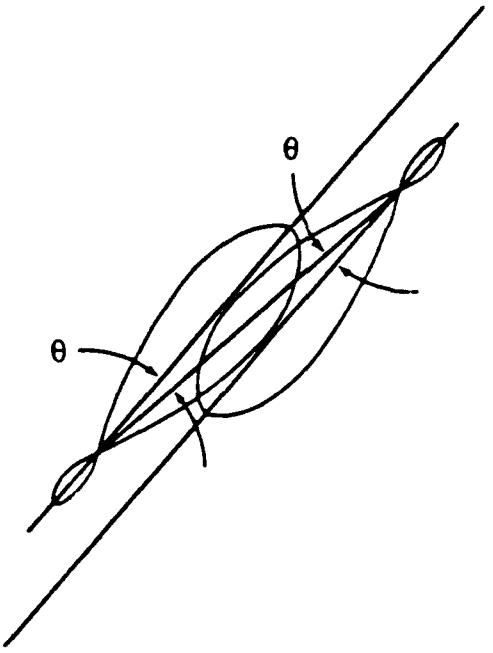


FIG. 3. A schematic picture of the hybrids, treated as rigid, in a shear distortion leading to $C_{11} - C_{12}$.

one pair given by $\delta r_1 = \delta r_2 = \delta + \varepsilon$ and by $\delta r_3 = \delta r_4 = -\delta + \varepsilon$ for the other pair, where $\delta = (e - \eta)d/3$ and $\varepsilon = (\eta + e/2)^2 d/9$. If again one assumes that the metallization is only through V_1, then the strain energy density can be shown to be given by

$$\begin{aligned} U &= 9B\delta^2/(2d^2) + 3(C_{11} - C_{12})(\delta\theta)^2 \\ &= B(e - \eta)^2/2 + (C_{11} - C_{12})(\eta + e/2)^2/3. \end{aligned} \tag{116}$$

For a given strain e, U can be minimized with respect to η, which yields the Kleiman displacement parameter $\zeta = \eta/\varepsilon$ with ζ given by

$$\zeta = (B - C/3)/(B + 2C/3), \tag{117}$$

where $C = C_{11} - C_{12}$. Finally, from $U = C_{44}e^2/2$, the following relationship[7] is established:

$$9/C_{44} = 6/C + 4/B, \tag{118}$$

or

$$I_{\text{BOM}} = 9BC/[C_{44}(6B + 4C)] = 1. \tag{119}$$

[7] This is an expression corrected from one published previously (van Schilfgaarde and Sher, 1987a and b).

If one includes the effect of V_1^x, the energy density will involve an additional term that couples δr and $\delta\theta$. Then the analysis is no longer simple.

The preceding explicit formulas for the elastic constants in BOM are not much more complicated than the valence force-field model. However, they relate macroscopic forces to intrinsic atomic interactions. It is interesting to note that the simple identity relation of Eq. (119) holds very well. As can be seen in Table III, this result is certainly better than the I_k of Eq. (78), and is very competitive to Martin's identity I_M, which requires the inclusion of his particular treatment of Coulomb forces. Because the Coulomb energy is not included explicitly in Eq. (119), its contribution to the elastic constants is probably small.

c. Numerical Results and Quantitative Applications

To study the quantitative aspect of the theory, one first needs to establish the TB parameters. Table IV lists the term values we will use. The values of the outermost valence levels are taken to be minus the experimental first ionization energies listed in Kittel's (1986) book, and the other term values are deduced from calculated extraction or promotion energies using norm-conserved atomic pseudopotentials (Bachelet *et al.*, 1982). These term values are very similar to Mann's (1967) Hartree-Fock calculations used by Harrison (1980), which are also given in Table IV. The major difference in the two sets occurs in heavy elements, where relativistic *s*-shifts are important, but were not included in Mann's results.

Table V lists the values of bond lengths, bond energies E_{bond}, the elastic coefficients B, $C \equiv C_{11} - C_{12}$, and C_{44}, and the zone-center TO phonon frequencies ω for a selected group of systems to be examined in the remainder of this section. Experimental values are also presented, with the exception of extrapolations for the elastic constants of AIP and AIAs. Table VI compares results between the BOM and the full band-structure (BS) calculation using Harrison's universal TB parameters. Except for the Ge and HgTe values, the agreement between the two calculations for E_{bond} is within 10% or better. The calculated E_{bond} values are also in fair agreement with the experimental values except for diamond. Since diamond has a much smaller bond length than the rest of the systems, this discrepancy is an indication of a limit to the scaling rules for both $V_{\alpha\alpha'}$ and u. Although the trends for the bulk moduli from both calculations are similar, the calculational errors in the BOM can be as large as 50%. Also note that the calculated values of B for most systems are only about one-half of the experimental values. The tabulated values of C and C_{44} for BOM are based on Eqs. (115) and (118), respectively. Considering the simplicity of these formulas, the agreement with the band structure calcula-

TABLE IV

TERM VALUES USED IN THE PRESENT CALCULATIONS AND MANN'S (1967) HARTREE-FOCK VALUES AS USED BY HARRISON (1980).

Element	Present		Mann	
	ε_s	ε_p	ε_s	ε_p
Cu	−7.72	−2.96	−7.72	−2.37
Ag	−7.57	−3.54	−7.06	−2.61
Au	−9.22	−3.91	−6.98	−2.67
Be	−9.32	−5.41	−8.41	−5.79
Mg	−7.62	−2.97	−6.88	−3.84
Zn	−9.39	−4.09	−7.96	−4.02
Cd	−8.99	−4.17	−7.21	−3.99
Hg	−10.43	−4.35	−7.10	−3.95
B	−14.00	−8.30	−13.46	−8.43
Al	−11.78	−5.98	−10.70	−5.71
Ga	−13.23	−5.90	−11.55	−5.67
In	−12.03	−5.56	−10.14	−5.37
C	−19.81	−11.26	−19.37	−11.07
Si	−15.03	−8.15	−14.79	−7.58
Ge	−16.40	−7.75	−15.15	−7.33
Sn	−14.53	−7.03	−13.04	−6.76
Pb	−15.25	−6.45	−12.48	−6.53
N	−26.08	−14.54	−26.22	−13.84
P	−19.62	−10.57	−19.22	−9.54
As	−20.02	−9.93	−18.91	−8.98
Sb	−17.56	−8.77	−16.02	−8.14
O	−28.55	−13.561	−34.02	−16.72
S	−21.16	−10.39	−24.01	−11.60
Se	−21.41	−9.90	−22.86	−10.68
Te	−19.12	−9.32	−19.12	−9.54
F	−36.23	−17.44	−42.78	−19.86
Cl	−25.81	−13.05	−29.19	−13.78
Br	−24.95	−12.01	−27.00	−12.43
I	−21.95	−10.79	−22.34	−10.97

tions is remarkable. However, the overall calculated values for these two shear coefficients are also consistently smaller than the experimental values.

The above comparisons show that the BOM and BS calculations predict similar qualitative trends for the binding energies and the elastic constants. In this regard, the BOM has the advantage of providing explicit forms to show the dependences on bond lengths and polarities. However, the merit of the TB theory over the valence-force model is in its ability to incorporate atomic quantities to mimic quantum mechanical effects. To be useful for specific material science applications (for example in alloy surface segregation, see Patrick *et al.*, 1987 and 1988; Sher *et al.*, 1988), the theory has to be more

TABLE V

EXPERIMENTAL PROPERTIES OF SEMICONDUCTORS[a]

	d	E_{bond}	B	C	C_{44}	ω
C	1.540	−3.68	44.227	95.120	57.740	1332
Si	2.352	−2.32	9.923	10.274	8.036	520
Ge	2.450	−1.94	7.653	8.189	6.816	301
AlP	2.367	−2.13	8.600	6.900	6.150	440
AlAs	2.451	−1.89	7.727	7.160	5.420	361
AlSb	2.656	−1.76	5.817	4.428	4.076	366
GaP	2.360	−1.78	9.143	7.780	7.143	367
GaAs	2.448	−1.63	7.690	6.630	6.040	269
GaSb	2.640	−1.48	5.792	4.946	4.440	231
InP	2.541	−1.74	7.247	4.460	4.600	304
InAs	2.622	−1.55	5.794	3.803	3.959	219
InSb	2.805	−1.40	4.831	3.130	3.132	185
ZnS	2.342	−1.59	7.637	3.990	4.558	279
ZnSe	2.454	−1.29	6.457	3.560	3.984	213
ZnTe	2.637	−1.20	5.090	3.060	3.120	177
CdTe	2.806	−1.10	4.210	1.680	2.040	141
HgTe	2.798	−0.81	4.759	1.817	2.259	116

[a]Values of bond length d, bond energy E_{bond}, bulk modulus B, and shear coefficient $C = C_{11} - C_{12}$ used to determine the parameters in Tables VII through XI. Also listed are the experimental values of C_{44} and the TO optical phonon mode ω at Γ to be compared with the calculations. All the elastic constants are in units of 10^{11} dynes/cm^2, d in Å, E_{bond} in eV, and ω in terms of wave numbers in 1/cm. All bond lengths are deduced from the lattice constants quoted by Zallen (1982). The values of E_{bond} are taken from Harrison (1980), Table 7-3, except for AlSb, ZnTe, CdTe, and HgTe, which are deduced from the Phillips (1973) Table 8.2. The elastic constants are taken from Table III, and the phonon frequencies are taken from values compiled in "Landolt-Borstein Numerical Data and Functional Relationships in Science and Technology," New Series, edited by K.-H. Helllwidge, Vols. 17 and 22.

quantitative. Successful quantitative application of the TB theory has been made by Chadi (1978, 1979, and 1984) in his study of semiconductor surfaces. The comparisons in Table VI indicate that the BOM should be treated differently from the BS calculation when considered for quantitative applications. If one wishes to calculate the properties using the local picture, one should use the BOM. If one wants to carry out the TB Hamiltonian precisely, one needs to adopt a different set of parameters based on the BS calculation. For this reason, we shall next consider quantitative applications of BOM and BS calculations separately.

There are many ways to parametrize the TB theory. Chadi (1979, 1984) used a simple form for the repulsive energy $u = a + b(d - d_0) + c(d - d_0)^2$ and the same $1/d^2$ scaling for the TB parameters $V_{\alpha\alpha'}$. To keep the theory as close to Harrison's (1983a and 1983b) form as possible but free it from the

TABLE VI

COMPARISON OF BOM AND BS MODELS' PREDICTIONS.[a]

	E_{bonds}		B		C		C_{44}	
	BS	BOM	BS	BOM	BS	BOM	BS	BOM
C	−7.17	−7.08	49.31	47.39	67.15	67.10	44.45	51.78
Si	−2.39	−2.50	4.65	4.18	7.19	8.07	5.10	5.29
Ge	−2.03	−2.34	2.89	1.91	4.72	6.58	3.69	3.00
AlP	−2.59	−2.54	3.81	4.21	4.45	4.83	3.45	4.10
AlAs	−2.22	−2.20	3.12	3.66	3.85	4.07	2.91	3.50
AlSb	−1.64	−1.62	1.99	2.47	2.75	3.12	2.11	2.54
GaP	−2.32	−2.16	3.63	4.40	4.48	5.15	3.69	4.34
GaAs	−1.99	−1.85	2.83	3.76	3.75	4.30	3.02	3.66
GaSb	−1.53	−1.43	1.76	2.45	2.63	3.40	2.17	2.65
InP	−2.21	−2.06	2.35	2.91	2.69	2.93	2.21	2.63
InAs	−1.89	−1.74	1.93	2.67	2.38	2.52	1.90	2.32
InSb	−1.43	−1.28	1.28	1.97	1.81	2.17	1.48	1.87
ZnS	−1.83	−1.79	3.46	3.74	3.53	3.62	2.71	3.30
ZnSe	−1.52	−1.48	2.67	3.06	2.75	2.67	2.08	2.53
ZnTe	−1.15	−1.05	1.82	2.16	1.88	1.87	1.50	1.78
CdTe	−1.06	−0.97	1.24	1.47	1.22	1.14	0.96	1.13
HgTe	−0.72	−0.49	1.23	1.72	1.33	1.31	1.11	1.30

[a]Comparison of the tight-binding theory using the full band structures (BS) and the bond orbital model (BOM for bond energyes E_{bond}, bulk moduli B, and shear coefficients $C = C_{11} - C_{12}$ and C_{44}. All energies are in eV and elastic constants in 10^{11} dynbes/cm^2.

$1/d^2$ and $1/d^4$ scaling rules for $V_{\alpha\alpha'}$ and u, respectively, we assume the following forms:

$$V_{\alpha\alpha'} = V^0_{\alpha\alpha'}(d_0/d)^n \tag{120}$$

and

$$u = u_0(d_0/d)^m, \tag{121}$$

where the superscript and subscript 0 indicate the values evaluated at the equilibrium bond length d_0. For simplicity, the values of $V^{(0)}_{\alpha\alpha'}$ are taken to be Harrison's values given in Eq. (106) scaled by a factor f:

$$V^{(0)}_{\alpha\alpha'} = fV^{\text{Harr}}_{\alpha\alpha'}. \tag{122}$$

Thus, there are four parameters for each system: the scaling parameter f, the powers n and m, and the value u_0. These parameters can be determined by

requiring that the model produce the correct experimental values for E_{bond}, d_0, $C_{11} - C_{12}$ and B. Since $C_{11} - C_{12}$ is only governed by $V_{aa}^{(0)}$ in both BOM and band calculation, it alone determines the scaling factor f. Then the bond energy E_{bond} can be used to determine u_0. The requirement that the first derivative of E_T be zero at d_0 then determines the ratio of the powers n/m, which couples with the equation for the bulk modulus to yield the values for n and m. One can then use these sets of parameters to check the validity of the model by calculating other quantities not employed in the fitting, e.g., C_{44}, the internal displacement parameter ζ, and the optical phonon frequencies ω at the zone center. If the results are acceptable, the model can be extended to more complicated systems such as alloys and superlattices with local environments similar to the bulk crystals.

Table VII shows the results for f, n, m, and u_0 obtained from the preceding fitting procedure by using the full band-structure calculations, and the corresponding values of C_{44}, ζ, and ω calculated for consistency checks. The scaling factor f ranges from 1 to 1.4 and tends to decrease with an increase in

TABLE VII

FULL BAND STRUCTURE CALCULATION[a]

	f	n	m	u_0	C_{44}	ζ	ω
C	1.390	2.840	3.767	21.924	48.393	0.121	1,459
Si	1.326	3.040	5.001	6.938	8.013	0.511	572
Ge	1.388	3.204	5.278	6.415	6.841	0.487	342
AlP	1.294	3.530	5.598	6.435	5.827	0.516	447
AlAs	1.464	3.524	5.430	7.089	5.598	0.459	384
AlSb	1.337	3.268	5.668	4.838	3.944	0.564	354
GaP	1.395	3.705	5.683	7.285	6.857	0.501	382
GaAs	1.397	3.633	5.716	6.530	5.791	0.500	292
GaSb	1.431	3.471	5.717	5.519	4.515	0.536	256
InP	1.323	4.240	6.633	5.603	4.260	0.584	304
InAs	1.300	3.997	6.427	4.962	3.564	0.552	220
InSb	1.353	3.773	6.399	4.350	3.092	0.602	200
ZnS	1.062	3.308	5.996	4.225	3.727	0.632	325
ZnSe	1.134	3.420	5.994	4.260	3.164	0.576	233
ZnTe	1.284	3.306	5.828	4.285	2.813	0.590	205
CdTe	1.171	3.656	6.761	3.092	1.701	0.694	156
HgTe	1.173	3.[illegible]	7.074	3.080	2.040	0.716	152

[a]The results for the parameters f, n, m, and u_0 obtained from the fitting of the bond energy, bond length, bulk modulus, and shear coefficient $C_{11} - C_{12}$ of Table V using the full band structure calculation. Also listed are the calculated C_{44}, internal displacement parameter ζ, and the TO optical phonon mode ω at Γ. All the elastic constants are in units of 10^{11} dynes/cm^2, u_0 is in eV, and ω are wave numbers in 1/cm.

polarity. In the power dependence of $V_{\alpha\alpha'} \propto (d_0/d)^n$, n ranges from 2.8 to 4.3, which is larger than the $n = 2$ used in Harrison's universal TB parameters. For the repulsive pair energy $u = u_0(d_0/d)^m$, the power m ranges from 3.8 to 6.8. The ratio m/n falls in the range from 1.3 to 1.9, which is smaller than the $m/n = 2$ used by Harrison. The calculated values of C_{44} for most systems agree with the experimental data to 10% or better, except for diamond and ZnS. Note that the experimental data for ZnS are rather dispersed. The calculated TO optical phonon modes at Γ in 1/cm for most group IV and III-V systems also agree with experiments to 10% or better. The discrepancies for the II-VI systems are larger, about 15%. Reliable results for ζ from both experiments and first-principles calculations are available only for a limited number of systems. The calculated ζ in the TB model agrees very well with those results, as shown in Table I. The overall results for C_{44}, ζ, and ω in the TB calculations are equivalent to those based on the valence force model, including Martin's Coulomb force corrections. By construction, the TB model also produces the correct cohesive energies, bond lengths, bulk moduli and shear coefficient $C_{11} - C_{12}$, because these quantities are used to fit the parameters.

The results in Table VII are based on the term values given in Table IV and the TB parameters scaled from Harrison's universal parameters. It is useful to know how the predictions are influenced by these parameters and the fitting procedure. Table VIII shows the results based on Chadi procedure in which the TB matrix elements $V_{\alpha\alpha'}$ are scaled as $1/d^2$, and the repulsive pair energy is taken to be $u = u_0 + u_1(d - d_0) + u_2(d - d_0)^2$. The parameter u_0 is set to produce the correct bond energy, u_1 is determined by requiring the correct equilibrium bond length, and u_2 is fixed by the bulk modulus. Two sets of TB parameters are tabulated for each system: One is the set used by Chadi (1978, 1979, and 1984), and the other is the set obtained by multiplying Harrison's $V_{\alpha\alpha'}$ by scaling factor f listed in Table VIII. For convenient comparison, the zero of the term values is set equal to the anion s energy. Despite considerable differences in these two sets of TB parameters, the results of the predictions from both sets are very similar and also very similar to those predicted from the other procedure used for the results in Table VII. The only noticeable difference between the predictions in Table VIII and Table VII is that the present procedure produces larger phonon frequencies and slightly smaller C_{44} values. We also note that the fitted parameters u_0, u_1, and u_2 and the predicted values for Chadi's elastic constant set in Table VIII are not the same as Chadi's published (1979) values; these give bulk moduli about 20% smaller than the experimental values, but give considerably better phonon frequencies.

To parametrize the BOM, several different stages of approximations can be made. However, for a general application, the full BOM steps presented in

TABLE VIII

TWO SETS OF TB PARAMETERS.[a]

	SI							
	ε_s^A	ε_p^A	ε_s^C	ε_p^C	$V_{ss\sigma}$	$V_{sp\sigma}$	$V_{pp\sigma}$	$V_{pp\pi}$
Chadi	0.0	7.20	0.0	7.20	−2.03	2.55	4.55	−1.09
Present	0.0	6.88	0.0	6.88	−2.41	2.59	4.05	−1.15
	u_0	u_1	u_2	C	C_{44}	$C_{44}^{(0)}$	ζ	ω
Chadi	7.29	−9.98	23.90	10.66	7.89	11.38	0.49	620
Present	6.93	−9.70	23.42	10.27	7.83	11.39	0.51	592

	GaAs								
	ε_s^A	ε_p^A	ε_s^C	ε_p^C	$V_{ss\sigma}$	$V_{sp\sigma}^{AC}$	$V_{pp\sigma}^{CA}$	$V_{pp\pi}$	$V_{pp\pi}$
Chadi	0.0	9.64	5.12	11.56	−1.70	2.40	1.90	3.44	−0.89
Present	0.0	10.09	6.79	14.12	−2.34	2.52	2.52	3.94	−1.12
	u_0	u_1	u_2	C	C_{44}	$C_{44}^{(0)}$	ζ	ω	
Chadi	5.12	−7.12	18.22	6.36	5.60	8.77	0.54	339	
Present	6.53	−8.39	19.90	6.63	5.70	8.53	0.54	322	

[a]Comparison between the two different sets of TB parameters described in the text, the resultant expression coefficients u_0, u_1, u_2 of the repulsive pair energy u, and the predicted elastic constants, Kleinmann internal displacement parameters ζ, and phonon frequency ω from Chadi fitting scheme.

Eqs. (100) to (105) for calculating E_b should be followed, regardless of approximations. The simplest model, referred to as BOM(1), is to include only V_{1A} and V_{1C} in the matrix element $\langle b|H|a'\rangle$ for the calculation of the metallization energy in Eq. (104). The next approximation, BOM(2), is to include V_1^x as well. Finally, one can include all the first-neighbor interatomic TB parameters in $\langle b|H|a'\rangle$; this will extend the $|a'\rangle$ to those belonging to the second-neighbor bonds. This last approximation will be referred to as BOM(3).

Table IX shows the results for BOM(1) following the parameterization procedure described in Eqs. (120) to (122). The fittted parameters f, n, m, and u_0 are substantially different from those based on the BS calculations. The predicted C_{44} for the group IV and III-V systems are slightly larger than the experimental values, but good for the II-VI systems. The calculated ζ values, although not all smaller than those in Table VII, are smaller on the average. The predicted phonon frequencies are too high.

The parameters and the predicted results from BOM(2) are listed in

TABLE IX

BOM(1) CALCULATIONS[a]

	f	n	m	u_0	C_{44}	ζ	ω
C	1.440	2.896	3.809	22.345	61.098	0.142	2,103
Si	1.356	3.208	5.166	7.023	9.729	0.447	695
Ge	1.395	3.666	5.854	6.383	8.240	0.506	416
AlP	1.047	3.334	5.842	4.418	6.627	0.460	571
AlAs	1.264	3.530	5.685	5.538	6.668	0.431	491
AlSb	1.179	3.170	5.831	3.818	4.482	0.511	411
GaP	1.154	3.439	5.744	5.239	7.500	0.453	496
GaAs	1.179	3.337	5.647	4.863	6.340	0.456	358
GaSb	1.274	3.354	5.797	4.456	4.950	0.500	302
InP	0.999	3.579	6.490	3.413	4.558	0.522	388
InAs	0.995	3.167	5.960	3.081	3.787	0.496	260
InSb	1.126	3.228	6.105	3.091	3.267	0.543	224
ZnS	0.692	2.750	7.262	1.642	4.312	0.580	401
ZnSe	0.750	2.823	6.839	1.821	3.685	0.544	287
ZnTe	0.888	2.715	6.484	1.987	3.156	0.531	248
CdTe	0.734	2.720	8.394	1.018	1.949	0.660	189
HgTe	0.732	2.202	8.810	0.842	2.140	0.677	177

[a]The results for the parameters f, n, m, and u_0 obtained from the fitting of the bond energy, bond length, bulk modulus, and shear coefficient $C_{11} - C_{12}$ of Table V using the BOM(1) described in the text. Also listed are the calculated C_{44}, internal displacement parameter ζ, and the TO optical phonon mode ω at Γ. All the elastic constants are in units of 10^{11} dynes/cm^2, u_0 is in eV, and ω is given in terms of wave number in 1/cm.

Table X. These parameters more closely resemble those in Table VIII than do the values from BOM(1). However, the predicted C_{44} values are still too small, and the ζ values are too large, but the ω values are better than those from BOM(1).

Table XI contains the results from BOM(3); these approach those of the BS calculations. In comparison with Table VII and the experimental values in Table V, the BOM(3) does well for ω, produces slightly smaller C_{44}, and probably slightly larger ζ values.

In conclusion, the TB method is a reasonable approach to the static elastic properties of semiconductors. If carried out rigorously, the TB parameters in Table VII will provide quantitative results for superlattices, alloys, and possibly surfaces in which the local environments are similar to those in the bulk. The quantitative predications of BOM are not as good as the BS calculations, but are still reasonable. The fitted parameters given in Tables IX to XI allow different stages of approximations to be made using the BOM. This is especially useful for more complicated systems, because computationally the BOM is about two orders faster than the band structure calculations.

TABLE X

BOM(2) CALCULATIONS*

	f	n	m	u_0	C_{44}	ζ	ω_{TO}
C	1.440	2.896	3.809	22.345	55.161	0.319	1672
Si	1.356	3.208	5.166	7.023	7.520	0.652	597
Ge	1.395	3.666	5.854	6.383	6.106	0.720	358
AlP	1.283	3.440	5.493	6.171	5.094	0.677	511
AlAs	1.472	3.543	5.419	6.979	5.103	0.640	440
AlSb	1.342	3.249	5.580	4.771	3.256	0.706	373
GaP	1.371	3.540	5.490	6.901	5.916	0.688	428
GaAs	1.387	3.455	5.452	6.306	4.864	0.681	317
GaSb	1.414	3.398	5.593	5.305	3.681	0.717	264
InP	1.307	3.846	6.115	5.339	3.446	0.744	354
InAs	1.298	3.528	5.731	4.834	2.806	0.706	243
InSb	1.342	3.464	5.910	4.204	2.351	0.747	207
ZnS	1.062	3.102	5.728	4.075	3.016	0.726	373
ZnSe	1.141	3.150	5.587	4.181	2.577	0.688	271
ZnTe	1.283	3.067	5.498	4.140	2.270	0.704	231
CdTe	1.178	3.271	6.216	3.011	1.319	0.764	177
HgTe	1.169	3.046	6.116	2.916	1.457	0.793	162

*The results for the parameters f, n, m, and u_0 obtained from the fitting of the bond energy, bond length, bulk modulus, and shear coefficient $C_{11} - C_{12}$ of Table V using the BOM(2) described in the text. Also listed are the calculated C_{44}, internal displacement parameter ζ, and the TO optical phonon model ω at Γ. All the elastic constants are in units of 10^{11} dynes/cm^2, u_0 is in eV, and ω is wave number in 1/cm.

11. SEMICONDUCTOR ALLOYS

The systems to be considered in this section are alloys of the diamond and zincblende semiconductors, both the ordered and disordered alloys. The ordered alloys include binary compounds such as SiC, and ternary compounds, such as $GaInAs_2$, in three crystal structures of the types CuAuI, chalcopyrite, and CuPt as shown in Fig. 4 (Bernard *et al.*, 1988). The disordered alloys include binary solutions such as $Si_{1-x}Ge_x$ and pseudobinaries such as $Hg_{1-x}Cd_xTe$ and $GaAs_{1-x}Sb_x$, where x is the fractional concentration. These alloys have been widely used and studied; however, detailed information about their elastic constants is scarce both experimentally and theoretically. One reason for the lack of rigorous calculation is that the elastic constants of these systems are more complex; existing theories are not as accurate, particularly for disordered alloys. Another reason may be attributed to the fact that most properties of these alloys, including their elasticity, were thought to be reasonably well approximated by the concen-

TABLE XI

BOM(3) CALCULATIONS[a]

	f	n	m	u_0	C_{44}	ζ	ω
C	1.336	2.901	3.872	20.814	47.691	0.135	1,531
Si	1.262	3.251	5.346	6.484	7.472	0.580	562
Ge	1.302	3.886	6.264	5.942	6.171	0.658	333
AlP	1.193	3.43	5.77	5.722	4.936	0.541	452
AlAs	1.369	3.559	5.507	6.498	4.812	0.490	387
AlSb	1.249	3.253	5.669	4.402	3.248	0.619	343
GaP	1.276	3.550	5.575	6.422	5.738	0.542	373
GaAs	1.290	3.445	5.511	5.855	4.724	0.549	281
GaSb	1.318	3.415	5.704	4.923	3.670	0.621	239
InP	1.210	3.812	6.153	4.929	3.398	0.604	304
InAs	1.199	3.456	5.724	4.442	2.755	0.579	216
InSb	1.245	3.411	5.943	3.871	2.367	0.657	188
ZnS	0.968	3.062	5.848	3.618	3.120	0.630	343
ZnSe	1.040	3.097	5.671	3.727	2.611	0.580	247
ZnTe	1.178	3.019	5.573	3.371	2.301	0.601	212
CdTe	1.065	3.198	6.368	2.621	1.403	0.690	165
HgTe	1.059	2.937	6.262	2.523	1.563	0.730	154

[a]*The results for the parameters* f, n, m, and u_0 obtained from the fitting of the bond energy, bond length, bulk modulus, and shear coefficient $C_{11} - C_{12}$ of Table V using the BOM(3) described in the text. Also listed are the calculated C_{44}, internal displacement parameter ζ, and the TO optical phonon mode ω at Γ. All the elastic constants are in units of 10^{11} dynes/cm^2, u_0 is in eV, and ω is wave number in 1/cm.

tration-weighted averages of their constituents. Because of the rudimentary state of the theory, we shall deal only with the simplest elastic constant, the bulk modulus. Our focus is on the difference between the alloy bulk modulus B and the concentration-weighted averaged value $\bar{B}$, i.e., $\Delta B = B - \bar{B}$. There are several fundamental questions that can be addressed. Is ΔB positive or negative? Do the sign and the magnitude of ΔB depend on the state of order? How can we calculate the bulk modulus in a disordered system? Analysis of these questions constitutes the content of this section.

a. Ordered Alloys

Long-range ordering has been found to exist in many epitaxially grown III-V semiconductor alloys (Kuan *et al.*, 1985; Jen *et al.*, 1986; Kuan *et al.*, 1987; Ihm *et al.*, 1987; Klem *et al.*, 1987; Huang *et al.*, 1988; Gomyo *et al.*, 1987 and 1988; Norman *et al.*, 1987; Shahid *et al.*, 1987). Wei and Zunger (1989) have given a rather complete list of these ordered alloys, most of which

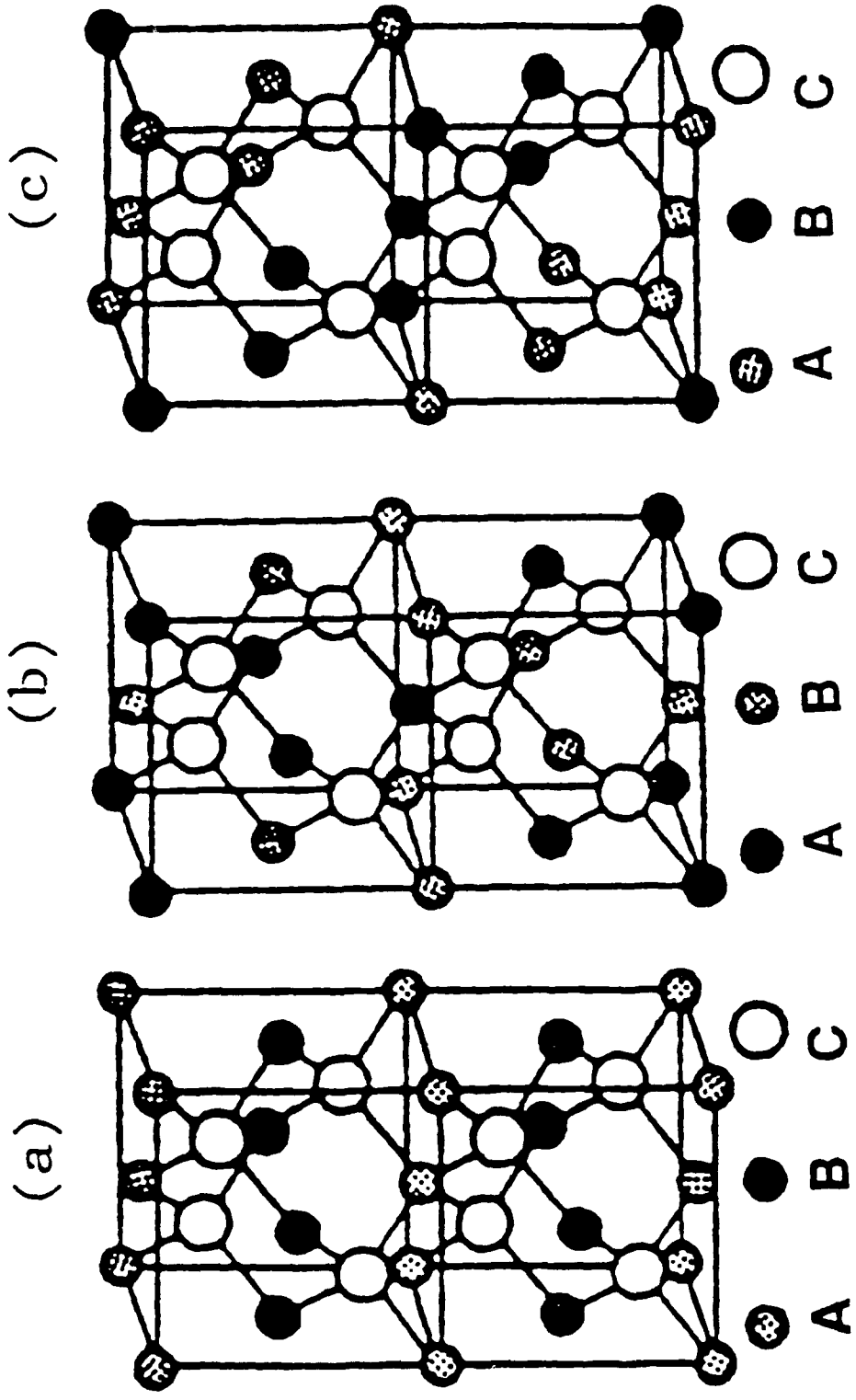


FIG. 4. Three ABC_2 structures studied: (a) CuAuI structure ordered in [001] direction. (b) Chalcopyrite structure ordered in [201] direction. (c) CuPt structure ordered in [111] direction. (Extracted from Wei and Zunger, 1989.)

are of the form ABC_2 existing in three different types of structures, CuAuI, chalcopyrite, and CuPt; all have their atomic planes stacked as ACBCACBC, but along three different directions, (100), (201), and (111), respectively, as shown in Fig. 4. There have been several first-principles calculations made to study the structural properties of these compounds; these focused mainly on the cohesive energies and bond lengths. These results have been compared with TB calculations by Yeh *et al.* (1990). As indicated in the preceding section, one virtue to fitting the TB model for the bulk semiconductors is to use it for interpolating alloy properties. Table XII lists the results for the bulk moduli of a number of III-V and II-VI alloys derived from full TB band structure calculations using the parameters given in Table VII. Also listed are the average values and percentage deviations from the mean $\Delta B/\bar{B}$. Note that all ΔB values are negative and that most of the magnitudes are small, except for Ga_2AsSb and Ga_2PSb; the latter has the largest difference in the constituent compounds among the alloys listed. Although the magnitudes of ΔB get larger for systems with larger differences in the bond lengths, the dependence does not seem to be a simple function of the bond length difference. The uniformly negative ΔB values also appeared in the first-principles local-density functional calculations for ordered GaAsSb alloys by Ferreira *et al.* (1989), as shown in Table XIII.

TABLE XII

ORDERED ALLOYS: TB ELASTIC CONSTANT CALCULATIONS[a]

					$\Delta B/\bar{B} \times 100$		
Alloy	B_{Ch}	B_{Ca}	ZB_{Cp}	$\bar{B}$	Ch	Ca	Cp
AlGaAs	7.695	7.693	7.689	7.009	−0.18	−0.20	−0.25
AlGaP	8.858	8.858	8.854	8.872	−0.15	−0.15	−0.20
GaInSb	5.226	5.202	5.156	5.312	−1.61	−2.07	−2.92
AlInAs	6.705	6.691	6.661	6.761	−0.83	−1.03	−1.48
InGaAS	6.610	6.579	6.508	6.742	−1.96	−2.42	−3.47
InAlP	7.876	7.860	7.774	7.924	−0.61	−0.08	−1.88
GaInP	8.007	8.035	8.878	8.195	−1.44	−1.95	−3.87
GaAsP	8.328	8.291	8.294	8.417	−1.05	−1.50	−1.46
GaAsSb	6.314	6.198	6.157	6.741	−6.34	−8.05	−8.66
GaPSb	6.584	6.297	6.188	7.468	−11.84	−15.68	−17.14
HgCdTe	4.470	4.472	4.471	4.485	−0.33	−0.28	−0.31
HgZnTe	4.890	4.887	4.632	4.925	−0.71	−0.76	−5.93
CdZnTe	4.611	4.604	4.338	4.650	−0.85	−1.00	−6.72

[a]Bulk moduli (in 10^{11} dynes/cm^2) of ordered alloys calculated using the full TB band-structure method described in Section 10. The three structures are chachopyrite (Ch), CuAul (Ca), and CuPt(Cp) types. $\bar{B}$ is the average value of the constituent compounds, and $\Delta B = B - \bar{B}$.

The reason for the negative values of ΔB, in a very qualitative argument, is that the bulk moduli of semiconductors scale inversely as high powers of the lattice constant, and, at the same time, the alloy lattice constant is approximated well by the mean value, by Vegard's law (1921). This implies that the value of B at the mean lattice constant should lie below the straight-line average. Since the TB results for the bulk moduli constants should not be qualitatively different from the valence force-field model (VFF) predictions, most of the key physics for the bowing of B should be contained in a VFF analysis. The major effects in the VFF can in turn be realized from the following simple analysis.

Consider the local structure of a CuAuI or chalcopyrite crystal ABC_2. Focus on a local tetrahedral cluster A_2B_2C with the two A atoms and two B atoms on the vertices of the tetrahedron and the C atom near the center. Let the coordinates of the two A atoms be $(-1, 1, -1)d/\sqrt{3}$ and $(-1, 1, -1)d/\sqrt{3}$ and the two B atoms be at $(1, 1, 1)d/\sqrt{3}$ and $(1, -1, -1)d/\sqrt{3}$. Let the force constants be k_A and k_B, and equilibrium bond lengths be d_A and d_B for the AC and BC bonds, respectively. To attain equilibrium, the central C atom is displaced by $(\varepsilon, 0, 0)d/\sqrt{3}$. We further define mean values $\bar{d} = (d_A + d_B)/2$ and $\bar{k} = (k_A + k_B)/2$, relative differences $\delta_0 = (d_A - d_B)/\bar{d}$ and $\Delta_0 = (k_A - k_B)/\bar{k}$, and $d = \bar{d}(1 + \delta)$. Then the AC bond is stretched by an amount $\bar{d}(\delta + \varepsilon/3 - \delta_0/2)$ from its equilibrium value, and similarly, the BC bond is compressed by $\bar{d}(\delta - \varepsilon/3 + \delta_0/2)$. The strain energy for any arbitrary δ and ε is given by $\Delta E = \bar{d}^2[k_A(\delta + \varepsilon/3 - \delta_0/2)^2 + k_B(\delta - \varepsilon/3 + \delta_0/2)^2]$. When ΔE is minimized with respect to δ and ε, one finds $\delta = 0$ and $\varepsilon = 3\delta_0/2$, and the minimum ΔE is zero. If the crystal expands uniformly with δ having a fixed small value, then ε becomes $\varepsilon = 3\delta_0/2 - 3\Delta_0\delta$ and $\Delta E = 2\bar{k}(1 - \Delta_0^2/4)\bar{d}^2\delta^2$. Thus, the effective spring constant is

$$k_{\text{eff}} = \bar{k}(1 - \Delta_0^2/4), \tag{123}$$

which is smaller than the average value $\bar{k}$. This weakening of the restoring force constant in the alloy is due to the internal displacement, represented by ε in the preceding model, which provides an extra degree of freedom for relaxation in response to the external stress. The bulk modulus B is proportional to k_{eff}/d, so the alloy bulk modulus minus the mean B is then given by

$$\Delta B = \bar{B}(\delta_0\Delta_0 - \delta_0^2 - \Delta_0^2/4), \tag{124}$$

where we recall the definitions $\delta_0 = (d_1 - d_2)/\bar{d}$ and $\Delta_0 = (k_1 - k_2)/\bar{k}$. Since and δ_0 and Δ_0 tend to have different signs, the bond length difference gives an

extra negative contribution to ΔB (the first two terms). If both the bond-stretching force constant α and the bond-angle restoring force β in the VFF are included, the equilibrium value of ΔE is no longer zero, but the deviation ΔB can also be shown to be similar to Eq. (123) and is given by

$$\Delta B = \bar{B}\left[\delta_0\left(\frac{3\Delta\alpha + \Delta\beta}{3\alpha + \beta}\right) - \delta_0^2 - 3\frac{(\Delta\alpha)^2}{(\alpha + 2\beta)(3\alpha + \beta)}\right]\Big/4, \qquad (125)$$

where $\Delta\alpha = (\alpha_1 - \alpha_2)$ and $\alpha = (\alpha_1 + \alpha_2)/2$ and similarly for $\Delta\beta$ and β. Equation (125) reduces to (124), if β is set equal to zero.

The above descriptions illustrate two mechanisms for the negative ΔB values, the $1/a^q$ scaling of B, with q ranging from 3.5 to 9, and more degrees of freedom for internal relaxation. Quantitative results should be described by TB, because in addition to the strain energy, there is also some chemical effect built into the TB theory. Although these ordered compounds have been found from epitaxial growth, the bulk moduli are probably difficult to measure, because these alloys are not single bulk crystals and because the ordering is only partial. It is interesting to note that the B values for a SiC/AlN alternating layer superlattice along (100) and for the constituent compounds have been calculated by Lambrecht and Segall (1990) using the LMTO; their percentage deviation from the mean $\Delta B/\bar{B}$ was found to be about -2%, which falls in the range of the ternary alloys in Tables XII and XIII.

Not all the mechanisms considered above apply to the ordered compounds of the elemental semiconductors, because internal relaxation under pressure may not be allowed, e.g., if the structure is assumed to be zincblende. Unfortunately, a simple analysis of the elastic constants of the 4-4 compounds cannot yet be made, because the tight-binding and VFF parameters have not yet been extended to deal with the atomic pairs not existing in the

TABLE XIII

ORDERED ALLOYS: BULK MODULI[a]

	GaAs	GaSb	Ga_2AsSb			Ga_4As_3Sb		Ga_4AsSb_3	
Structure	zb	zb	CA	CH	CP	LU	FA	LU	FA
B	7.46	5.18	6.10	5.92	5.96	6.52	6.58	5.40	5.31
$\Delta B/\bar{B} \times 100$			−3.5	−6.3	−5.7	−5.4	−4.5	−6.1	−7.7

[a]Calculated bulk moduli for GaAs, GaSb, and GaAsSb ordered alloys by Ferreira *et al.* (1989), and the corresponding percentage deviation from the concentration-weighted average. The structures are zincblende (zb), CuAl (CA), chachopyrite (CH), CuPt (CP), luzonite (LU), and famatinite (FA).

constituent crystals. However, several first-principles calculations have been made on the ordered SiC and SiGe (Martins and Zunger, 1986; Qteish and Resta, 1988; van Schilfgaarde, 1990). The main results are listed in Table XIII; the theoretical results were calculated for the zincblende structure. The plane-wave pseudopotential (PP-PW) calculation of Martins and Zunger (1986) for SiC gave a -21% value for $\Delta B/\bar{B}$, which is in reasonable agreement with the experimental value of -17%. This is consistent with the qualitative argument based on the $1/d^q$ ($q \geq 3.5$) scaling of B. The theory also yielded a negative formation energy, which also agrees with experiment. For SiGe, the theoretical calculations cited in Table XIV gave positive formation energies, which are consistent with the fact that no ordered bulk compounds of SiGe have been grown. However, some weak ordering has been found in the epitaxial SiGe films (Ourmazd and Bean, 1985). The calculated values of $\Delta B/\bar{B}$ for the zincblende SiGe are either slightly above or just below zero. These differences, however, fall within the uncertainties of the present first-principles theory. The best conclusion that can be drawn from these results is that the B value for SiGe should be very close to the mean value.

b. Disordered Alloys

Disordered binary alloys $A_{1-x}B_x$ of diamond semiconductors and pseudobinary alloys $A_{1-x}B_xC$ of zincblende semiconductors AC and BC are considered in this section; they are not amorphous materials, as they still

TABLE XIV

BULK MODULI OF ORDERED BINARY ALLOYS AB OF THE DIAMOND SEMICONDUCTORS A AND B FROM THEORIES AND EXPERIMENT

SiC	PP-PW[a]	Experiment		
B(C)	50.3	44.23		
B(Si)	9.53	9.92		
B(SiC)	23.4	22.4		
$\Delta B/\bar{B}$(%)	−21	−17		
SiGe	**PP-PW[a]**	**PP-PW[b]**	**ASA[c]**	**FP-LMTO[c]**
B(Si)	9.53	9.8	8.80	9.58
B(Ge)	7.75	7.7	6.25	7.05
B(SiGe)	8.73	8.7	7.38	8.31
$\Delta B/\bar{B}$(%)	1	0	2	0

[a]Martins and Zunger (1986).
[b]Qteish and Resta (1988).
[c]van Schilfgaarde (1990).

possess their constituent diamond and zincblende lattices, respectively, as characterized by their crystal diffraction patterns. We shall consider the pseudobinaries first. The alloying atoms A and B in these alloys belong to a fcc sublattice, and the C atoms to the other sublattice. However, the positions of the A and B atoms are not necessarily locked precisely on the lattice sites. The extended x-ray absorption fine-structure spectroscopy (EXAFS) data (Mikkelsen and Boyce, 1982 and 1983; Boyce and Mikkelsen, 1985; Balzarotti *et al.*, 1985) have consistently shown a bimodal distribution of the bond lengths in these alloys, although the average lattice constant follows the Vegard (1921) law $a = (1 - x)a_{AC} + xa_{BC}$. Figure 5 shows an example of the results for the bond lengths in $Ga_{1-x}In_xAs$ deduced by Mikkelsen and Boyce

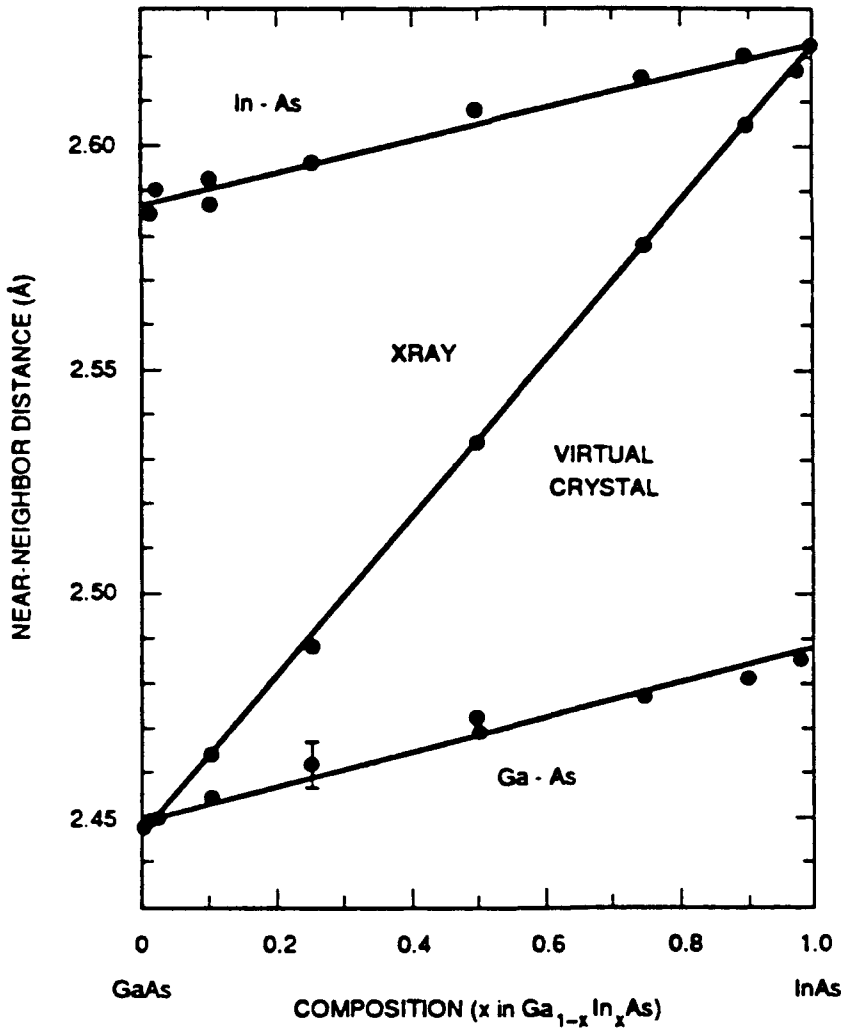


FIG. 5. Near-neighbor bond lengths (GaAs and InAs in the $Ga_{1-x}In_xAs$ alloy) as a function of composition x, measured by EXAFS (Mikkelsen and Boyce, 1982).

(1982) from their EXAFS data. In a first approximation, the crystal structure of an alloy can be viewed as having the A and B atoms on their fcc sublattice with an average lattice constant, while the C atoms are distorted away from their lattice sites, in a way similar to the local structures of the three ordered superlattices considered earlier. The difference is that there is no long-range superlattice ordering of the A and B atoms in the disordered state. This simple crystal picture is only a first approximation; the EXAFS experiments just cited also indicate that the sublattice of the A and B atoms is less than a perfect fcc. There are also theoretical calculations (e.g., Sher *et al.*, 1987; Wei *et al.*, 1990) that suggest a certain degree of short-range ordering in these alloys, namely that the arrangement of the A and B atoms is not random. Figure 6 shows an example of the calculated deviations (Sher *et al.*, 1987) of the probabilities from random distribution, $\Delta p_n = p_n - p_n^0$ for $Ga_{1-x}In_xAs$ as a function of alloy concentration x, where p_n is the probability of having n Ga atoms and 4-n In atoms on the vertices of a local tetrahedral cluster in the alloy, and $p_n^0 = {}_4C_n(1-x)^n x^{4-n}$, where ${}_4C_n$ is a binomial coefficient, is the corresponding value for the random distribution.

The structural energy needed for calculating the elastic constants of a disordered alloy is an ensemble average of the total energy over the distribution of the alloying atoms under strains. It has been demonstrated (Ferreira *et al.*, 1989) that the total energy of a semiconductor can be decomposed into the sum of multisite correlation energies, from the single-site, the pair, and up to a cluster containing a handful of sites. In other words, the multisite correlation energies converge to zero quickly at a manageable number of sites. This implies that the structural energy of an alloy is an average of these multisite correlation energies. Connolly and Williams (1983), working on metal alloys, proposed that these multisite correlation energies be deduced from the ordered systems that are composed of the same atoms. This scheme allows a direct application of the first-principles theory in the calculation of the energy parameters. These energetics can then be used in the alloy statistics such as in the cluster variational method of Kikuchi (1951) or in the Monte Carlo calculations to deduce the distribution functions or the average properties. This theory has been carried out extensively for semiconductor alloys by Ferreira *et al.* (1989), and respectable results have been obtained for the phase diagrams and alloy equilibrium properties. For this theory to fit the elastic constants requires detailed dependences of the multisite energies under different strains that have yet to be worked out. Also, the validity of using the energy parameters deduced from ordered alloys in the disordered systems needs to be examined further.

There is a different cluster approach, which directly relates the alloy Hamiltonian to the distribution function (Gautier *et al.*, 1975; Chen *et al.*, 1987; Berera *et al.*, 1988; Dreysse *et al.*, 1989). In this approach one focuses

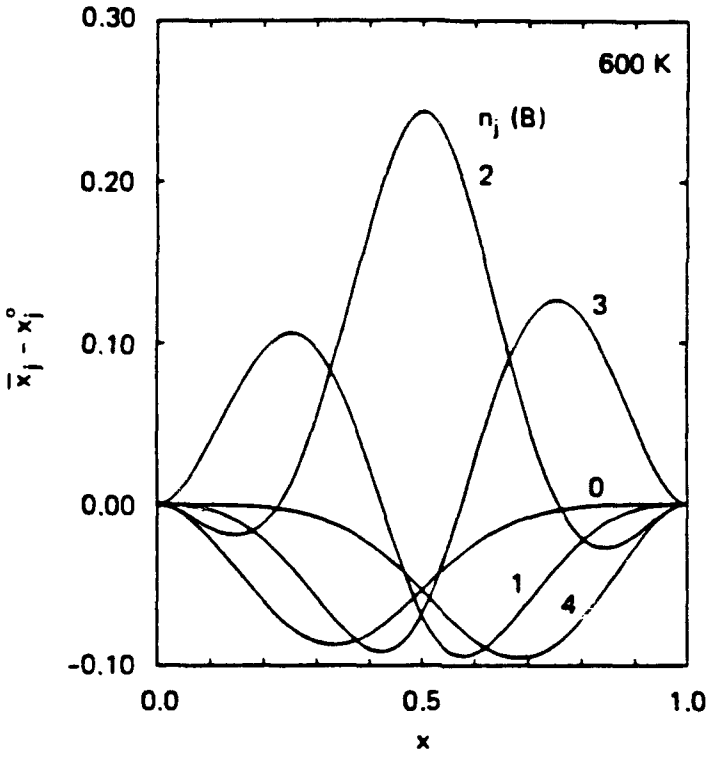


FIG. 6. Cluster populations relative to those in a random alloy $x_j - \bar{x}_j$ for clusters with $n_j = 0, 1, 2, 3, 4$ B atoms for a $Ga_{1-x}In_xAs$ alloy equilibrated at 600 K.

a particular cluster in an alloy ensemble. The average energy per cluster can be written (Chen *et al.*, 1988) as

$$\langle \varepsilon \rangle = \sum_n \sum_m \left(\varepsilon_n + \frac{1}{2} h_{nm} \right) p_n P_{nm}. \tag{126}$$

on where ε_n is the energy of a cluster detached from a given alloy configuration, and h_{nm} is the interaction energy, which is the change of energy of the combined system when the cluster is put back into the alloy. In Eq. (126), p_n is the probability that a cluster is of the type n, specified by the number of A and B atoms and their arrangements, and P_{nm} is a conditional probability that the surrounding environment is in state m when the cluster is in state n. The factor 1/2 in Eq. (126) is to eliminate the double counting in the total average alloy energy $\langle E \rangle = M \langle \varepsilon \rangle$, where M is the ratio between the size of the alloy and the cluster. If one writes $\langle \varepsilon \rangle = \Sigma\, p_n \varepsilon(n)$, then an effective cluster energy can be defined as

$$\varepsilon(n) = \sum_m \left(\varepsilon_n + \frac{1}{2} h_{nm} \right) P_{nm}. \tag{127}$$

This procedure is particularly useful when the interaction energy is short-ranged. One can start with a small cluster and a given probability distribution, then calculate the effective cluster energies from Eq. (126). These energies are then used in a statistical theory to deduce a cluster distribution, which in turn is used to calculate a new set of $\varepsilon(n)$ and distribution functions, and the process is iterated until it converges. It should be pointed out that to compute the total energy of an alloy quantum mechanically, one needs to solve the Schroedinger equation for a Hamiltonian that does not have the lattice translational symmetry so indispensable in traditional band-structure theory. If the fluctuation of the alloy potential from the virtual crystal approximation (VCA), where the alloy potential is approximated as the concentration-weighted average, is small, then the next leading correction to VCA can be obtained from perturbation theory. This should work for most semiconductor alloys except for systems with large potential fluctuations such as $Hg_{1-x}Cd_xTe$ (Chen and Sher, 1982; Spicer *et al.*, 1982; Hass *et al.*, 1983). A more general but more difficult approach is to extend the present molecular coherent potential approximation (MCPA) (Hass *et al.*, 1983) to clusters and to achieve a triple self-consistency (Chen *et al.*, 1987): consistency between cluster distribution and Hamiltonian, between the Hamiltonian and electron density, and between the self-energy operator Σ in the cluster CPA theory and the potential fluctuations. To date this theory has been carried out only for metal alloys, and then only within the single-site KKR-CPA with a random distribution (Schwartz and Bansil, 1975; Gyorffy and Stocks, 1978). Major work is needed to determine if this approach can achieve the same degree of rigor for disordered alloys as self-consistent density functional theory, which has been successfully used in dealing with crystalline semiconductors.

The above idea has been applied to an elastic medium model to deduce a mean field theory for the internal strain and bulk modulus in semiconductor alloys (Chen *et al.*, 1988). This theory starts by assuming that the alloy has an effective lattice constant and effective modulus. When part of the effective alloy medium is replaced by a specified cluster, there will be strain energy introduced. It was shown that this strain energy can be taken as the effective energy $\varepsilon(n)$ for that cluster, the probability distribution p_n within a statistical theory can then be deduced. The internal strain energy is calculated as $E = M\langle\varepsilon\rangle = M\,\Sigma\, p_n\varepsilon(n)$. When the alloy is under an external pressure δP, the effective cluster energy will change by an amount $\delta\varepsilon(n)$, which implies a change of the total energy by an amount $\delta E = M\langle\delta\varepsilon(n)\rangle$. Then the bulk modulus of the alloy can be obtained from $\Delta E = \frac{1}{2}(\delta P)^2 V/B$, where V is the alloy volume. The mean-field nature of this approach is evident from the fact that the calculation requires knowledge of the alloy lattice constant and

elastic constants that are only assumed and are required to be calculated self-consistently. To illustrate this self-consistency procedure, let us consider the following simple spring model for a random pseudobinary alloy $A_{1-x}B_xC$. The cluster corresponds to the four bonds surrounding an "impurity" atom A or B, and the environment of the cluster corresponds to the 12 bonds that connect inwardly to the cluster and outwardly to a rigid lattice of the effective alloy. It is worth mentioning that there have been detailed analyses of the valence-field force models for the strain energies of semiconductor alloys in regard to the range allowed for lattice relaxation (Martins and Zunger, 1984; Chen and Sher, 1985). It was found that by neglecting the bond-angle forces, one can use a shorter range of lattice relaxation to obtain the correct mixing enthalpies and bond lengths. The simple model considered here works amazingly well for these properties.

Let the spring constants for the pure AC and BC compounds be k_A and k_B, respectively, and the effective alloy spring constant be k, with similar notations for the bond lengths d_A, d_B, and d. When an A atom is embedded in the medium, all 16 bonds under consideration will relax, and the strain energy is given by

$$\varepsilon(\mathrm{A}) = \frac{1}{2} k_1 (d - d_A)^2, \tag{128}$$

where

$$k_1 = 4k_A k/(3k_A + k). \tag{129}$$

A similar energy $\varepsilon(B)$ is obtained, when a B atom is embedded. The effective bond length d is obtained from a minimization of the average cluster energy $E = (1 - x)\varepsilon(A) + x\varepsilon(B)$ with respect to d, which yields

$$d = [(1 - x)k_1 d_A + xk_2 d_B]/[(1 - x)k_1 + xk_2]. \tag{130}$$

When the alloy is compressed, the alloy bond length is reduced to $d(1 - e)$, where e is a macroscopic strain corresponding to the external pressure. The pressure-induced strain energy for the 16 bonds in the medium is $\delta E = 8k(de)^2$, and $\delta E = 2k(de)^2$ for each cluster. Embedding an A atom in this compressed medium, one finds the total strain energy for the 16 bonds to be

$$E_A = \frac{1}{2} k_1 (d - d_A - 4de)^2. \tag{131}$$

To obtain the extra cluster energy $\delta\varepsilon(A)$ induced by the pressure, we subtract $\varepsilon(A)$ of Eq. (127) and the background energy for the surrounding 12 bonds from E_A to give

$$\delta\varepsilon(A) = -4k_1(d - d_A)de + 8k_1(de)^2 - 6k(de)^2 . \tag{132}$$

Similarly, the following expression for $\delta\varepsilon(B)$ is obtained when the embedded atom is B:

$$\delta\varepsilon(B) = -4k_2(d - d_B)de + 8k_2(de)^2 - 6k(de)^2 . \tag{133}$$

Thus, the change of the average cluster energy that is due to the pressure is given by $\delta E = \langle\delta\varepsilon(n)\rangle = (1 - x)\delta\varepsilon(A) + x\delta\varepsilon(B)$, which, when equated to $2k(de)^2$, leads to the following self-consistent equation for the effective spring constant k : $k = (1 - x)k_1 + xk_2$. The k can now be solved analytically when both the expression for k_1 in Eq. (129) and the similar expression for k_2 are used. The result is

$$k = \langle k\rangle[1 - 3x(1 - x)(\delta k/\langle k\rangle)^2], \tag{134}$$

where $\langle k\rangle = (1 - x)k_A + xk_B$ is the mean spring constant and $\delta k = k_A - k_B$ the difference. It is interesting to compare this result for the 50/50 alloy, i.e., $k = \bar{k}(1 - \frac{3}{4}\Delta_o^2)$ with the value $\bar{k}(1 - \Delta_o^2/4)$ in Eq. (123) for the ordered alloys in the CuAuI and chalcopyrite structures. The alloy spring constant is slightly below the straight-line average, and the bowing is larger for a disordered alloy than for the corresponding ordered compound. Using the effective spring constant of Eq. (134), we find that the effective bond lengths for most alloys also bow slightly below their mean value,

$$d = \langle d\rangle + 4x(1 - x)(d_A - d_B)(k_A - k_B)k/[(3k_A + k)(3k_B + k)] \tag{135}$$

because the spring constant tends to increase as the bond length decreases.

To compare the calculations above with experimental data for pseudobinary alloys, we were able to find results for GaAlAs (Landolt-Bornstein, 1988), CdZnTe, CdMnTe, and HgCdTe (Quadri *et al.*, 1986). For GaAlAs, the following linear x dependences were measured (Landolt-Bornstein, 1988): $C_{11} = 11.85 + 0.14x$, $C_{12} = 5.38 + 0.32x$, and $C_{44} = 5.94 - 0.05x$. This lack of detectable bowing is expected, because of the nearly equal bond lengths of the two constituent compounds and small differences in the elastic constants. The bulk moduli in the three II-VI alloy systems mentioned were obtained from high-pressure x-ray diffraction data. For HgCdTe, the results are similar to those for the GaAlAs in that both the bond lengths and the bulk moduli of

TABLE XV

MEASURED ELASTIC CONSTANTS IN 10^{11} DYNES/CM^2 OF SiGe ALLOYS BY BUBLIK *ET AL* (1974)

Alloy	C_{11}	C_{12}	C_{44}
$Si_{0.28}Ge_{0.72}$	16.1 ± 0.8	8.35 ± 0.8	8.55 ± 0.4
$Si_{0.54}Ge_{0.46}$	17.0 ± 0.8		
$Si_{0.46}Ge_{0.36}$	17.1 ± 0.8		

HgTe and CdTe are so close that the differences in B between the alloys and the pure crystals were beyond the experimental resolution. However, a 5% Zn in CdZnTe alloy was found to give a 15% increase in the B value from the pure CdTe value and a 10% Mn in CdMnTe gave a 21% decrease (Quadri *et al.*, 1986). These significantly large changes in the B values caused by smaller concentrations cannot be explained from the above considerations.

The qualitative model considered above does not apply to the binary alloys $A_{1-x}B_x$, because in these alloys both the A and B atoms can be found in both sublattices, and the local bond length arrangement is more complicated than the pseudobinary alloys. However, one can expect that there are still more degrees of relaxation in the disordered binaries than in the ordered compounds. Therefore, one would conclude that the bulk modulus of the disordered 50-50 SiGe alloy would have a smaller value than those tabulated in Table XIII for the ordered compounds. At least one would not expect the alloy B values to be significantly larger than the mean values $\bar{B}$. However, the only experimental data available (Bublik *et al.*, 1974), Table XV, show that all three elastic constants for these alloys at three different concentrations exceed the values for Si, and that $\Delta B/\bar{B}$ is as large as 20%, despite the fact that the bond length difference between Si and Ge is only about 3% and the measured alloy lattice constants are only bowed slightly below the average. This, and the unexplained results for the II-VI alloys, point to the need for a more systematic study of the elastic properties of semiconductor alloys, both experimentally and theoretically.

IV. Dislocations and Hardness[8]

Hardness has proven to be a useful probe of the mechanical properties of the brittle semiconductors. Here we will use the term hardness to refer specifically to Vickers' hardness, unless otherwise noted. In the Vickers'

[8]Much of this section is adapted from "Final Report" (AFOSR-F49620-85-0023) by M. A. Berding (1988), SRI International, Menlo Park, California.

hardness measurement, a square pyramidal indenter is used, and the hardness number is given by the applied load divided by the area of the indentation (i.e., units of pressure). Hardness has been found to be an intrinsic property of the material, because it is relatively independent of the applied load. One advantage of hardness measurements for semiconductors is that, in contrast to bending tests, only small samples are necessary for conventional Vickers' hardness measurements, or for nanoindenter measurements (Fang *et al.*, 1990), and relatively thin epitaxial films can be probed. Additionally, unlike conventional tests used to measure yield stress, hardness measurements can be made at room temperature, which is far below the usual plastic regime for most semiconductors. As such, the hardness measurement provides a convenient and usable probe.

The question remains, though, as to the interpretation of the hardness measurement in semiconductors: Just what property or properties of a semiconductor are we measuring when we measure hardness? In metals, an empirical relationship is found between the hardness H and yield stress Y, such that $H = 3Y$.

In metals, this relationship can be justified on the basis of continuum theory, as discussed in McClintock and Argon (1966). In semiconductors such a simple relationship between H and Y is not necessarily appropriate for several reasons. During deformation in metals, many slip planes can be active because the Peierls barriers for dislocation motion in most directions are low. In contrast, because the bonds in semiconductors are strongly covalent, the Peierls barriers are high and dislocation in the $(111)1/2\langle\bar{1}10\rangle$ slip system dominate.

To date, there is no complete quantitative theory of hardness in the semiconductors in which the temperature dependence, photoplastic effect, and the alloy hardening effect are included. Sher *et al.* (1985) proposed a model of hardness for the semiconductor compound that gives good quantitative agreement with experiment, but this model does not provide an explanation for several of the observed dependences of hardness. This model of hardness in semiconductors differs from more conventional interpretations and suggests that hardness is dominated by dislocation-dislocation interactions, as opposed to dislocation activation and motion terms. We discuss the results of an improved quantitative model of hardness below.

12. Slip Systems

In the dislocation interaction hardness model, Vickers' hardness is found to be dominated by the interaction energy of an idealized array of dislocations that has been generated by the indenter. The idealized array

can be considered as a first approximation to the more realistic dislocation tangles found experimentally, leading in higher order to an expansion in dislocation configurations. In this idealized array, no account is taken of the true slip systems active in the semiconductors. Experiments (Hirsch *et al.*, 1985) have demonstrated that, in Vickers' hardness, slip occurs primarily on the $\{111\}\frac{1}{2}\langle\bar{1}10\rangle$ glide set, where the threefold symmetry of slip and rosette lines occurs at the intersection of the $\{111\}$ planes with the (111) surface.

For indentation on the (111) plane, dislocations can glide on the (111) plane parallel to the surface or on one of the three other $\{111\}$ planes with a total of four active slip planes. Although the detailed analysis differs from that given previously (Sher *et al.*, 1985), the contribution to the hardness from the interaction energy is comparable to that previously calculated. This contribution to H is directly proportional to the shear coefficient.

13. Peierls Energy

The Peierls energy is difficult to calculate precisely because of dislocation charge effects and reconstruction at the dislocation core. In the context of the hardness measurement we calculate the Peierls energy in order to evaluate the importance of this contribution to the Vickers' hardness number. Although it is generally agreed that dislocations in semiconductors move through the generation and propagation of double kinks, in the hardness measurement, the region about the indenter is grossly plastically deformed. Because the dislocation velocity is low at room temperature (see below), the large dislocation pile-up model proposed by Sher *et al.* (1985) may be appropriate. If the dislocation separation is small, dislocation motion through kink processes will be suppressed and the dislocations will propagate as a complete unit.

To get from Configuration A to Configuration B in Fig. 7, we must break a row of bonds. Since the long-range strain fields should be comparable in the two configurations as well as in intermediate configurations, the Peierls force can be calculated from local energy considerations only. The energy to break a bond at the dislocation core is approximately given by

$$U_b = 2\sqrt{V_2^2 + V_3^2} + 2\varepsilon_{met} - V_0 + \frac{n}{4}(\varepsilon_h^a - \varepsilon_h^c), \tag{136}$$

where V_2 is the covalent energy, V_3 the ionic energy, ε_{met} the metallization energy, V_o the bond overlap energy, n is 1 for III-V and 2 for II-VI compounds, and ε_h^a and ε_h^c the hybrid energy for the anion and cation, respectively. The first two terms in Eq. (136) account for the loss of the

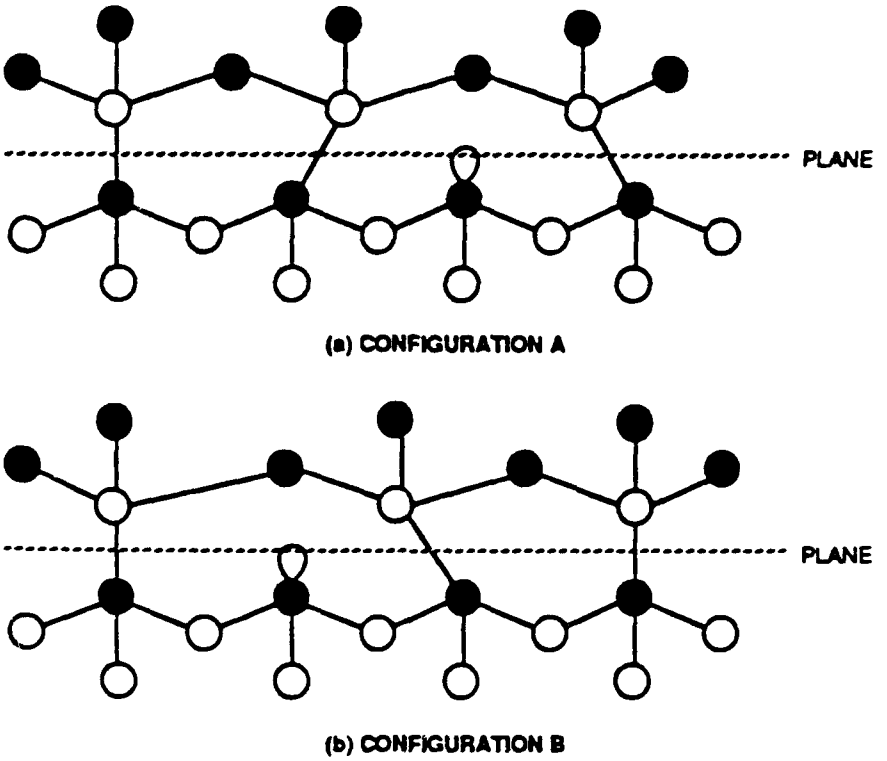


FIG. 7. Atom configurations during the slip of a dislocation. (a) Configuration A; (b) Configuration B.

bonding energy of the two electrons in the breaking bond, the third term accounts for regaining the repulsive interaction energy of the bond, and the fourth term accounts for the energy gain to transfer electrons back from the cation to the anion. We note that the electron orbitals of the atoms at the dislocation core are left in the sp^3 hybrids after the bond breaking. The expression in Eq. (136) represents a theoretical maximum of the Peierls energy, since no reconstruction at the core has been included.

To calculate the Peierls energy per unit length, we consider a primary dislocation in the $\langle \bar{1}10 \rangle$ direction in a zincblende compound. The number of bonds per unit length in $\langle \bar{1}10 \rangle$ is given by $1/b$, where b is Burger's vector. Thus, the Peierls energy per unit length is given by

$$E_p = \frac{U_b}{b} = \sqrt{\frac{3}{8}} \frac{U_b}{d}, \tag{137}$$

where d is the bond length.

We can now calculate the Peierls force, or the force per unit length necessary to move a dislocation over the potential barrier, as illustrated in

Fig. 8. The Peierls energy is related to the Peierls force through

$$E_p = F_p L, \tag{138}$$

where

$$L = \frac{b}{2} \tag{139}$$

is the distance between Configuration A and Configuration B. Solving for F_p in terms of U_b and d, we arrive at:

$$F_p = \frac{3}{4}\frac{U_b}{d^2}, \tag{140}$$

or

$$\tau_p = \sqrt{\frac{3}{2}}\frac{3}{8}\frac{U_b}{d^3}. \tag{141}$$

Values for U_b and τ_p are summarized in Table XVI.

Now we incorporate the Peierls energy into the hardness model for low temperature where the full barrier must be surmounted. The Sher model is based on energy considerations. The Vickers' hardness number is given by the applied force divided by the area of indentation. Multiplying the numerator and denominator by h, the depth of indentation, we have

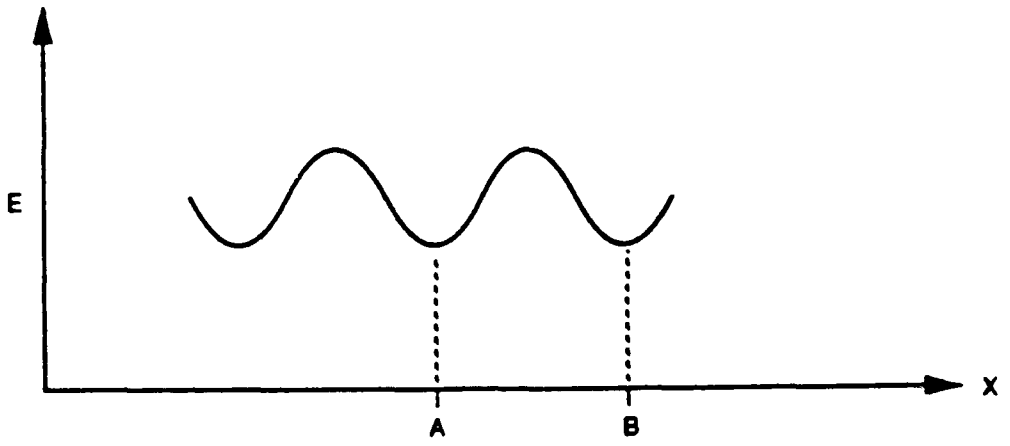


FIG. 8. Schematic of the dislocation potential as a function of its position.

TABLE XVI

CALCULATED PEIERLS STRESS AND HARDNESS FOR VARIOUS ZINCBLENDE SEMICONDUCTORS, WITH EXPERIMENTAL HARDNESS VALUES FOR COMPARISON[a]

	U_b	τ_p	H_p	H_{int}	$H_p + H_{int}$	H_{exp}
C	11.35	23,300	1940	9244	11,184	10,000
Si	5.95	3430	286	1098	1384	1370
Ge	6.67	3440	286	893	1179	1000
Sn	5.66	1930	161	—	—	—
AlP	6.04	3410	284	—	—	—
GaP	6.15	3500	292	903	1195	940
InP	5.78	2640	220	548	768	520
AlAs	5.90	3000	250	—	—	505
GaAs	6.03	3000	256	750	1006	580
InAs	5.64	2340	195	469	664	430
AlSb	5.08	2840	237	524	761	400
GaSb	5.36	2180	182	553	735	450
InSb	4.92	1670	139	365	504	230
ZnS	5.14	3000	250	515	765	—
CdS	4.77	2750	229	288	517	—
HgS	5.68	2590	216	—	—	—
ZnSe	4.92	2500	209	462	671	137
CdSe	4.59	1890	157	254	411	—
HgSe	4.523	1840	154	232	386	—
ZnTe	4.42	1720	143	374	517	82
CdTe	4.01	1350	113	222	335	60
HgTe	3.99	1360	113	230	343	25

[a]The U values are in eV, and the others in kg/mm^2.

$H = E/(W^2h)$, where $E = Fh$ is the energy of indentation, and h is the depth of the indentation. Including the interaction energy only we have

$$H = H_{int} = \frac{G \cot\theta}{6\pi(1-\nu)}\left[-\ln\left(\frac{\cot\theta}{2}\right) + \frac{4}{3} + \sin^2\frac{\theta}{2}\right], \tag{142}$$

where θ is one half the indenter angle. To include the Peierls energy, we consider the total energy necessary to move the dislocations from their initial to final positions in the idealized model. The hardness is then given by:

$$H = H_{int} + H_p, \tag{143}$$

where

$$H_p = \frac{E_p}{W^2 h} \tag{144}$$

is the Peierls contribution to the hardness, and E_p is the total Peierls energy expended. The total length of dislocation to be moved is calculated to be

$$L_T = \frac{1}{6}\frac{W^3}{b^2}\cos^2\theta. \tag{145}$$

The total Peierls energy is given by

$$E_p = \frac{U_b}{b} L_T = \frac{1}{6}\frac{U_b W^3}{b^3}\cos^2\theta. \tag{146}$$

Thus, we have

$$H_p = \frac{1}{3}\frac{U_b}{b^3}\cos\theta\sin\theta. \tag{147}$$

For $\theta = 45°$,

$$H_p = \frac{1}{6}\frac{U_b}{b^3}. \tag{148}$$

Values of H_p are summarized in Table XVI. Several features of H_p should be noted. First, we have used a zero-temperature value of the Peierls energy. Because hardness measurements are typically done at room temperature, one should take the thermal energy into account; this will reduce the values of H_p from those listed in Table XVI. Also shown in Table XVI are H_{int}, $H_{int} + H_p$, the best theoretical estimate for H, and H_{exp}. Note that, like H_{int}, H_p is independent of the applied load, in agreement with experiment. Also note that H_p improves the agreement between theory and experiment for the hard, nonpolar materials. For the softer, more ionic materials, H is overestimated by the theory. The overestimation of H may be because of neglect of dislocation velocity effects and their temperature dependence, as discussed below.

14. Temperature Dependence

Here we summarize the experimental results and discuss a tentative theory of the temperature dependence of the hardness.

Several recent studies on the temperature dependence of hardness serve to illustrate the behavior. Results for GaAs and Ge are shown in Fig. 9. The (111) and (100) faces of GaAs have been examined by Hirsch *et al.* (1985) and Guruswamy *et al.* (1986), respectively. Results for the Knoop hardness on the (100) face of *n*-type Ge are also shown (Roberts *et al.*, 1986). The (100) face of GaAs and the Ge show a definite temperature dependence with a relatively temperature-independent region for $T < 450$ K and an exponential temperature dependence for $T > 550$ K:

$$H = H_0 e^{U/kT}, \tag{149}$$

with

$$U \cong 0.24 \text{ eV} \tag{150}$$

for (100) GaAs. The results for GaAs (111) appear to follow a similar behavior.

The temperature dependence of hardness suggests that two different mechanisms may determine hardness in the two temperature regimes. At low temperature, the hardness is nearly independent of temperature and may be limited primarily by dislocation interactions. Dislocation mobility is low at low temperatures, and the tendency for dislocation pile-up is high. At elevated temperature, the dislocation mobility is increased, so that dislocations move more readily under an applied stress. Therefore, at higher temperatures, dislocation pile-up is reduced and the hardness is limited by lattice friction, which shows a strong temperature dependence.

V. Concluding Remarks

The experimental methods available to measure elastic constants vary greatly in their accuracy and in the size of samples required. Generally, those that measure the velocity of sound are quite accurate, some yielding elastic constants to one part in 10^6. These methods are also capable of measuring higher-order elastic constants, a subject not treated in this paper. However, the samples required for these measurements must be large, of the order of several centimeters, and must be perfect bulk single crystals. Many semiconductors, alloys in particular, are only grown as thin films on disparate

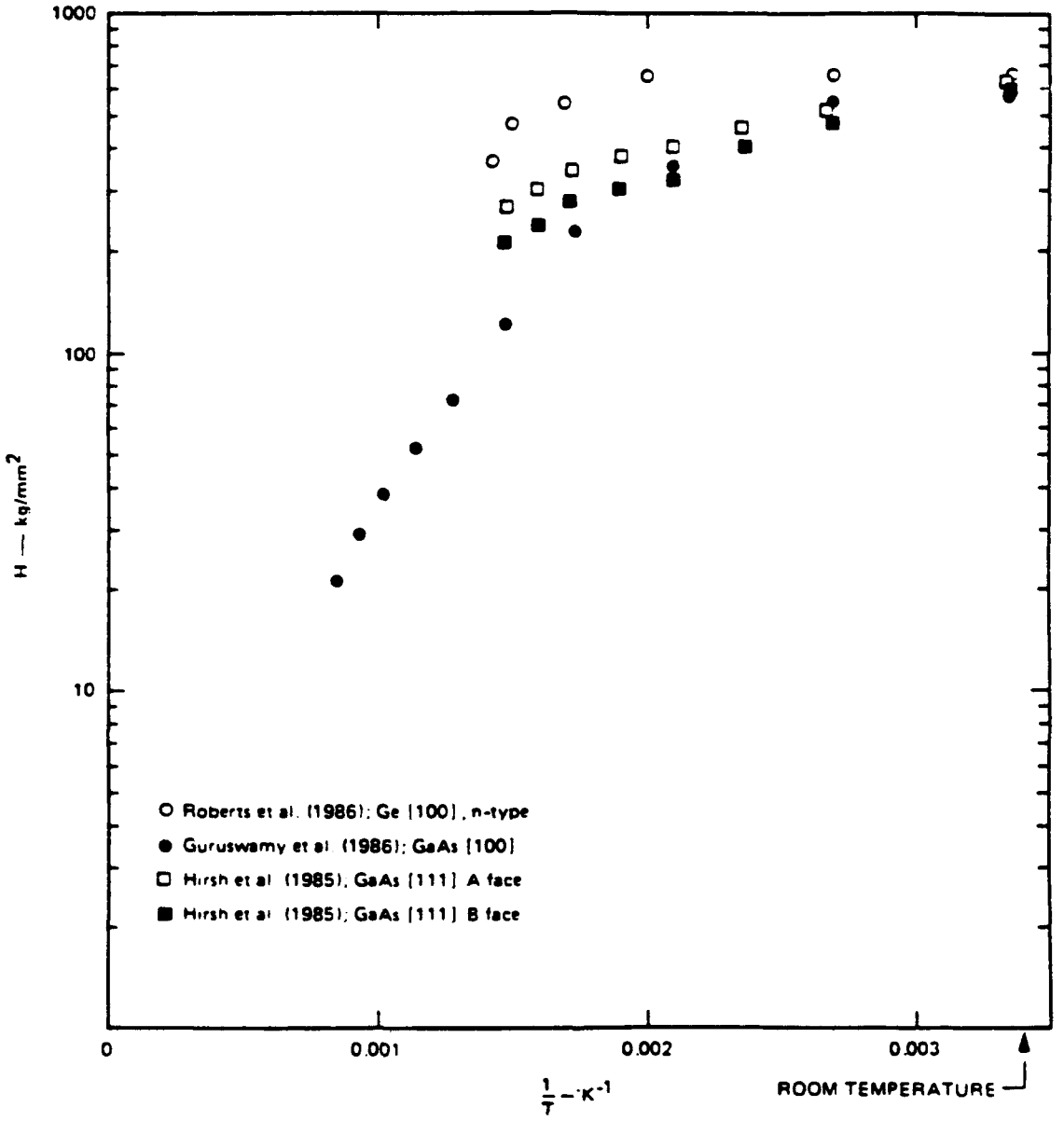


FIG. 9. Measured hardness of Ge and several GaAs samples as a function of temperature.

substrates. For these examples, the velocity-of-sound methods fail, and the less accurate Raman and Brillouin scattering techniques become the methods of choice. Their accuracy is about 1 to 4%, which is adequate for many practical applications.

Most group IV, III-V compound, and II-VI compound semiconductors have been studied, and their elastic constants tabulated. A few remain to be examined, and several should be reexamined because different experimenters do not agree on the results. The situation in the pseudobinary alloys is quite

different. Few alloy systems have been adequately studied; those studied have mostly fallen into the class of materials in which the bond lengths of the constituents nearly match. More interesting results are expected from alloys with a bond length mismatch. Such studies would yield a wealth of information on mechanisms responsible for correlations in these alloys, and perhaps even on those responsible for producing the ordered alloys that have been grown recently.

We have emphasized the utility of various parameterized models for treating nonideal situations. However, the most powerful new theoretical developments are in the area of first-principles theories. The advent of self-consistent local density theories (Hohnberg and Kohn, 1964) more than 20 years ago, and the advances in methods to solve the Schroedinger equation, are making real applications, as evident from the excellent structural and elastic properties produced by these theories (see Table II and Anderson *et al.*, 1985). Recent progress in the full-potential LMTO method (Methfessel and van Schilfgaarde *et al.*, 1990), in the non-self-consistent approach (Harris, 1985; Foules and Haydock, 1989; van Schilfgaarde *et al.*, 1991), and the Carr-Parrinello (1985) quantum molecular dynamics extension has increased solution speeds into the realm in which it is practical to attack many mechanical property problems. Moreover, LDA has now been extended to include many-body corrections (Hybertsen and Louie, 1987), so properties that are sensitive to conduction bands can now be computed accurately. As computational speeds continue to increase, these methods will evolve into practical engineering tools.

Acknowledgments

This work was supported in part by AFOSR Contract F49620-88-K-0009 and ONR Contract N00014-88-C0096. We are especially indebted to M. A. Berding for permitting us to include in Section IV unpublished ideas from a report she wrote. We also wish to acknowledge contributions from J. Heyman.

References

Anderson, O. K., Jepson, O., and Glötzel, D. (1985). *In* "Highlights of Condensed Matter Theory," (F. Bassiani *et al.*, eds.). North Holland, Amsterdam.

Ashcroft, N. W., and Mermin, N. D. (1976), p. 445. "Solid State Physics." Saunders College, Philadelphia.

Bachelet, G. B., Hamman, D. R., and Schluter, M. (1982). *Phys. Rev.* **B26**, 4199.

Balzarotti, A., Czyzk, M. T., Kisel, A., Motta, N., Podgorny, M., and Zimnal-Starnawska, M. (1985). *Phys. Rev.* **B31**, 7526.

Barnes, J. M., and Hiedemann, E. A. (1957). *J. Acoust. Soc. Am.* **28**, 1218-1221; **29**, 865.
Baroni, S., Giannozzi, P., and Testa, A. (1987). *Phys. Rev. Lett.* **58**, 1861.
Benedek, G. B., and Fritsch, K. (1966). *Phys. Rev.* **149**, 647-662.
Benson, G. C., and Kiyohara, O. (1974). *J. Acoust. Soc. Am.* **55**, 1184-185.
Berera, A., Dreysse, H., Wille, L. T., and de Fontaine, D. (1988). *Phys. Rev.* **F18**, 149.
Bernard, J. E., Ferreira, L. G., Wei, S.-W., and Zunger, A. (1988). *Phys. Rev.* **B38**, 6338.
Beyer, R. T., and Letcher, S. V. (1969). "Physical Ultrasonics," Vol 32 in *Pure and Applied Physics*. (H. S. W. Massey and K. A. Brueckner, consulting editors), pp. 47-50. Academic Press, New York.
Bhatia, A. B., and Noble, W. J. (1953). *Proc. Roy. Soc. London, Ser. A* **220**, 356-368, 369-385.
Blackman, M. (1959). *Phil. Mag.* **3**, 831.
Bloomfield, P. E., Ferren, R. A., Radice, P. F., Stefanou, H., and Sprout, O. S. (1978). *Nav. Res. Rev.* (May), 1-15.
Bolef, D. I., and Miller, J. G. (1971). *In* "Physical Acoustics," (W. P. Mason and R. N. Thurston, eds.), Vol. VIII, pp. 96-201, Academic Press, New York.
Bolef and Miller (1974). Table 2.
Bolef, D. I., deKlerk, J., and Gosser, R. B. (1962). *Rev. Sci. Instrum.* **33**, 631.
Born, M., and Oppenhimer, J. R. (1927). *Ann. Phys. 4*, **84**, 457.
Born, M., and Wolf, E. (1970). "Principles of Optics," 4th ed., pp. 593-610. Pergamon Press, Oxford. A concise treatment of this subject appears in Chapter 12.
Boyce, J. B., and Mikkelsen, J. C. Jr. (1985). *Phys. Rev.* **B31**, 6903.
Breazeale, M. A., and Ford, J. (1965). *J. Appl. Phys.* **36**, 3486. For the pure longitudinal mode directions and for isotropic materials, the nonlinear wave equation has been solved here.
Breazeale, M. A., Cantrell, J. H. Jr., and Heyman, J. S. (1981). "Methods of Experimental Physics, Vol. 19, Ultrasonics," 67-135. Academic Press, New York.
Brewer, R. (1965). *Appl. Phys Lett.* **8**, 165.
Brewer, R. G., and Rieckhoff, K. E. (1964). *Phys. Rev. Lett.* **13**, 334-336.
Brugger, K. (1964). *Phys. Rev.* **133**, A1611.
Bublik, V. T., Gorelik, S. S., Zaitsev, A. A., and Polyakov, A. Y. (1974). *Phys. Stat. Soli. B* **66**, 427.
Callaway, J., and March, N. H. (1984). *Solid State Phys.* **38**, 135.
Cantrell, J. H., Jr. (1982). *J. Testing and Eval.* **10**, 223.
Cantrell, J. H., and Breazeale, M. A. (1977). *J. Acoust. Soc. Am.* **61**, 403.
Carr, M., and Parrinello, M. (1985). *Phys. Rev. Lett.* **55**, 2471.
Ceperly, D. M., and Alder, B. J. (1980). *Phys. Rev. Lett.* **45**, 566.
Chadi, D. J. (1978). *Phys. Rev. Lett.* **41**, 1062.
Chadi, D. J. (1979). *Phys. Rev.* **B19**, 2074.
Chadi, D. J. (1984). *Phys. Rev.* **B29**, 785.
Chadi, D. J., and Cohen, M. L. (1973). *Phys. Rev.* **B8**, 5747.
Chen, A.-B., and Sher, A. (1982). *J. Vac. Sci. & Technol.* **21**(1), 138.
Chen, A.-B., and Sher, A. (1985). *Phys. Rev.* **B32**, 3695.
Chen, A.-B., van Schilfgaarde, M., Krishnamurthy, S., Berding, M. A., and Sher, A. (1987). *In* "Ternary and Multinary Compounds" (S. K. Deb and A. Zunger, eds.). Materials Research Society, Pittsburgh, Pennsylvania.
Chen, A.-B., Berding, M. A., and Sher, A. (1988). *Phys. Rev.* **B37**, 6285.
Chern, E. J., Cantrell, J. H., and Heyman, J. S. (1981). *J. Appl. Phys.* **52**, 3200-3204.
Chiao, R. Y., Townes, C. H., and Stoicheff, B. P. (1964). *Phys. Rev. Lett.* **12**, 592-595.
Cohen, M. L. (1985). *Phys. Rev.* **B32**, 7988.
Connolly, J. W.D., and Williams, A. R. (1983). *Phys. Rev.* **B27**, 5169.
Conradi, M. S., Miller, J. G., and Heyman, J. S. (1974). *Rev. Sci. Instrum.* **45**, 358-360.
Cousins, C. S.G., Gerward, L., Staun, L. J., Selsmark, B., and Sheldon, B. J. (1987). *J. Phys.* **C20**, 29.

Dreysse, H., Berera, A., Wille, L. T., and de Fontaine, D. (1989). *Phys. Rev.* **B39**, 2442.
Elmore, W. C., and Heald, M. A., (1969). "The Physics of Waves," pp. 98-104. McGraw-Hill, New York.
Fang, S., Farthing, L. J., Tang, M.-F.-S., and Stevenson, D. A. (1990). *J. Vac. Sci. Technol.* **A8**(2), 1120.
Ferreira, L. G., Wei, S.-H., and Zunger, A. (1989). *Phys Rev.* **B40**, 3197.
Fleury, P. A. (1970). *In* "Physical Acoustics" (W. P. Mason and R. N. Thurston, eds.), Vol. 6, Chap. 1. Academic Press, New York.
Foulkes, M., and Haydock, R. (1989). *Phys. Rev.* **B39**, 12520.
Froyen, S., and Harrison, W. A. (1979). *Phys. Rev.* **B20**, 2420.
Gautier, G., Ducatelle, F., and Giner, J. (1975). *Phil. Mag.* **31**, 1373.
Goldstein, H. (1965). "Classical Mechanics," p. 347. Addison-Wesley, Reading, Massachusetts.
Gomyo, A., Suzuki, T., Kobayashi, K., Kawata, S., Hino, I., and Yuasa, T. (1987). *Appl. Phys. Lett.* **50**, 673.
Gomyo, A., Suzuki, T., and Iijima, S. (1988). *Phys. Rev. Lett.* **60**, 2645.
Green, R. E., Jr. (1973). "Treatise on Materials Science and Technology, Vol. 3, Ultrasonic Investigation of Mechanical Properties," pp. 7-8, 11-25. Academic Press, New York.
Guruswamy, S., Hirth, J. P., Faber, K. T. (1986). *J. Appl. Phys.* **60**, 4136.
Gyorffy, B., and Stocks, G. M. (1978). *Inst. Phys. Conf. Ser.* **39**, 394.
Harris, J. (1985). *Phys. Rev.* **B31**, 1770-1790.
Harrison, W. A. (1980). "Electronic Structure and the Properties of Solids." Freeman and Co., San Francisco.
Harrison, W. A. (1983a). *Phys. Rev.* **B27**, 3592.
Harrison, W. A. (1983b). "The Bonding Properties of Semiconductors," SRI International, Menlo Park, California.
Hass, K. C., Ehrenreich, H., and Velicky, B. (1983). *Phys. Rev.* **B27**, 1088.
Hedin, L., and Lunquist, B. I. (1971). *J. Phys.* **C4**, 2064.
Hiedemann, E. A., and Hoesch, K. H. (1934). *Z. Phys.* **90**, 322-326.
Hiedemann, E. A., and Hoesch, K. H. (1937). *Z. Phys.* **107**, 463-473.
Heyman, J. S. (1976). *Proc. IEEE Ultrason. Symp., Annapolis, Maryland*, p. 113.
Hirsch, P. B., Pirouz, P., Roberts, S. G., and Warren, P. D. (1985). *Philo. Mag.* **B52**, 759.
Hohenberg, P., and Kohn, W. (1964). *Phys. Rev.* **136**, B864.
Huang, D., Chyi, J., Klem, J., and Morkoc, H. (1988). *J. Appl. Phys.* **63**, 5859.
Hybertsen, M. S., and Louie, S. G. (1987). *Phys. Rev. Lett.* **58**, 1551.
Ihm, Y. E., Otuska, N., Klem, K., and Morkoc, H. (1987). *Appl. Phys. Lett.* **51**, 2013.
Jen, H. R., Cherng, M. J., and Stringfellow, G. B. (1986). *Appl. Phys. Lett.* **48**, 1603.
Johnson, G. C., and Mase, G. D. (1984). *J. Acoust. Soc. Am.* **75**, 1741.
Keating, P. N. (1966). *Phys. Rev.* **145**, 637.
Khimunin, A. S. (1972). *Acustica* **27**, 173-181.
Kikuchi, R. (1951). *Phys. Rev.* **81**, 988.
Kittel, C. (1986). "Introduction to Solid-State Physics," Table 5. John Wiley & Sons, Inc., New York.
Kleiman L. (1962). *Phys. Rev.* **128**, 2614.
Klein, W. R., Cook, B. D., and Mayer, W. G. (1965). *Acustica* **15**, 67-74. This work is a theoretical analysis and defines Q.
Klem, J., Huang, D., Morkoc, H., Ihm, Y. E., and Otuska, N. (1987). *Appl. Phys. Lett.* **50**, 1364.
Kohn, W., and Sham, L. J. (1965). *Phys. Rev.* **140**, A 1133.
Krahauer, H., Posternak, M., and Freeman, A. J. (1979). *Phys. Rev. Lett.* **43**, 1885.
Krischer, C. (1968). *Appl. Phys. Lett.* **13**, 310.
Kuan, T. S., Kuech, T. F., Wang, W. I., and Wilkie, E. L. (1985). *Phys. Rev. Lett.* **54**, 201.

Kuan, T. S., Wang, W. I., and Wilkie, E. L. (1987). *Appl. Phys. Lett.* **51**, 51.
Lambrecht, W. R.L., and Segall, B. (1990). Private communication.
Landolt-Bornstein (1982). "Numerical Data and Functional Relationships in Science and Technology" (K.-H. Hellwidge, ed.), Vol. 17. Springer-Verlag, Berlin.
Landolt-Bornstein (1988). "Numerical Data and Functional Relationships in Science and Technology" (K.-H. Hellwidge, ed.), Vol. 22. Springer-Verlag, Berlin.
Landau, L. D., and Lifshitz, E. M. (1959). "Theory of Elasticity." Pergamon Press, New York.
Landau, L. D., and Lifshitz, E. M. (1986). "Theory of Elasticity." 3rd ed. Pergamon Press, New York. Equation 16 of text is in slightly different form than given on p. 2. Using $x_i = a_i + u_i$, and substituting into Eq. (16), one obtains the strain tensor found in this reference. This reference stresses physics concepts to derive the important equations.
Love, A. E.H. (1944). "A Treatise on the Mathematical Theory of Elasticity." Dover, New York.
Mann, J. B. (1967). "Atomic Structure Calculations, 1: Hartree-Fock Energy Results for Elements from Hydrogen to Lawrencium." Clearinghouse for Technical Information, Springfield, Virginia.
Martin, R. M. (1970). *Phys. Rev.* **B1**, 4005.
Martins, J. L., and Zunger, A. (1984). *Phys. Rev.* **B30**, 6217.
Martins, J. L., and Zunger, A. (1986). *Phys. Rev. Lett.* **56**, 1400.
Mayer, W. G., and Hiedemann, E. A. (1958). *J. Acoust. Soc. Am.* **30**, 756-760.
Mayer, W. G., and Hiedemann, E. A. (1959). *Acta Crystallogr.* **12**, 1.
McClintock, F. A., and Argon, A. S. (1966). "Mechanical Behavior of Materials." Addison-Wesley, Reading, Massachusetts.
McSkimin, H. J. (1950). *J. Acoust. Soc. Am.* **22** 413-418.
McSkimin, H. J. (1957). *IRE Trans. Ultrason. Eng.* **5**, 25.
McSkimin, H. J. (1961). *J. Acoust. Soc. Am.* **33**, 12.
McSkimin, H. J., and Andreatch, P. (1964). *J. Appl. Phys.* **35**, 3312-3319.
Methfessel, M., and van Schilfgaarde, M. (1990). Private communication.
Methfessel, M., Rodriguez, C. O., and Andersen, O. K. (1989). *Phys. Rev.* **B40**, 2009.
Michard, F., and Perrin, B. (1978). *J. Acoust. Soc. Am.* **64**, 1447-1456.
Mikkelsen, J. C., Jr., and Boyce, J. B. (1982). *Phys. Rev. Lett.* **49**, 1412.
Mikkelsen, J. C., Jr., and Boyce, J. B. (1983). *Phys. Rev.* **B28**, 7130.
Miller, J. G. (1973). *J. Acoust. Soc. Am.* **53**, 710-713.
Mitra, S. S., and Massa, N. E. (1982). *In* "Handbook on Semiconductors" (T. S. Moss, ed.), Vol. 1, Chap. 3. North-Holland, Amsterdam.
Murnaghan, F. D. (1951). "Finite Deformation of an Elastic Solid." Wiley, New York. This reference uses matrices throughout and is mathematically more formal.
Musgrave, J. P., and Pople, J. A. (1962). *Proc. Roy. Soc. (London)* **A268**, 474.
Nielsen, O. H., and Martin, R. M. (1983). *Phys. Rev. Lett.* **50**, 697.
Nielsen, O. H., and Martin, R. M. (1985a). *Phys. Rev.* **B32**, 3792.
Nielsen, O. H., and Martin, R. M. (1985b). *Phys. Rev.* **B32**, 3780.
Nomoto, O. (1942). *Proc. Phys. Math. Soc. Japan* **24**, 380-400, 613-639. This work covers an experimental investigation into the overlap region.
Norman, A. G., Mallard, R. E., Murgatroyd, I. J., Brooker, G. R., More, A. H., and Scott, M. D. (1987). *In* "Microscopy of Semiconductor Materials," *Inst. Phys. Conf. Ser.* **87**, 77.
O'Donnell, M., Busse, L. J., and Miller, J. G. (1981). In "Methods of Experimental Physics, Vol. 19, Ultrasonics," pp. 29-65. Academic Press, New York.
Ourmazd, A., and Bean, J. C. (1985). *Phys. Rev. Lett.* **55**, 765.
Papadakis, E. P. (1963). *J. Acoust. Soc. Am.* **35**, 490-494.
Papadakis, E. P. (1964a). *J. Appl. Phys.* **35**, 1474.
Papadakis, E. P. (1964b). *J. Acoust. Soc. Am.* **36**, 414-422.

Papadakis, E. P. (1966). *J. Acoust. Soc. Am.* **40**, 863-876.
Papadakis, E. P. (1967). *J. Acoust. Soc. Am.* **42**, 1045.
Papadakis,. E. P. (1972). *J. Acoust. Soc. Am.* **52**, 843.
Papadakis, E. P. (1976). *In* "Physical Acoustics" (W. P. Mason and R. N. Thurston, eds.), Vol. XII, pp. 277-374. Academic Press, New York.
Patrick, R. S., Chen, A.-B., and Sher, A. (1987). *Phys. Rev.* **B36**, 6585.
Patrick, R. S., Chen, A.-B., Sher, A., and Berding, M. A. (1988). *J. Vac. Sci. and Technol.* **6**, 2643.
Perdew, J. P., and Zunger, A. (1981). *Phys. Rev.* **B23**, 5048.
Phillips, J. C. (1973). "Bonds and Bands in Semiconductors." Academic Press, New York.
Pollard, H. F. (1977). "Sound Waves in Solids," pp. 308-309. Pion Limited, London.
Prosser, W. H., and Green, R. E., Jr. (1985). *In* "Proceedings of the 1985 SEM Spring Conference on Experimental Mechanics," pp. 340-346. Society for Experimental Mechanics.
Qteish, A., and Resta, R. (1988). *Phys. Rev.* **B37**, 1308.
Quadri, S. B., Skelton, E. F., and Webb, A. W. (1986). *J. Vac. Sci. and Technol.* **A4**(4), 1971, 1974.
Raman, C. V., and Nath, N. S.N. (1935a). *Proc. Indian Acad. Sci. Sect. A, Part I* **2**, 406-412.
Raman, C. V., and Nath, N. S.N. (1935b). *Proc. Indian Acad. Sci. Sect. A, Part II* **2**, 413-420.
Raman, C. V., and Nath, N. S.N. (1936a). *Proc. Indian Acad. Sci. Sect. A, Part III* **3**, 75-84.
Raman, C. V., and Nath, N. S.N. (1936b). *Proc. Indian Acad. Sci. Sect. A. Part IV* **3**, 119-125.
Raman, C. V., and Nath, N. S.N. (1936c). *Proc. Indian Acad. Sci. Sect. ... Part V* **3**, 459-465.
Roberts, S. G., Warren, P. D., and Hirsch, P. B. (1986). *J. Mater. Res.* **1**, 162.
Rogers, P. H., and Van Buren, A. L. (1974). *J. Acoust. Soc. Am.* **55**, 724-728.
Schwartz, L., and Bansil, A. (1975). *Phys. Rev.* **B10**, 3261.
Scott, W. R., and Bloomfield, P. E. (1981). *Ferroelectrics* **32**, 79-83. Lead attachment to KYNAR films is covered in this paper.
Seki, H., Granato, A., and Truell, R. (1956). *J. Acoust. Soc. Am.* **28**, 230-238.
Shahid, M. A., Mahajan, S., Laughlin, D. E., and Cox, H. M. (1987). *Phys. Rev. Lett.* **58**, 2567.
Sher, A., Chen, A.-B., and Spicer, W. E. (1985). *Appl. Phys. Lett.* **46**, 54.
Sher, A., van Schilfgaarde, M., Chen, A.-B., and Chen, W. (1987). *Phys. Rev.* **B36**, 4279.
Sher, A., Berding, M. A., van Schilfgaarde, M., Chen, A.-B., and Patrick, R. S. (1988). *J. Crystal Growth* **86**, 15.
Simondet, F., Michard, F., and Toquet, R. (1976). *Opt. Commun.* **16**, 411-416.
Slater, J. C., and Koster, G. F. (1954). *Phys. Rev.* **94**, 1498.
Spicer, W. E., Silberman, J. A., Morgan, P., Lindau, I., Wilson, J. A., Chen, A.-B., and Sher, A. (1982). *Phys. Rev. Lett.* **47**, 948.
Truell, R., Elbaum, C., and Chick, B. B. (1969). "Ultrasonic Methods in Solid State Physics." Academic Press, New York.
Tu, L. Y., Brennan, J. N., and Sauer, J. A. (1955). *J. Acoust. Soc. Am.* **27**, 550.
van Schilfgaarde, M. (1990), unpublished.
van Schilfgaarde, M., and Sher, A. (1987a). *Phys. Rev.* **B36**, 4375.
van Schilfgaarde, M., and Sher, A. (1987b). *Appl. Phys. Lett.* **51**, 175.
van Schifgaarde, M., Methfessel, M., and Paxton, A. T. 1991. In preparation.
Van Vechten, J. A. (1969). Phys. Rev. **187**, 1007.
Vasile, C. R., and Thompson, R. B. (1977). *In* "IEEE Procedings, 1977 Ultrasonics Symposium" (J. deKlerk and B. R. McAvoy, eds.), p. 84.
Vegard, L. (1921). *Z. Phys.* **5**, 17.
Voight, W. (1928). "Lehrbuch der Kristallphysik." Teubner, Leipzig. This contraction of the indices is also covered in Landau and Lifschitz (1986) and Green (1975), page 6. The subscripts are contracted as $11 \rightarrow 1$, $22 \rightarrow 2$, $33 \rightarrow 3$, $23 \rightarrow 4$, $31 \rightarrow 5$, $12 \rightarrow 6$.
Wei, S.-H., and Krahauer, H. (1985). *Phys. Rev. Lett.* **55**, 1200.
Wei, S.-H., Ferreira, L. G., and Zunger, A. (1990). *Phys. Rev.* **B41**, 8240.

Wei, S.-W., and Zunger, A. (1989). *Phys. Rev.* **B39**, 3279.
Wigner, E. P. (1934). *Phys. Rev.* **46**, 1002.
Williams, A. O., Jr., (1970). *J. Acoust. Soc. Am.* **48**, 285 (1970).
Williams, J., and Lamb, J. (1958). *J. Acoust. Soc. Am.* **30**, 308–313.
Wimmer, E., Krahauer, H., Weinert, M., and Freeman, A. J. (1981). *Phys. Rev.* **B24**, 864.
Wood, D. M., Wei, S.-W., and Zunger, A. (1988). *Phys. Rev.* **B37**, 1342.
Yeh, C.-Y., Chen, A.-B., and Sher, A. (1990). *Phys. Rev.* **B43**, 9138.
Zallen, R. (1982). In "Handbook on Semiconductors," (T. S. Moss, ed.), Vol. 1, Chap. 1. North-Holland, Amsterdam.

Ordering in GaAlAs

Srinivasan Krishnamurthy, A. Sher
SRI International, Menlo Park, CA-94025
and
A.-B. Chen
Auburn University, Auburn, AL-36849

Abstract

We calculated excess pair energies of GaAlAs alloys in the bulk and on the surface of a semi-infinite crystal in a generalized perturbation method using a third-neighbor tight-binding Hamiltonian parameters derived from an *ab-initio* density functional theory. This Hamiltonian has produced good electronic and structural properties for GaAs and AlAs. Calculated excess pair energy for the bulk alloy is nearly zero, suggesting a disordered equilibrium bulk state, and that for the Ga-Al pair on the (100) surface is found to be strongly negative. The σ bonds are found to be three times more attractive than the π-bonds, which strongly suggest the tendency of ordering in the CuAu I structure observed experimentally. From the magnitude of this attractive pair interaction the critical temperature for the order-disorder transition is estimated to be 900K.

Ordering in semiconductors is of both scientific and technological interest. The phase transition induced by strain, chemical energy differences, surfaces are of vigorous scientific study. Properties of alloys grown for materials engineering such as improved band gap, structural and electronic properties are altered, and mostly favorably, by ordering. For example, alloys grown for improved electron mobility, will exhibit further dramatic increase due to ordering.

(An Ban:: Please write here a para or two on status of experimental and theoretical research on ordering in semiconductors and difficulties to explain ordering GaAlAs)

Previous calculation of sublimation energy of surface atoms from two extreme surface coverages concluded that less energy was required to remove a Ga atom from a completed than from nearly empty (100) GaAs surface (Krishnamurthy et al., 1990a). This suggested a possibility of repulsive Ga-Ga surface interactions with Ga surrounded by vacancies as a lower energy state. The repulsive interactions were interpreted to be caused by charge transfer between dangling Ga and newly exposed As bonds. In such a case, if Al atoms are preferred at vacant sites due to its lower dangling bond level than Ga, then growth of ordered GaAlAs is possible by low temperature growth methods such as MBE. A preliminary calculations carried out with empirical band structures strongly suggested such a possibility (krishnamurthy et al., 1991).

The empirical band structures, even when they are carefully constructed for the cases studied, are limited in versatility and accuracy, while most of the existing first principles calculations are computationally too demanding to study realistic problems. The best and numerically tractable first principle theory is given by the local density approximation (LDA) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965). Widely used accurate methods to solve LDA equations either by plane wave method (Ihm and Cohen, 1980; Nielson and Martin, 1985; Wang et al., 1989) or by linearized muffin-tin orbitals (Anderson, Jepsen, and Sob, 1987; Methfessel and van Schilfgaarde, 1990, Christensen et al., 1990), and linearized augmented plane waves (Zunger and Freeman, 1977; Mbaye, Ferreira, and Zunger, 1987) are discouragingly time consuming for the approach considered here.

We use the linear combination of atomic orbitals (LCAO) method in a tight-binding form whose parameters are obtained from first-principles (Jenson and Sankey, 1987, 1989; Sankey and Niklewski, 1989; Chelikowsky and Louie, 1984; Vanderbilt and Louie, 1984). This Hamiltonian uses pseudo-atomic-orbitals, computed self-consistently from a free-atom calculation within LDA, as the basis. The bulk band structures and total energies are calculated with an assumption that the charge density in the solid is the superposition

of atomic charge densities(Harris, 1985). Due to charge transfer in the solid, the wave function in the III-V compounds are more localized than the corresponding atomic wave function. Linear combination of atomic function do not adequately describe the wave function in the solid. The basis functions are made more compact by multiplying with a function (Methfessel,1990)

$$\frac{1}{(e^{\beta}(r-r_0)+1)}$$

and renormalizing it. These compress parameters, β and r_0 have been varied to obtain a local maximum in cohesive energy. As a general rule of guidance, we start with twice the value of Pauling radius for r_0, and a value of 2.5 for β. The obtained electronic structure and structural properties such as equilibrium bond length, cohesive energy per bond, and bulk moduli are in excellent agreement (Krishnamurthy et al., 1992) with experiments and other calculations.

The studies of preferred phases during layer-by-layer crystal growth involves minimization free energy with various parameters such as alloy concentration, order parameter, temperature and so on. Free energy contains two parts, viz., enthalpy and entropy. Enthalpy is calculated from the excess energy necessary to form a cluster of two or more atoms. Entropy is obtained from the number of ways of arranging such a cluster. Many models are used to calculate these cluster energies (Zunger, Das sarma, Krishnamurthy). In this letter, we report an accurate evaluation of these energies with a mathematically rigorous formalism and first quantitative explanation for observed ordering in GaAlAs alloys.

We calculate the excess energies in an approach known as generalized perturbation method (GPM), which is computationally straightforward with TB Hamiltonian and is based on Green's function (GF) formalism. This method can efficiently calculate cluster energies of arbitrary size and has been successfully used in phase diagram studies(Ducastelle et al., 1976,Turchi et al., 1987). Its generalization to use first principles Hamiltonian with nonorthogonal basis is given here in brief.

With orthogonal basis set a Greens function G is conventionally defined such that $(E+i\delta-H)G=1$ is satisfied. When the overlap matrix S is different from unit matrix, the corresponding GF G^m will satisfy $(ES+i\delta-H)G^m=S$. It can be shown (JAM, Lohez+Lannoo) that G^m plays the same role as conventional G in the calculation of physical quantities such as density of states and total energy. For example, the density of states and total density of states are respectively $\frac{-1}{\pi}$ Im Tr(G^m) and $\frac{1}{\pi}$ Im Tr $(\ln(G^m))$. If we define an intermediate GF, G^k such that $(ES+i\delta-H)G^k=1$, the required G^m is

G^kS. We calculate the surface and bulk G^k by simply replacing (E-H) by (ES-H) in the difference equation method (Chen et al., 1989).

The calculation of excess energies in GaAlAs invariably involves clusters of Ga and Al. In GPM, the reference medium is random alloy and most suitable method of obtaining band structure of such an alloy is by the coherent potential approximation (CPA). Once a CPA-alloy is constructed, the cluster, whose excess energy is to be determined, is placed in the CPA medium. For example, if Ga-Al 'near-neighbor' pair energy is required, two near-neighbor CPA atoms are removed and Ga-Al pair is placed there. Then the new Hamiltonian is $H = \bar{H} + V$ where $\bar{H}$ is CPA Hamiltonian and V is the perturbation potential due to substitution.

When the change in S due to substitution is neglected, new GF using Dyson expansion, is

$$\begin{aligned} G^k &= \bar{G}^k + \bar{G}^k V G^k \\ &= (1 - \bar{G}^k V)^{-1} \bar{G}^k \end{aligned} \tag{1}$$

multiplying both sides of Eq.(1) by S,

$$\begin{aligned} G^m &= \bar{G}^m + \bar{G}^k V G^m \\ &= (1 - \bar{G}^k V)^{-1} \bar{G}^m \end{aligned} \tag{2}$$

Then the change electronic energy,

$$\begin{aligned} \Delta E_{ele} &= -\int_{-\infty}^{\epsilon_F} \frac{1}{\pi} ImTr(ln(G^m) - ln(\bar{G}^m)) dE \\ &= \frac{1}{\pi} ImTr \int_{-\infty}^{\epsilon_F} ln(1 - \bar{G}^k V) dE \end{aligned} \tag{3}$$

To simplify further, we write

$$\begin{aligned} (1 - \bar{G}^k V) &= (1 - \bar{G}^k_d V)(1 - t\bar{G}^k_{nd}) \\ t &= V(1 - \bar{G}^k_d V)^{-1} \end{aligned} \tag{3}$$

where $\bar{G}^k$ is sum of diagonal ($\bar{G}^k_d$) and non-diagonal ($\bar{G}^k_{nd}$) in site index. From Eq. (2) and (3) and using the expansion that ln(1-x) = $-\sum_n \frac{x^n}{n}$, we get the expression in the familiar form,

$$\Delta E_{ele} = \frac{1}{\pi} ImTr \int ln(1 - G_d^k V) - \frac{1}{\pi} ImTr \sum_m^{\infty} \int \frac{tG_{nd}^k tG_{nd}^k \cdots}{m}$$
$$\equiv h_0 + \sum_{ijk..} v_{ijk..} t_i t_j t_k \cdots \quad (4)$$

If the presence of both bulk and surface are considered, the h_0 is approximately the difference in site term values and is often a dominant mechanism for surface segregation. h_0 is zero when we consider only surface or only bulk atoms. All constant terms can be conveniently removed from the free energy calculations. $v_{ijk..}$ are the cluster energies that determine the phase diagrams. Because the excess energies are calculated with reference to random alloy, just from the *sign* of $v_{ijk..}$ for various cluster energies, one can immediately write-down the phase at T=0 K (Turchi et al., 1987). A minus sign for unlike-atom pair energies imply ordering and a plus sign for the same pair imply segregation and so on. However, for higher temperatures the distribution of these clusters should be considered. The pair interaction energy, for example, between Ga and Al at sites n and m respectively is,

$$v_{nm}^{Ga-Al} = -\frac{1}{2\pi} ImTr \int t^{Ga} t^{Al} \bar{G}_{nm}^k \bar{G}_{mn}^k \delta C_n^{Ga} \delta C_m^{Al} \quad (5)$$

where $\delta C_n^{Ga/Al}$ is the change in the concentration at site n due to replacement of a CPA atom by Ga or Al. A factor of 2 should be included in the cases where spin is not explicitly considered. It should be noted that Eq.(4) and Eq.(5) appear much similar to previously obtained results (Ducastle et al., 1976). However, contrary to intuition, conventional $\bar{G}$ is replaced by $\bar{G}^k$ and *not* by $\bar{G}^m$.

We find that that the term values and interaction matrix elements of surface atoms are much different (as much as 2 eV in some cases) from the corresponding bulk values. In such cases, The surface GF is obtained in two step. First by assuming bulk term values on the surface, the difference equation method used to obtain the G^k. Then the difference in term values are considered as local perturbation and the final GF is obtained with Dyson's expansion. With different self energies for bulk and surface atoms, CPA equations are solved. This self consistent procedure requires some generalization, due to nonorthogonality, and it has been included. An important observation is that the surface states occur in the gap when surface atoms are treated same as bulk atoms. However, when the difference in term values are included, the surface states are pushed into the valence bands, justifying the use of LDA studies of surface related phenomena. Incorrect conduction bands do not affect the total energy results.

Above described procedure is used to calculate pair energies in the bulk and on (100), (111)B surfaces of $Ga_{0.5}Al_{0.5}As$. In the case bulk $Ga_{0.5}Al_{0.5}As$, the pair interaction between Al-Al and Ga-Ga are attractive and that between Ga-Al are repulsive with excess pair energy of about 1 meV. This tendency towards randomness had been suggested by other calculations (Zunger) and experiments(?). Similar calculations are carried out on (111)B surface of $Ga_{0.5}Al_{0.5}As$. The excess pair energy is zero within the accuracy of our calculations. This energy has to be less than zero (in our sign convention) to obtain ordering. Experiments on (111)B surface of this alloy do not see ordering(Kuan, 1985).

On the (100) surface, there are two kinds of surface bonds. The surface atoms connected by σ bonds are also connected by first neighbor As on the layer below. The π bonds connects the atoms that are connected also by third-neighbor As atoms on the layer below. The excess pair energies which are nearly zero in the bulk, reduces substantially and become strongly attractive for unlike-pairs on the surface. The pair energy for σ bonds is -100 meV and that for π bonds -30 meV. The pair energy for the like pairs have the same magnitude but repulsive. Such an interaction will yield to ordered growth where every Ga(Al) will be surrounded by Al(Ga) in CuAg-I phase. The energy parameter, v is defined as $(v_{Ga-Al} - \frac{(v_{Ga-Ga}-v_{Al-Al})}{2})$. Then using average value of the π and σ bond energies and the exact expression of $0.57v/k_B$ (k_B is Boltzmann constant) for the critical temperature for order-disorder transition on a square lattice (Onsagar, 1944), we get a value of 860 K. This is in excellent agreement with experiment where CuAg-I ordering is observed during the growth on (100) direction at 600 K (Kuan et al., 1985). In the same experiment, no ordering was observed on growing (111)As surface, also consistent with our calculations.

The cause for attractive interaction between unlike pairs is not transparent from these calculations as the surface states are dispersed well in the valence bands. However, with a fitted second-neighbor TB Hamiltonian, one can see that Al dangling states are lower in energy than Ga dangling states. Thus by charge transfer from Ga to 'near-neighbor' Al, surface can lower energy, leading to preference for Ga-Al pairs over other pairs. However, the energy gained should be larger than Coulomb energy associated with charge transfer. For pairs beyond near-neighbor sites, Coulomb energy will be larger as the charge transfers over larger distance. In addition, in the absence of Al, Ga-vacancy pairs cause lower energy as the charge transfer occur between Ga and much deeper As dangling bonds. As AlAs bonds are slightly stronger than GaAs bonds (cohesive energy -1.84 ev vs -1.64 ev), filling those vacancies with Al, rather than Ga, will lower energy. Such advantages are not available for the growth in (111)B direction, as the surface will always be As terminated. Our quantitative evaluation of excess energies and critical temperature are consistent with

these plausible physical mechanisms and explain the observed ordering.

In conclusion, we carried out first-principles studies of cluster energies in bulk and on surfaces. The calculated pair energies on (100) GaAs surfaces, indicate a strongly attractive interaction for Ga-Al pairs and repulsive interaction for Ga-Ga or Al-Al pairs. The magnitude and sign of these interactions yield a critical temperature of 860 K and in excellent agreement with experiments. The charge transfer between the neighboring pairs are believed to cause attractive interaction for unlike atoms leading to ordering.

This work was supported by ONR contract N00014-89-K132 and by NASA contract NAS1-18226-11. A part of the work was done while SK was at Max Planck Institut für festkörperforschung, Stuttgart. He thanks Prof. M. Cardona for his hospitality and Alexander von Humboldt foundation, Bonn, Germany for the fellowship.

VI. REFERENCES

Andersen, O.K., O. Jepsen, and M. Sob, 1987: *Electronic band structures and applications*, M. Yussouf, eds. (Springer Lecture Notes).

Chelikowsky, J.R., and S. Louie, 1984: Phys, Rev. B **29**, 2470.

Chen, A.-B., Y.M. Lai-Hsu, and W. Chen, 1989: Phys. Rev. B **39**, 923.

Ducastelle, F., and F. gautier, 1976. J. Phys. F, **6**, 2039.

Harris, J., 1985: Phys. Rev., **31**, 1770.

Hohenberg, P.C.. W. Kohn, 1964: Phys. Rev., **136**, B864.

Ihm, J., and M.L. Cohen, 1980: Phys. Rev. B **21**, 1527.

Jensen, R.W., and O.F. Sankey, 1987: Phys. Rev. B **36**, 6520.

Jensen, R.W., and O.F. Sankey, 1989: Phys. Rev. B **39**, 3192.

Kohn, W., and L. Sham, 1965: Phys. Rev., **140**, A1133.

Krishnamurthy, S., M.A. Berding, A. Sher, and A.-B. Chen, 1900a: Phys. Rev. Lett. **64**, 2531.

Krishnamurthy, S., A. Sher, and A.-B. Chen, 1991: Bull. Am. Phys. Soc., **00**, 000.

Krishnamurthy, S., A. Sher, and A.-B. Chen, 1992b: in preparation

Kuan.....1985:

Mbaye, A.A., L.G. Ferreira, and A. Zunger, 1987: Phys. Rev. Lett., **58**, 49.

Nielsen, O.H., and R.M. Martin, 1985: Phys. Rev., B **32**, 3792.

Onsagar, L, 1944: Phys. Rev. **65** 117.

Sankey, O.F., D.J. Niklewski, 1989: Phys. Rev., **40**, 3979.

Turchi, P., M. Sluiter, and D. de Fontaine, 1987: Phys. Rev., **36**, 3161.

van Schilfgaarde, M., and M. Methfessel, 1990: Bull. Am. Phys. Soc., **35**, 667.

Vanderbilt, D., and S. Louie, Phy. Rev. B **30**, 6118.

Zunger, A., and A.J. Freeman, 1977: Phys. Rev., B. **15**, 5049.

Comparison of $In_{1-x}Tl_xSb$ and $Hg_{1-x}Cd_xTe$ as Long Wavelength Infrared Materials

A.-B. CHEN

Physics Department, Auburn University, AL 36849

M. VAN SCHILFGAARDE and A. SHER

SRI International, Menlo Park, CA 94025

Cohesive energies, elastic constants, band structures, and phase diagram are calculated to evaluate the $In_{1-x}Tl_xSb$ alloy (ITA) as a long-wavelength infrared (LWIR) material compared to $Hg_{1-x}Cd_xTe$ (MCT). To obtain a 0.1 eV gap at zero temperature, the x value for ITA is estimated to be x = 0.083 as compared to x = 0.222 for MCT. At this gap, ITA is more robust than MCT because the cohesive energies order as InSb>TlSb>CdTe>HgTe, and ITA has the stronger bonding InSb as the majority component. Although TlSb is found to favor the CsCl structure, ITA is a stable alloy in the zincblende structure for low x values. However, our phase diagram indicates that it is difficult to grow the 0.1 eV gap ITA from the melt, because above the eutectic the liquidus curve is flat, and the solidus drops rapidly. Moreover, the width of the stable concentration range of the zincblende solid phase shrinks at low temperatures due to the presence of the CsCl structure.

Key words: Elastic constants, InTlSb, HgCdTe, long-wavelength infrared materials, phase diagram

INTRODUCTION

As growth technology continues to progress, researchers are always in pursuit of better semiconductors. In the area of long-wavelength infrared (LWIR) semiconductors, efforts have been directed into three fronts: refinement of the $Hg_{1-x}Cd_xTe$ (MCT) technology, search for alternative semiconductors, and use of superlattices and quantum-well structures.[1] Among alternative LWIR semiconductors, obvious III-V candidates are alloys of InSb with TlSb or InBi. An energy-assisted epitaxial growth technique[2] has been used to obtain the $InSb_{1-x}Bi_x$ alloy with a gap in the LWIR range (8 to 12 μ), but that method has never evolved into a practical device technology. Thus it is interesting to examine the prospects of the $In_{1-x}Tl_xSb$ (ITA) alloy.

We have calculated[3] the structural properties, band structures, and phase diagram to evaluate ITA as an infrared (IR) material as compared to MCT. At the same gap for LWIR application, the band structure of ITA is similar to that of MCT, but ITA is a more robust material. The phase diagram of ITA is more complicated than MCT, because of the presence of a stable CsCl structure in TlSb. Our phase diagram indicates that it is difficult to grow ITA from the melt. However, the zincblende solid solution of ITA is stable at low thallium compositions, which may permit the growth of ITA at the desired concentration using epitaxial growth methods. These studies show that ITA is a potential but challenging IR material deserving more attention.

(Received October 12, 1992; revised January 13, 1993)

LATTICE CONSTANT, COHESIVE ENERGY, AND ELASTICITY

Table I shows the cohesive energies, atomic volumes, and some elastic constants of CdTe, HgTe, InSb, and TlSb calculated[3] from the full-potential linear muffin-tin orbital method[4] (FP-LMTO) within the local density approximation (LDA). The calculated atomic volumes and elastic constants are in good agreement with available experimental values. The calculated cohesive energies are consistently larger than the experimental values by about 0.9 eV, which is typical in LDA owing mainly to errors in the free atom. However, the relative cohesive energies

Table I. Calculated vs Experimental Values of Atomic Volume (V), Cohesive Energy (E), Shear Modulus (c_{11}-c_{12}), and Bulk Modulus (B)

	CdTe		HgTe		InSb			TlSb	
	ZB*	Expt	ZB*	Expt	ZB	Expt	CsCl†	ZB	CsCl†
V (Å³)	33.4	34.2	34.2	33.4	34.1	33.8	25.6	36.4	27.1
E (eV/pr. of atoms)	5.28	4.40	4.24	3.24	6.46	5.60	6.08	5.70	5.79
c_{11}-c_{12} (10^{12} erg/cm³)	0.20	0.17	0.19	0.18	0.31	0.31	—	0.17	—
B (10^{12} erg/cm³)	0.44	0.42	0.44	0.47	0.45	0.48	0.72	0.38	0.49

*zincblende structure
†CsCl structure

Table II. Direct Band Gaps (in eV) of Several Semiconductors Calculated in the Local Density Approximation (LDA) and Comparison with Experimental Gaps (Expt)

	CdTe	GaAs	InP	InAs	InSb	HgTe	TlSb
LDA	0.54	0.44	0.47	–0.39	–0.46	–0.95	–2.20
LDA-SO	0.23	0.32	0.42	–0.53	–0.73	–1.28	(–2.50)
Expt	1.60	1.52	1.42	0.42	0.26	–0.31	(–1.52)
Underest.	1.37	1.20	1.00	0.95	0.99	0.97	(0.98)

Note: Systematic LDA errors are used to deduce the bandgap for TlSb. The numbers inside the parentheses are the estimated values

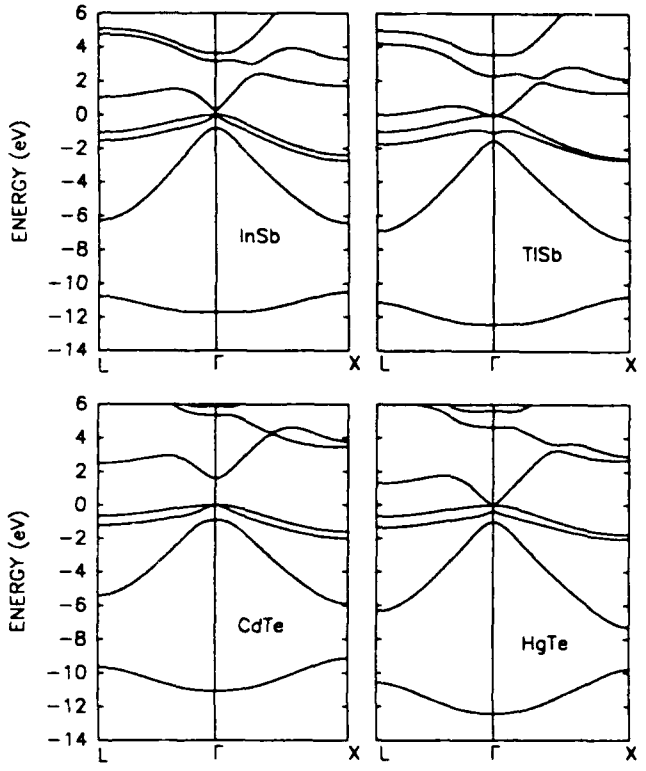


Fig. 1. The parametrized band structures for InSb, TlSb, CdTe, and HgTe along the Λ (Γ-L) and Δ (Γ-X) directions.

are reliable. Our result shows that TlSb slightly favors the CsCl over the zincblende structure at zero temperature. In the zincblende structure, the cohesive energy magnitudes order as InSb > TlSb > CdTe > HgTe. Although TlSb has a weaker shear modulus than either HgTe or CdTe, InSb is stronger than both. As will be shown latter, InSb is the majority component in ITA for LWIR applications. These results combined imply that ITA is more robust structurally than MCT as a LWIR material.

BAND STRUCTURE

Although LDA is known to produce band gaps for semiconductors that are too small, the gap for TlSb can be deduced by exploiting the systematic nature of LDA errors as shown in Table II. The first row shows the band gaps calculated from LDA for several direct-gap semiconductors. These values do not include the spin-orbit coupling. The second row shows the LDA values with the spin-orbit (SO) correction added. The third row lists the available experimental values. The fourth row lists the underestimated values, ranging from 0.95 to 1.37 eV, in LDA. CdTe, and to a certain extent GaAs, has a larger underestimate owing to its smaller dielectric constant. TlSb should have a dielectric constant close to InSb and HgTe, so an average of the underestimates of the two, i.e. 0.98 eV, can be used for TlSb to obtain an energy gap of –1.52 eV.

This inversion to a negative gap can also be estimated and understood in terms of tight-binding concepts. Firstly, due to a stronger scalar relativistic potential, the s-level of thallium is about 1.4 eV deeper than that of indium, which lowers the conduction minimum of TlSb. Secondly, the valence d-level of thallium is higher than that of indium by about 3 eV, and is only 10 eV below the top of the valence band. In the crystal, the cation d-level interacts with the neighboring p-state of antimony to push up the valence band maximum (VBM). This p-d repulsion raises the VBM in TlSb by 0.3 eV higher than that in InSb. Starting with a 0.25 eV gap for InSb, we subtract from this value the 0.3 eV from the p-d interaction, the 1.4 eV of the s-shift, and 0.05 eV due to a larger spin-orbit

coupling to arrive at a gap of -1.5 eV for TlSb. This p-d interaction plus the spin-orbit coupling also produces a 0.35 eV valence band offset between TlSb and InSb. The same mechanism is also responsible for the band inversion and valence band offset between HgTe and CdTe.

To obtain more details, we have used a hybrid psudopotential tight-binding method[3] to parametrize the band structures for the hosts and calculate the band structures for the alloys. The band structures for the four hosts along Λ (Γ-L) and Δ (Γ-X) directions are plotted in Fig. 1. The gaps for these band structures are 0.25, -1.50, 1.60, and -0.31 eV, respectively, for InSb, TlSb, CdTe, and HgTe. Based on a scaled-virtual crystal approximation (SVCA),[5] the calculated band gaps for both MCT and ITA at zero temperature are plotted in Fig. 2 as a function of the alloy concentration x. The predicted x values for the 0.1 eV gap at zero temperature are x = 0.083 for ITA and, in close agreement with experiment, x = 0.222 for MCT. The band structures in the vicinity of the band gap for ITA and MCT at the 0.1 eV gap are compared in Fig. 3. We can see these two alloys have very similar band structures at the band gap. This result implies that ITA has electrical and optical properties similar to MCT for LWIR applications.

PHASE DIAGRAM

The phase diagram of $Hg_{1-x}Cd_xTe$ is very simple;[6] one in which the liquidus-solidus curves have a simple lens shape and the zincblende solid solution is completely miscible except at very low temperatures. The presence of a stable CsCl solid phase in TlSb considerably complicates the phase diagram of $In_{1-x}Tl_xSb$. Because of the small lattice mismatch between InSb and TlSb, the free energies of the solid solutions in both zincblende and CsCl structures can be approximated by a regular solution model given by $F(x) = E(x)+kT[x\ln x +(1-x)\ln(1-x)]$, and the internal energy can be written as $E(x) = (1-x)E(0)+xE(1)+x(1-x)\Omega$. The end point energies are given in Table I. The mixing enthalpies Ω are calculated within LDA using quasi-random structures[7] and the results[3] give Ω_{ZB} = 32 meV per two atoms for the zincblende structure and Ω_{CsCl} = -4 meV for the CsCl structure. These two free energy functions produce a large miscibility gap separating the zincblende (ZB) alloy from the nearly pure TlSb crystal in the CsCl structure as shown in Fig. 4. Also shown are the solidus-liquidus curves calculated from the regular solution model. In addition to the above data for the solid solution, the following data were adopted for the liquid phase: mixing enthalpy Ω = -1000 cal/mole, melting temperatures of 808 and 558K, and enthalpies of melting of 11878 and 8629 cal/mole respectively for InSb and TlSb. The Ω value is taken to be the value previously used for MCT,[8] and the other values for the ZB TlSb are extrapolated from those of InSb[9] assuming they scale as Ω(TlSb)/Ω(InSb) = Ω(HgTe)/Ω(CdTe). Our phase diagram shows that it is very difficult to grow the 0.1 eV (x = 0.083) ITA from the melt, because above the eutectic the liquidus curve is rather flat, and the solidus drops rapidly. Moreover, the width of the stable concentration range of the zincblende solid

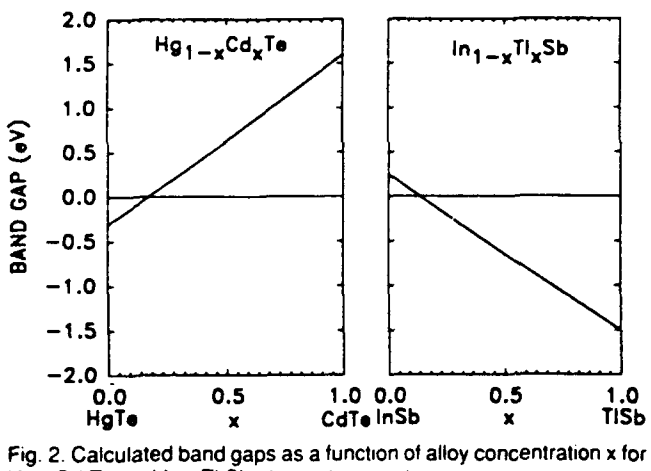


Fig. 2. Calculated band gaps as a function of alloy concentration x for $Hg_{1-x}Cd_xTe$ and $In_{1-x}Tl_xSb$ at zero temperature.

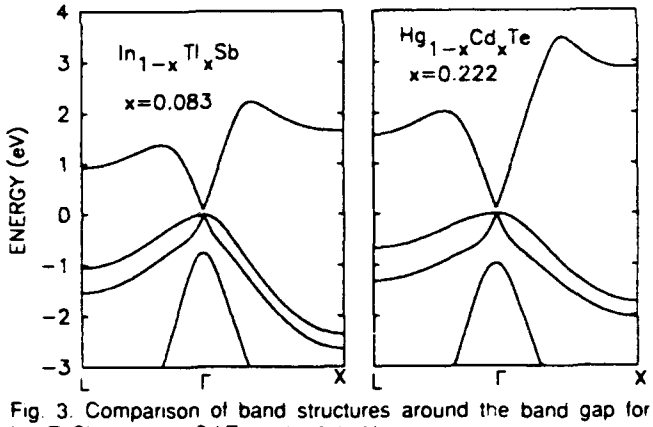


Fig. 3. Comparison of band structures around the band gap for $In_{1-x}Tl_xSb$ and $Hg_{1-x}Cd_xTe$ at the 0.1 eV gap.

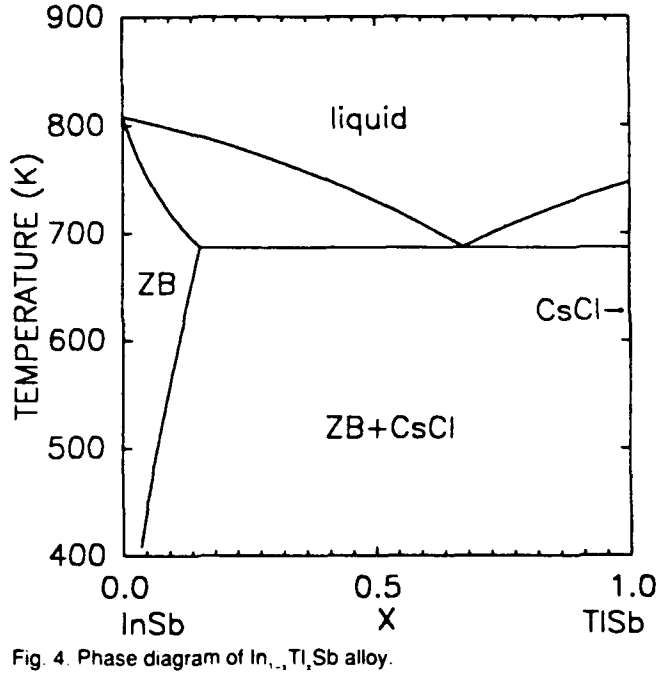


Fig. 4. Phase diagram of $In_{1-x}Tl_xSb$ alloy.

phase shrinks at low temperatures due to the presence of the CsCl structure. However, the ZB solid ITA solution is stable at low x values and moderate temperatures. A metastable alloy at the desired concentration may be obtained using epitaxial growth techniques.

CONCLUSION AND DISCUSSION

Several properties of ITA have been studied. Our results show that, when compared to MCT as LWIR material, ITA is superior in structural bonding, comparable in electrical and optical properties, but more difficult to grow. However, our work represents the first reasonable materials evaluation. The predicted phase diagram depends heavily on the accuracy of the energy difference between the CsCl and the zincblende structures of both InSb and TlSb, and to a less extent, on the other parameters used. Further studies are called for to test these suggestions and to explore the potential of this LWIR material.

ACKNOWLEDGMENT

We wish to thank J.M. MacCallum, G. Cinader, A. Sanjurjo, and M.A. Berding for their helpful suggestions. The support by ONR Contract N00014-89K-0132 and NASA Contract NAS1-18226 is gratefully acknowledged.

REFERENCES

1. B.J. Levine, C.G. Becthea, K.G. Glogovsky, J.W. Stayt and R.E. Leibenguth, *Semicond. Sci. Technol.* 6, C114 (1991).
2. J. Zilko and J.E. Green, *Appl. Phys. Lett.* 33, 254 (1978); *J. Appl. Phys.* 51, 1549 (1980).
3. M. van Schilfgaarde, A. Sher and A.-B. Chen, submitted for publication.
4. M. Methfessel and M. van Schilfgaarde, unpublished.
5. A.-B. Chen and A. Sher, *Phys. Rev.* B23, 5360 (1981).
6. R.S. Patrick, A.-B. Chen, A. Sher and M.A. Berding, *J. Vac. Sci. Technol.* A 6(4), 2643 (1988) and references therein.
7. S.-H. Wei, L.G. Ferreira, J.E. Bernard and A. Zunger, *Phys. Rev.* B42, 9622 (1990).
8. F.R. Szofran and S.L. Lohoczky, *J. Electron. Mater.* 10, 1131 (1981).
9. G.B. Stringfellow and P.E. Green, *J. Phys. Chem. Solids* 30, 1779 (1969).

InTlSb: An infrared detector material?

Mark van Schilfgaarde and Arden Sher
SRI International, 333 Ravenswood Ave., Menlo Park, California 94025

An-Ban Chen
Physics Department, Auburn University, Alabama 36849

(Received 27 July 1992; accepted for publication 25 January 1993)

$In_{1-x}Tl_xSb$ is proposed as promising infrared material. A number of optical and structural properties are studied within local density-functional theory. The alloy at $x=0.09$ is estimated to have a gap of 0.1 eV. Although TlSb is found to favor the CsCl structure, the zinc blende alloy is stable for low x values. A phase diagram is calculated to estimate the regions of stable phases and explore the conditions for growing narrow-gap $In_{1-x}Tl_xSb$ alloys.

The alloy $Hg_{1-x}Cd_xTe$ with $x=0.22$ is currently the most widely used material for long-wave infrared (LWIR, 8–12 μm) focal-plane arrays (FPA). However, $Hg_{1-x}Cd_xTe$ is a weakly bound II-V compound with materials growth and processing problems[1] dominated by native defects that limit FPA performance and yield. A sturdier alternative to $Hg_{1-x}Cd_xTe$ offering a comparable IR response would significantly benefit FPA costs and performance. Here we argue that the III-V alloy $In_{1-x}Tl_xSb$ holds much promise as a suitable, sturdier alternative to $Hg_{1-x}Cd_xTe$.

InSb, a III-V zinc blende (ZB) semiconductor with a band gap of 0.18 eV at room temperature, responds to the midwave (3–5 μm) infrared band and is usually preferred[2] to $Hg_{0.69}Cd_{0.31}Te$ with a comparable band gap, because InSb is the more strongly bound. However, the band gap of InSb is too wide for LWIR applications. Two alternatives are currently under study to produce III-V based materials capable of LWIR response: strained layer superlattices[3] and multiple quantum well structures.[4] These materials must be prepared either by metalorganic vapor deposition (MOCVD) or by molecular beam epitaxy (MBE). Both methods are expensive and while advances have been demonstrated, it is not clear that the individual device performance, uniformity over an array, and material stability will better those of $Hg_{1-x}Cd_xTe$.

InSb has the narrowest band gap of the ZB semiconductors. Alloying the Sb with Bi,[5] or alloying the In with Tl can potentially narrow the band gap. In its ground state, InBi is a tetragonal structure, which limits concentrations over which the equilibrium alloy is miscible.[6] Concentrations high enough to narrow the gap into the LWIR range have been prepared, but only using a difficult Ar energy-assisted epitaxial growth technique[7] that never evolved into a practical method for a LWIR material.

Very little work has been reported on the $In_{1-x}Tl_xSb$ alloy. This letter is devoted to calculation of the materials properties of $In_{1-x}Tl_xSb$ within the framework of density-functional theory. Questions about the properties of $In_{1-x}Tl_xSb$ center around two key issues. First, stipulating that the alloy assumes a ZB lattice with either Tl or In occupying the cation sites, how do the structural and optical properties compare with $Hg_{1-x}Cd_xTe$? Second, our calculations show that TlSb slightly prefers the CsCl lattice to the ZB. Under what thermodynamic conditions can $In_{1-x}Tl_xSb$ be grown in a ZB structure?

For the calculations presented here, except where otherwise stated, an all-electron, soft core, full-potential method of linear muffin tin orbitals (LMTO) was employed.[8] A basis of 22 orbitals/atom was used, including three s, three sets of p, and two sets of d states. This was sufficient to converge the total energy to an absolute precision of about 1 mRy/atom. In all cases, enough k-points were used to converge the mesh to an absolute precision of better than 0.1 mRy; this required between 30 and 128 points, depending on the crystal structure. Thus, the results presented here should be very close to exact solutions of the local density approximation (LDA).

Table I shows some calculated properties of CdTe, HgTe, InSb, and TlSb. Experimental data tabulated with the first three give the reader a sense of the errors to be expected from these calculations. As is typical in the LDA, the atomic volume is predicted to within about 2%; the cohesive energy is overestimated by about 0.9 eV (owing mainly to errors in the free atom; gradient corrections to the LDA largely eliminate this error); and the band gap is underestimated by approximately 0.7 eV, the underestimate larger for CdTe owing to its smaller dielectric constant. Exploiting the systematic trends in errors for the band gap and cohesive energy, column "Adj" adjusts those calculated values for TlSb, were it to assume a ZB lattice.

Table I illustrates several points of interest. TlSb is lattice-matched to InSb to within 2%: it has a comparatively strong bond, considerably stronger than in HgTe. Its shear modulus $c_{11}-c_{12}$ is rather small for a III-V material, possibly indicative of its preference for the CsCl structure (see below). Probably the most important point concerns the band gaps. The difference $E_g(\text{CdTe})-E_g(\text{HgTe})=1.9$ eV, is comparable to $E_g(\text{InSb})-E_g(\text{TlSb})=1.8$ eV. Assuming a linear dependence on the gap, $Hg_{1-x}Cd_xTe$ has a 100 meV gap at $x=0.22$, while for $In_{1-x}Tl_xSb$ the gap reaches 100 meV at $x=0.09$ at 0 K. (At 300 K InSb has a gap of 0.18 eV and it reaches 100 meV at $x=0.05$.) Thus the predominant cation in $Hg_{1-x}Cd_xTe$ is the heavier, weakening Hg, while in $In_{1-x}Tl_xSb$ it is the lighter and strengthening In. The cohesive strength of $In_{1-x}Tl_xSb$ will be about 50% greater than $Hg_{1-x}Cd_xTe$; as a consequence, we expect it to be far more stable and freer of defects. Because the material is more robust, processing should be

 0003-6951/93/161857-03$06.00

TABLE I. Calculated and experimental optical and structural properties for CdTe, HgTe, InSb, and TlSb in a ZB lattice at 0 K. Atomic volume Ω is in Å^3, cohesive energy per cation E and band gap E_g are in eV; $c_{11}-c_{12}$ and bulk modulus B are 10^{12} erg/cm^3. Calculated values for InSb in the CsCl and TlSb in both ZB and CsCl lattices are shown. Column Adj adjusts the calculated values of TlSb in the ZB structure for errors in the LDA.

	CdTe		HgTe		InSb			TlSb		
ZB	Calc	Expt	Calc	Expt	Calc	Expt	CsCl	ZB	Adj	CsCl
Ω	33.4	34.2	34.2	33.4	34.1	33.8	25.6	36.4		27.1
E	5.28	4.40	4.24	3.24	6.458	5.60	6.078	5.704	4.8	5.792
$c_{11}-c_{12}$	0.20	0.17	0.19	0.18	0.31	0.31		0.17		
B	0.44	0.42	0.44	0.47	0.45	0.48	0.72	0.38		0.49
E_g	0.54	1.60	−0.95	−0.30	−0.46	0.26		−2.20	−1.5	

much simpler. For example, we anticipate no significant problem of Tl loss in an annealing step.

The valence band offset is one other point of interest concerning optical properties of $In_{1-x}Tl_xSb$. A simple estimate can be made in the context of the atomic spheres approximation (ASA). Because the ASA aligns the bands with respect to a common reference potential (the Madelung potential), it forms a good trial potential for an interface between similar materials, and the dipole with respect to this potential is generally small. If the dipole is zero, the band offset of two materials is the difference in the valence band edges, calculated from separate bulk calculations. At the lattice parameter of InSb, the band offset of TlSb/InSb was calculated to be 0.35 eV. As a comparison, the band offset of HgTe/CdTe comes out to 0.30 eV in the ASA. In both cases, the origin of the band offset arises from coupling of the deep d states to the anion p orbitals at Γ. The heavier cation is shallower: it couples more strongly and exerts a stronger repulsion on the valence band maximum.

Next we turn to an examination of the $In_{1-x}Tl_xSb$ ZB alloy. Generally speaking, ZB alloys $A_{1-x}B_xC$ have a disposition to spinoidal decomposition into the bulk constituents $(AC)_{1-x}+(BC)_x$. But when the AC and BC lattices match well as in HgCdTe and AlGaAs, strain energy from accommodating bond-length mismatch is absent and the energy cost in forming the alloy is quite small. Indeed we find that to be the case in $In_{1-x}Tl_xSb$.

To show this for the $In_{0.5}Tl_{0.5}Sb$ alloy, three structures were considered: the "CuAu" structure,[9] and two quasirandom structures proposed by Zunger and co-workers which simulate a random alloy (see Table II and Ref. 9). In the ideal lattice, the energies of the three structures are found to be identical to within 1 meV/cation. In the ideal lattice these structures are found to be 12 meV less stable than the VCA average; on allowing the lattice to relax in the CuAu and the 4-cation quasirandom structure, this difference dropped to 8 meV/cation. The smallness in deviation from the VCA average and the independence of the energy on crystal structure suggests that the effective alloy Hamiltonian is weak and well approximated by a simple Ising model. The internal energy of a random alloy then assumes a form

$$E(x)=(1-x)\cdot E(x=0)+x\cdot E(x=1)+\Omega x(1-x) \quad (1)$$

with $\Omega=\Omega_Z=32$ meV.

We now turn to the second set of issues, namely the miscibility of Tl in $In_{1-x}Tl_xSb$ in the ZB structure. This miscibility is limited in InSb by the apparent fact that TlSb prefers the CsCl lattice to the ZB. Table I shows that, while InSb favors the ZB lattice by 380 meV/cation, TlSb favors the CsCl lattice by 88 meV/cation. To examine this in

TABLE I. Structural properties of three ordered alloys of $In_{0.5}Tl_{0.5}Sb$. The QR-4 and QR-8 structures are 4- and 8-cation superlattices taken from Zunger and co-workers.[a] Short-range correlation functions of QR-4 and QR-8 mimic those of the random alloy, and are named "quasirandom" structures. VCA is the virtual crystal average of this table. The energy of each structure was minimized with respect to the cell volume. E' is the energy of the structure on an ideal lattice; E'' is the energy when the lattice is allowed to relax.

	VCA	CuAu	QR-4	QR-8
a	6.554	6.554	6.554	6.554
E'	6.082	6.070	6.070	6.070
E''	6.082	6.074	6.074	

[a]Reference 9.

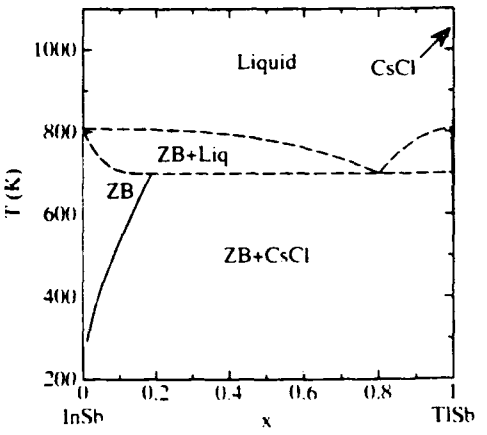


FIG. 1. Phase diagram of InTlSb along the $In_{1-x}Tl_xSb$ line. The shaded region is the desired solid solution of $In_{1-x}Tl_xSb$ in the ZB structure. The solid (changing into dotted) lines demarcate the separation of the ZB solid solution of $In_{1-x}Tl_xSb$ on the left, a very narrow solid solution in the CsCl structure on the extreme right (see arrow), and a two-phase eutectic of these two phases (marked "ZB+CsCl") in the center. (These lines are shown as dotted above the speculated melting point, as they are no longer meaningful there.) The dashed lines are our speculated liquidus and solidus; the eutectic point at $x=0.8$, $T=700$ K was chosen arbitrarily. The region marked "ZB+Liq" is a two phase mixture of $In_{1-x}Tl_xSb$ in the ZB phase and liquid. It is evident from this figure that, even while there exists a stable phase of $In_{1-x}Tl_xSb$ in the ZB structure, the only accessible phase from the liquid is the eutectic ZB+CsCl.

greater detail, we have calculated the total energies of some ordered alloys of $In_{1-x}Tl_xSb$ in the CsCl structure in addition to those studied in the ZB structure (Table I). As in the ZB structure, compounds studied in the CsCl structure exhibited a very small deviation in the mixing energy from the virtual crystal average; the deviation was found to be slightly negative, but smaller than in the ZB structure. Thus Eq. (1) also approximately describes the x dependence of the internal energy in the CsCl structure, with $\Omega=\Omega_C\simeq -4$ meV. Using these data, we have obtained that section through the ternary phase diagram which joins the InSb and TlSb binary compositions.

To describe the miscibility of the ZB and CsCl structures, we approximate the free energy of each structure with a regular solution model. The free energy of either ZB phase is then $F(x)=(1-x)E(0)+xE(1)+\Omega x(1-x)+kT[x\ln x+(1-x)\ln(1-x)]$, with $\Omega=\Omega_Z$ or Ω_C. By constructing common tangent lines between the two free-energy curves in the usual way, we obtain the solid phase portion of the phase diagram that exhibits three regions. As Fig. 1 shows, one region is found in the ZB structure near the InSb composition, and a corresponding region very near to the TlSb composition in the CsCl structure. A two-phase eutectic occupies the central and largest part of the diagram.

Also shown in dotted lines are the hypothetical solidus and liquidus. Explicit calculation of this portion of the phase diagram was not attempted here, particularly since we do not know the melting point of TlSb. Some general observations can be made, however. InSb is known to melt congruently at 808 K, which fixes the $x=0$ point of the liquids. Also our calculations show that the bulk modulus of TlSb in the CsCl structure is comparable to InSb in the ZB, owing to the compressed atomic volume. According to Lindeman's law, melting takes place when lattice distortions reach about 10% of the bond length. Crudely speaking, the vibrational amplitudes of InSb(AB) and TlSb(CsCl) are comparable, and thus we speculate that the melting points may also be comparable. It is almost certain, however, that the solidus near $x=0$ drops quickly, owing to a softening of the ZB phase on alloying with Tl.

Fortunately, the most important feature of Fig. 1 does not depend in any major way in the uncertainty in the liquidus and solidus. There exists below the melting point a stable region of the $In_{1-x}Tl_xSb$ alloy in the ZB lattice for Tl concentrations in the range of 5%–15%. On the basis of Table I this is sufficient to narrow the band gap to the 8–12 μm range. Figure 1 suggests why this region may be difficult or impossible to reach using liquid phase epitaxy. If the liquidus is flat for a wide concentration range as depicted in Fig. 1, and if the solidus falls off rapidly (as it almost certainly does), it is extremely difficult to solidify a crystal in the ZB lattice from the liquid phase with any significant amounts of Tl. Indeed, this has been observed experimentally.[10,11] Nevertheless, a stable phase could exist according to these calculations, and using other growth techniques it should be possible to reach it.

To summarize, our calculations show that $In_{1-x}Tl_xSb$ holds much promise as a viable, more robust alternative to $Hg_{1-x}Cd_xTe$ for use in long-wave infrared detectors. A number of electronic and mechanical properties were calculated for InSb, TlSb, and $In_{1-x}Tl_xSb$ to support this argument. Perhaps the most important finding is that a relatively small amount of Tl (approximately 10%) is needed to narrow the band gap to the required size. The III-V $In_{1-x}Tl_xSb$ is more robust than the II-VI $Hg_{1-x}Cd_xTe$ not only by virtue of it being a III-V compound, but also because it is composed predominantly of the lighter, strengthening cation. Equally important, these calculations show that there exists a stable region in the $In_{1-x}Tl_xSb$ plane of the phase diagram for which Tl is miscible in InSb in concentrations up to 15%–20%.

Note in proof: It has recently come to our attention[11] that the binary phase diagram of TlSb is more complicated than the early literature suggests. Nevertheless, we believe our central conclusion remains valid, namely that there will be a substantial stable phase of $In_{1-x}Tl_xSb$ in the ZB structure, since that conclusion mainly derives from the result that InSb strongly favors the ZB phase to the close-packed one, while TlSb only weakly favors a close-packed phase to the ZB.

This work was supported by ONR Contracts N00014-88-C-0096 and N00014-89-K-0132. We wish to thank A. T. Paxton, M. A. Berding, and J. M. MacCallum for helpful suggestions.

[1] A. Sher, M. A. Berding, M. van Schilfgaarde, and A. B. Chen, Semicond. Sci. Technol. 6, C59 (1991).
[2] R. Balcerak, Semcond. Sci. Technol. 6, C1 (1991).
[3] G. C. Osbourn, L. R. Dawson, R. M. Biefeld, T. E. Zipperian, I. J. Fritz, and B. L. Doyle, J. Vac. Sci. Technol. A 5, 3150 (1987).
[4] B. J. Levine, C. G. Bethea, K. G. Glogovsky, J. W. Stayt, and R. E. Leibenguth, Semicond. Sci. Technol. 6, C114 (1991).
[5] M. A. Berding, A. Sher, A. B. Chen, and W. E. Miller, J. Appl. Phys. 63, 107 (1988).
[6] J. Zelko and J. E. Greene, Appl. Phys. Lett. 33, 254 (1978); J. Appl. Phys. 51, 1549 (1980).
[7] M. Methfessel and M. van Schilfgaarde (unpublished).
[8] S.-H. Wei, L. G. Ferreira, J. E. Bernard, and A. Zunger, Phys. Rev. B 42, 9622 (1990).
[9] S. B. Evgen'ev, O. V. Sorokina, and V. V. Zinov'ev Inorg. Mater. 21, 1747 (1985). O. V. Sorokina, S. B. Sorokina, and I. V. Maschenko, Russ. J. Inorg. Chem. 28, 916 (1983).
[10] V. I. Fedorov and V. I. Machnev, Teplofizika Vysokikh, Temperature 8, 447 (1970).
[11] B. Predel and W. Schwermann, sl Z. Naturforschung A 25, 877 (1970).